National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”
Faculty of Informatics and Computer Science
Department of Computer Engineering

Security, Fault Tolerance, Intelligence

International Conference ICSFTI2018
May 10−12, 2018

Kyiv
Ukraine
Igor Sikorsky Kyiv Polytechnic Institute
2018

UDC 004.056+004.052.3+004.8](062)
T44

Recommended by the Academic Council of FICS
Igor Sikorsky Kyiv Polytechnic Institute
(Protocol No 11 dated May 30, 2018)

Security, Fault Tolerance, Intelligence: proceedings of the International Conference ICSFTI2018, Kyiv, Ukraine, may 10–12, 2018. – Kyiv : Igor Sikorsky Kyiv Polytechnic Institute, publisning house "Polytechnica", 2018. – 415 p.

Program committee
Chairman:
- M. Z. Zgurovsky, Professor, Rector of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Ukraine, Kyiv)
- Y. I. Yakymenko, Professor, First Vice-Rector of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Ukraine, Kyiv)

Vice-chairman:
- S. G. Stirenko, Professor, Head of the Department of Computing Engineering (Ukraine, Kyiv)
- H. M. Loutskii, Professor at the Department of Computing Engineering (Ukraine, Kyiv)

Committee members:
- O. A. Pavlov, Professor, Head of the Department of Computer-Aided Management And Data Processing Systems (Ukraine, Kyiv)
- I. A. Dychka, Professor, Dean of Faculty of Applied Mathematics (Ukraine, Kyiv)
- V.P. Tarasenko, Professor, Head of the Department of System Programming and Specialized Computer Systems (Ukraine, Kyiv)
- Y. O. Kulakov, Professor at the Department of Computing Engineering (Ukraine, Kyiv)
- O. V. Buzovsky, Professor at the Department of Computing Engineering (Ukraine, Kyiv)
- V. I. Zhabin, Professor at the Department of Computing Engineering (Ukraine, Kyiv)
- V. P. Simonenko, Professor at the Department of Computing Engineering (Ukraine, Kyiv)
- M. A. Novotarskyi, Professor at the Department of Computing Engineering (Ukraine, Kyiv)
- Y. H. Gordienko, Professor at the Department of Computing Engineering(Ukraine, Kyiv)
- V. Y. Mukhin, Professor at the Department of Mathematical Methods of System Analysis (Ukraine, Kyiv)
- Z. B. Hu, professor at the Central China Normal University (China) Vu Duc Thinh, Associate Professor at the Ho Chi Minh City University of Food Industry (Vietnam)

Organizational committee
Chairman
- A. M. Volokyta, Associate Professor at the Department of Computing Engineering (Ukraine, Kyiv)

Committee members:
- I. M. Vynogradov, Senior Lecturer at the Department of Computing Engineering(Ukraine, Kyiv)
- Y. I. Kornaga, Ph.D., Associate Professor at the Department of Technical Cybernetics(Ukraine, Kyiv)
- V. M. Shymkovych, Assistant at the Department of Automation and Control in Technical Systems (Ukraine, Kyiv)
- V. V. Steshyn, Assistant at the Department of Computer Engineering (Ukraine, Kyiv)
- P. G. Rehida, Assistant at the Department of Computer Engineering (Ukraine, Kyiv)

The collection presents the works of the scientific conference in the field of computer engineering, software engineering and technical education. The conference is devoted to the memory of the outstanding scientist, professor of the department of computer science prof. V. Shyrochyn.

For teachers, graduate students, students in the field of informatics and computer science.

ISBN 978-966-622-806-5

© Igor Sikorsky Kyiv Polytechnic Institute, 2018
© authors of papers, 2018

http://comsys.kpi.ua

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Факультет інформатики та обчислювальної техніки
Кафедра обчислювальної техніки

Безпека. Відмовостійкість. Інтелект

Міжнародна науково-практична конференція ICSFTI2018

10−12 травня 2018 р.

Київ
Україна
КПІ ім. Ігоря Сікоського
2018

УДК 004.056+004.052.3+004.8](062)
Б40

Рекомендовано Вченою радою ФІОТ КПІ ім. Ігоря Сікорського (Протокол № 11 від 30.05.2018 р.)

Безпека. Відмовостійкість. Інтелект: зб. пр. міжнародної наук.-практ. конф. ICSFTI2018. Київ, Україна, 10–12 трав. 2018 р. – Київ : КПІ ім. Ігоря Сікорського, Вид-во «Політехніка», 2018. – 415 с.

Програмний комітет

Почесні голови:

- Згуровський М. З., ректор КПІ ім. Ігоря Сікорського
- Якименко Ю. І., перший проректор КПІ ім. Ігоря Сікорського

Голови:

- Стіренко С. Г., завідувач кафедри обчислювальної техніки
- Луцький Г. М., почесний науковий керівник кафедри обчислювальної техніки

Члени комітету:

- Павлов О. А., завідувач кафедри автоматизованих систем обробки інформації і управління
- Дичка І. А., декан факультету прикладної математики
- Тарасенко В. П., завідувач кафедри системного програмування та спеціалізованих комп'ютерних систем
- Кулаков Ю. О., професор кафедри обчислювальної техніки
- Бузовський О. В., професор кафедри обчислювальної техніки
- Жабін В. І., професор кафедри обчислювальної техніки
- Симоненко В. П., професор кафедри обчислювальної техніки
- Новотарський М. А., професор кафедри обчислювальної техніки
- Гордієнко Ю. Г., професор кафедри обчислювальної техніки
- Мухін В. Є., професор кафедри математичних методів системного аналізу
- Z. B. Hu, професор Central China Normal University (China)
- Vu Duc Thinh, доцент Ho Chi Minh City University of Food Industry (Vietnam)

Організаційний комітет

Голова

- Волокита А. М., доцент кафедри обчислювальної техніки

Члени комітету:

- Виноградов Ю. М., заступник декана факультету інформатики та обчислювальної техніки по роботі з іноземними студентами
- Корнага Я. І., доцент кафедри технічної кібернетики
- Шимкович В. М., асистент кафедри автоматики та управління в технічних системах
- Стешин В. В., асистент кафедри обчислювальної техніки
- Регіда П. Г., асистент кафедри обчислювальної техніки

Подано праці міжнародної наукової конференції з комп'ютерної інженерії, інженерії програмного забезпечення та технічної освіти, яка присвячена видатному вченому, викладачу кафедри обчислювальної техніки професору В. П. Широчину.

Для викладачів, аспірантів, студентів галузі інформатики та обчислювальної техніки.

Представлены труды международной научной конференции по компьютерной инженерии, инженерии программного обеспечения и технического образования, посвященной выдающемуся ученому, преподавателю кафедры вычислительной техники профессору В. П. Широчину.

Для преподавателей, аспирантов, студентов в области информатики и вычислительной техники.

ISBN 978-966-622-806-5

© КПІ ім. Ігоря Сікорського, 2018
© автори статей, 2018





Пленарна секція.

Семён Широчин.
ФИЛОСОФИЯ ЗНАНИЙ В НАУЧНОЙ И ПЕДАГОГИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ В. П. ШИРОЧИНА 10

Новотарський М., Гулько А.
ГЕНЕРАТИВНІ ЗМАГАЛЬНІ МЕРЕЖІ ТА ЇХ ЗАСТОСУВАННЯ 15

Обозний Дмитро, Поштацька Катерина, Артем Волокита, Георгій Луцький
СТУДЕНТСЬКА ЕЛЕКТРОННА ПОШТА ТА СХОВИЩЕ ДАНИХ НА ОСНОВІ ХМАРНИХ ТЕХНОЛОГІЙ 21

Секція 1.SEC
(Безпека комп'ютерних систем та мереж. Відмовостійкі розподілені обчислення..)

Марковський О. П., Сербін О. Д.
АЛГОРИТМ ЦИФРОВОГО ПІДПИСУ З ВИКОРИСТАННЯМ ПОЛІНОМІАЛЬНОЇ АРИФМЕТИКИ НА ПОЛЯХ ГАЛУА 26

Марич Володимир, Волокита Артем
СПОСІБ ПЛАНУВАННЯ ПЕРЕВІРКИ РЕЗУЛЬТАТІВ ВОЛОНТЕРСЬКИХ ОБЧИСЛЕНЬ ДЛЯ ЗАДАЧ РЕАЛЬНОГО ЧАСУ 35

Інна Стеценко, Дмитро Огородников
МОДЕЛЮВАННЯ СЦЕНАРІЮ КІБЕРАТАКИ НА ОСНОВІ АТАКИ «PETYA», «NOTPETYA» 42

Анатоліївна О. Т., Борисович О. К.
ЗАГРОЗИ БЕЗПЕКИ В ХМАРНИХ СИСТЕМАХ ТА ШЛЯХИ ЇХ УНИКНЕННЯ 52

Троценко В. В., Остапченко К. Б.
ЗАХОДИ БЕЗПЕКИ В СИСТЕМАХ З ІНТЕРНЕТУ РЕЧЕЙ 57

Подтьопа Сергій
ПОМИЛКИ ПРИ РОБОТІ З WEB-СИСТЕМАМИ АВТЕНТИФІКАЦІЇ 63

Валерий Симоненко, Александр Слюсаренко
СИСТЕМА ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ВЫЧИСЛИТЕЛЬНОГО УЗЛА НА ОСНОВЕ РЕЖИМА АУДИТ 69

Тизунь Віталій
ВРАЗЛИВОСТІ БЕЗПЕКИ ТА КОНФІДЕНЦІЙНОСТІ ІНФОРМАЦІЇ ПРИ ВИКОРИСТАННІ ПОПУЛЯРНИХ JAVASCRIPT БІБЛІОТЕК 73

Олександр Марковський, Аббасі Шагін, Вікторія Максимук
ЕФЕКТИВНА ОРГАНІЗАЦІЯ БАГАТОРІВНЕВОЇ ХЕШ-ПАМ'ЯТІ 77

Марковский А.П., Искаков Эмиль Русланович., Гарасимович Г.В.
КОМБИНАТОРНЫЙ АНАЛИЗБУЛЕВЫХ ФУНКЦИЙ, СПЕЦИАЛЬНЫХ КЛАССОВ ДЛЯ СИСТЕМ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ 83

Artem Volokyta, Artem Kaplunov, Heorhii Loutskii
METHOD OF EVALUATING THE EFFECTIVENESS OF DISTRIBUTED COMPUTING SYSTEM TOOLS MONITORING 89

Оксана Кучмий, Виктор Корнейчук, Анна Дорошенко
НАДЕЖНОСТЬ ЭЛЕКТРОСНАБЖЕНИЯ
КОМПЬЮТЕРНЫХ СИСТЕМ И МЕТОДЫ ЕЕ ПОВЫШЕНИЯ 94

Секція 2.RT (Інтернет речей, системи реального часу).

Анатолій Сергієнко, Кадір Сафван Хусейн, Анастасія Сергієнко
ФІЛЬТРИ ЗІ СКІНЧЕННОЮ ХАРАКТЕРИСТИКОЮ
З МІНІМІЗОВАНИМИ АПАРАТНИМИ ВИТРАТАМИ..... 99

Анатолій Сергієнко, Хасан Мухамед Джамал, Павло Сергієнко
АЛГОРИТМ І СТРУКТУРА МОДУЛЯ
ДЛЯ ОБЧИСЛЕННЯ КВАДРАТНОГО КОРЕНЯ У ПЛІС..... 104

Анатолій Сергієнко, Мустафа Рекар Касім
АПАРАТНЕ МОДЕЛЮВАННЯ ПОШИРЕННЯ
УЛЬТРАЗВУКОВИХ ХВИЛЬ У ТВЕРДОМУ ТІЛІ 109

Демчук Дмитро, Валерій Сімоненко
РОЗРОБКА СИСТЕМИ АВТОМАТИЧНОГО ПРОГНОЗУВАННЯ
ЗРОШЕНЬ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР
З ВИКОРИСТАННЯМ МОДУЛЬНИХ ПРИСТРОЇВ 114

Даніленко Наталія, Сімоненко Андрій
МОДЕЛЬ ПІДВОДНОЇ ЖИТТЄЗДАТНОЇ СИСТЕМИ НА ОСНОВІ
МІКРОКОНТРОЛЕРА ATMEGA328 ПЛАТФОРМИ ARDUINO UNO..... 119

Владислав Панійван
ПОБУДОВА МОДЕЛІ СИСТЕМИ МАСОВОГО
ОБСЛУГОВУВАННЯ ДЛЯ ВИДАННЯ
РЕКОМЕНДАЦІЙ ЩОДО ОБРАННЯ ЧЕРГ.....128

Володимир Шимкович, Юрій Бердник
РЕАЛІЗАЦІЯ ТА ДОСЛІДЖЕННЯ НЕЙРОННИХ
МЕРЕЖ ПРЯМОГО РОЗПОВСЮДЖЕННЯ НА ПЛІС.....134

Сергій Бугай, Юрій Гордієнко
ОПТИМІЗАЦІЯ АЛГОРИТМУ ВИЯВЛЕННЯ ПАДІННЯ
НА ОСНОВІ АКСЕЛЕРОМЕТРА ТА ГІРОСКОПА.....141

Сергій Бугай, Юрій Гордієнко
СПОСІБ ІНТЕГРАЦІЇ РОЗПОДІЛЕНИХ
ОБЧИСЛЕНЬ НА БАЗІ НОСИМОЇ ЕЛЕКТРОНІКИ.....148

Киричек М. І., Волокита А. М.
ПОЛІПШЕННЯ ВХОДЖЕННЯ В МЕДИТАТИВНИЙ
СТАН ЗА ДОПОМОГОЮ МОБІЛЬНОГО ЗАСТОСУНКУ
З ЕЛЕМЕНТАМИ УПРАВЛІННЯ СВІТЛОДІОДНИМИ ЛАМПАМИ. 154

Ярослав Корнага, Альона Вітюк
ОЦІНКА ЯКОСТІ СТІЙКОГО ЗАХВАТУ, СПЛАНОВАНОГО
НА ОСНОВІ ТРИВИМІРНОЇ РЕКОНСТРУКЦІЇ ЦІЛЬОВОГО
ОБ'ЄКТУ ПО ЗОБРАЖЕННЯМ З МОНОКАМЕРИ 159

Олексій Гранюк, Юрій Гордієнко
ПИТАННЯ ПЕРЕДАЧІ ДАНИХ ВІД СЕНСОРІВ
НА МОБІЛЬНИЙ ТЕЛЕФОН КОРИСТУВАЧА
В КОНТЕКСТІ ІНТЕРНЕТУ РЕЧЕЙ.....165

Марковський О.П., Ван Хуаі, Гарасимович Г. В.
ОРГАНІЗАЦІЯ ЗАХИЩЕНОГО ОБЧИСЛЕННЯ МОДУЛЯРНОЇ
ЕКСПОНЕНТИ З ВИКОРИСТАННЯМ АДИТИВНОГО МАСКУВАННЯ 171

Гарасимович Г. В., Куц В. Ю.
МЕТОД ЗАХИСТУ ОПЕРАНДІВ МОДУЛЯРНОГО
ЕКСПОНЕНЦІЮВАННЯ ВІД ЇХ РЕКОНСТРУКЦІЇ
АНАЛІЗОМ ДИНАМІКИ СПОЖИВАННЯ ПОТУЖНОСТІ 177

Микола Серпученко, Ірина Клименко
ОГЛЯД СУЧАСНИХ ТЕХНОЛОГІЙ СТВОРЕННЯ
ЦИФРОВИХ ПРИСТРОЇВ НА БАЗІ МІКРОПРОЦЕСОРНИХ СИСТЕМ *184*

Соляніков Павло
ДОСВІД ЗАПУСКУ РОБОТИЗОВАНОГО
ЗВАРЮВАЛЬНОГО КОМПЛЕКСУ 194

В. І. Жабін, А. Г. Токар
АНАЛІЗ ЕФЕКТИВНОСТІ РЕАЛІЗАЦІЇ
ДОДАВАННЯ В НЕАВТОНОМНОМУ РЕЖИМІ 201

**Секція 3. AI (Інтелектуальні системи.
Машинне навчання, bigdata).**

Юрій Тимошин, Сергій Орленко
ВПЛИВ ПЛЯМИ СПРИЙНЯТТЯ НА ШВИДКІСТЬ
ТА ТОЧНІСТЬ РОЗПІЗНАВАННЯ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ 209

Пилипенко Павло, Пантелєєв Олександр, Олійник Володимир
МЕТОД ВИРІШЕННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ ЛЮДЕЙ
В РЕАЛЬНОМУ ЧАСІ В УМОВАХ УСКЛАДНЕНОЇ ВИДИМОСТІ 215

Дмитро Смішний, Олексій Алещенко
РОЗПІЗНАВАННЯ МОВИ З ВИКОРИСТАННЯМ
ТЕХНОЛОГІЇ ГЛИБОКОГО НАВЧАННЯ 220

Денис Смішний, Олексій Алещенко
ОБРОБКА ВІЗУАЛЬНИХ ДАНИХ
ЗА ДОПОМОГОЮ ТЕХНОЛОГІЇ COMPUTER VISION.....225

Пантелєєв Олександр, Пилипенко Павло, Олійник Володимир
СИСТЕМА ВІЗУАЛЬНОГО МУЛЬТІТРЕКІНГА В РЕАЛЬНОМУ ЧАСІ 230

Богдан Бугай, Юрій Гордієнко
МЕТОД ОПТИМІЗОВАНОГО ЗНАХОДЖЕННЯ
ПОДІБНИХ БІНАРНИХ РЯДКІВ У ПЕВНІЙ МНОЖИНІ 235

Богдан Бугай, Юрій Гордієнко
МЕТОД ВИМІРЮВАННЯ ПОДІБНОСТІ ВІДЕОМАТЕРІАЛІВ 239

Болдак А. О., Лисенко Д. В., Ханін М. Є.
СПОСІБ ОПРАЦЮВАННЯ ПРОПУСКІВ
В РЕЗУЛЬТАТАХ ЕКСПЕРТНИХ ОПИТУВАНЬ 246

Андрій Логвинчук, Олексій Алещенко, Олександр Зальотнов
ЗАСТОСУВАННЯ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ
ДО МОДЕЛЮВАННЯ ЕЛЕКТРИЧНИХ КІЛ ЗМІННОГО СТРУМУ 251

Данило Сергеєв
ІНТЕГРАЦІЙНА МОДЕЛЬ ПРЕДСТАВЛЕННЯ
ЗНАНЬ ДЛЯ СЕМАНТИЧНОГО WEB 256

Ярослав Корнага, Микола Шиманський, Андрій Барабаш
ЛОКАЛІЗАЦІЯ ТЕКСТУ НА ЗОБРАЖЕННЯХ
ЗА ДОПОМОГОЮ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ.....261

Ярослав Корнага, Юрій Тільняк
ДІАГНОСТИКА ТЕХНІЧНОГО СТАНУ АВТОМОБІЛІВ
З ВИКОРИСТАННЯМ ШТУЧНИХ НЕЙРОННОЇ МЕРЕЖІ 265

Вадим Мухін, Ярослав Корнага, Андрій Любівий
МОДЕЛЮВАННЯ Й РОЗПІЗНАВАННЯ ЖЕСТІВ СКЕЛЕТОНУ РУКИ 270

Ярослав Подзірей, Павло Регіда
КЛАСТЕРИЗАЦІЯ. ЗМЕНШЕННЯ РОЗМІРНОСТІ ОЗНАК ОБ'ЄКТІВ В
МАШИННОМУ НАВЧАННІ ТА ВІЗУАЛІЗАЦІЯ ДАНИХ. МЕТОД t-SNE. 276

Вадим Щур, Віктор Стешин
ВИКОРИСТАННЯ МЕТОДІВ РОЗПІЗНАВАННЯ СИМВОЛІВ
ДЛЯ ВИЗНАЧЕННЯ ДОСТОВІРНОСТІ ОСОБИСТОГО ПІДПИСУ 281

Коломієць Олег Олексійович
МЕТОДИ ІНТЕГРАЦІЇ ЕЛЕМЕНТІВ ДОПОВНЕНОЇ
РЕАЛЬНОСТІ З СИСТЕМОЮ РОЗПІЗНАВАННЯ ОБРАЗІВ 286

Віктор Порєв
ЕЛЕМЕНТИ ІНТЕЛЕКТУАЛЬНОЇ ПРОГРАМНОЇ
СИСТЕМИ ДЛЯ ПІДТРИМКИ НАВЧАЛЬНОГО ПРОЦЕСУ 291

Секція 4. GN (Глобальні мережі, grid та cloud системи).

Ярич Андрій
ПИТАННЯ ВИБОРУ CLOUD-ДОДАТКУ
НА ОСНОВІ АНАЛІЗУ GOOGLE, AMAZON ТА APPLE РІШЕНЬ.....299

Максим Бондарчук, Юрій Зорін
АЛГОРИТМ РОЮ ЧАСТИНОК ПОШУКУ
НАЙКОРОТШОГО ШЛЯХУ В ТЕЛЕКОМУНІКАЦІЙНІЙ МЕРЕЖІ 304

Олександр Калюжний, Юрій Кулаков, Михайло Діброва
СПОСІБ ОРГАНІЗАЦІЇ БАГАТОШЛЯХОВОЇ МАРШРУТИЗАЦІЇ
З ВИКОРИСТАННЯМ ЧАСТКОВО-ПЕРЕСІЧНИХ ШЛЯХІВ.....310

Бугай Андрій, Олійник Володимир
МОДЕЛЬ ПОБУДОВИ ІНТЕРФЕЙСУ КОРИСТУВАЧА 315
З ВИКОРИСТАННЯМ МІКРОСЕРВІСНОЇ АРХІТЕКТУРИ.....315

Кулаков Юрій, Шудренко Євгеній, Баринін Юрій
МОДИФІКАЦІЯ АЛГОРИТМУ
ЕНЕРГОЗБЕРІГАЮЧОГО УПРАВЛІННЯ ТРАФІКОМ
В ГІБРИДНИХ МАГІСТРАЛЬНИХ SDN/IP МЕРЕЖАХ.....319

Виноградов Ю. М., Байдиков М. О.
ПИТАННЯ ЯКОСТІ МОБІЛЬНИХ
ДОДАТКІВ НА ОСНОВІ ЇХ ТЕСТУВАННЯ 326

Демчик Валерій, Корочкін Олександр
ЗАСТОСУВАННЯ ДРІБНОЗЕРНИСТОГО
ПАРАЛЕЛІЗМУ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ
ПАРАЛЕЛЬНИХ ТА РОЗПОДІЛЕНИХ ОБЧИСЛЕНЬ 330

Юрій Кулаков, Євгеній Лопушен, Берест Руслана
СПОСІБ МОНІТОРИНГУ ОБЧИСЛЮВАЛЬНИХ
РЕСУРСІВ В ПРОГРАМНО-КОНФІГУРОВАНИХ МЕРЕЖАХ.....337

Ріпневський Олександр
ВИКОРИСТАННЯ СУБД ІЗ АРХІТЕКТУРОЮ IN-MEMORY DATA GRID 342

Олександр Марковський, Вікторія Максимук
МЕТОД СТРОГОЇ АВТЕНТИФІКАЦІЇ ВІДДАЛЕНИХ
АБОНЕНТІВ НА ОСНОВІ АЛГЕБРИ ПОЛІВ ГАЛУА.....347

Марковський О. П., Білашевська А. В.
МЕТОД КОРЕКЦІЇ БАГАТОКРАТНИХ ПОМИЛОК В КАНАЛАХ
ПЕРЕДАЧІ ЦИФРОВИХ ДАНИХ ЗІ СПЕКТРАЛЬНОЮ МОДУЛЯЦІЄЮ.....355

Л. Д. Міщенко, Ю. М. Виноградов
МЕТОД ПАРАЛЕЛЬНОГО ОБЧИСЛЕННЯ МОДУЛЯРНОЇ
ЕКСПОНЕНТИ З ВИКОРИСТАННЯМ ПЕРЕДОБЧИСЛЕНЬ 365

Елена Жданова, Алла Коган, Юрий Кулаков, Майя Сперкач
О РАЗБИЕНИИ ГРАФА НА ПОДГРАФЫ 372

Пленарна секція.
Plenary Section.

УДК 001.891.34

Семён Широчин.

ФИЛОСОФИЯ ЗНАНИЙ В НАУЧНОЙ И ПЕДАГОГИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ В. П. ШИРОЧИНА

Аннотация. В статье тезисно анализируется научно-педагогическая деятельность В. П. Широчина в контексте универсальности философии знаний, представленной в его творчестве.

Ключевые слова: интеллект, педагогика, психология, познание

Abstract. The article briefly analyzes the scientific and pedagogical activity of V.P. Shyrochyn in the context of universality of the philosophy of knowledge presented in his works.

Keywords:intellect, pedagogy, psychology, cognition

Валерий Павлович Широчин (1939-2016) всю жизнь был учёным и педагогом. Будучи сыном военного инженера связи, он с детства имел интерес к науке и технике. Это и определило его судьбу – практически всю свою жизнь он посвятил работе в КПИ. Именно здесь он полвека преподавал, обучив тысячи студентов, подготавливал аспирантов, многие из которых стали кандидатами и докторами наук, заняв научные позиции в различных странах мира, и именно здесь он занимался своей научной деятельностью, работая в учёных советах, оппонируя на защитах диссертаций, а также – создавая своё научное видение, материализованное в книгах, которые ещё предстоит оценить будущим поколениям учёных.

Будучи учёным в ряде различных сфер знаний, Валерий Павлович имел уникальный научный опыт, в котором структурированная совокупность знаний сформировалась в определённую философию знаний. Это означало не только, что Валерий Павлович мог поддержать разговор на любую тему, что делало его интересным собеседником, но и, имея структурное системное видение знаний в различных сферах, обнаруживать аналогии и доступно их описывать. Так, к примеру, во время лекций по техническим предметам Валерий Павлович часто использовал доступные и, что особенно важно, запоминающиеся аналогии объясняемых явлений в живом мире и поведении человека. Это позволяло не только представить информацию в максимально доступной слушателю форме, но и передать сам принцип универсальности закономерностей, заинтересовав слушателя философией знаний.

Важно вспомнить и о форме передачи знаний. Лекции, книги, стихи, обычные разговоры – все возможности коммуникаций Валерий Павлович всегда использовал для того, чтобы делиться знаниями. Проблема распространения научных знаний заключается в том, что идея, выраженная в научной терминологии и использующая принятые в научном стиле сугубо наукообразные формулировки, часто становится доступной лишь узкому кругу специалистов. Валерий Павлович стремился донести свои идеи максимально широкому кругу лиц. И тут нам стоит

вспомнить Альберта Эйнштейна, который говорил: «Всё должно быть изложено так просто, как только возможно, но не проще». Поэтому параллельно научным текстам у Валерия Павловича научные и философские идеи часто формулируются художественно, просто и доступно – в удобоваримой разговорной или же стихотворной форме.

Отдельное место здесь занимает сборник стихов Zodiac+, изданный Валерием Павловичем ещё в начале 1990-х. Сборник интересен тем, что стихи – это неформальный способ ёмко выразить обнаруженные в окружающем мире закономерности. В том числе, в сборнике Zodiac+ описана сама потребность в познании различных сфер знаний, как источнике мудрости:

«*Каким бы ни был быт наш прозаичным*
И ни сложился ритм труда привычным
Всегда мы рвёмся к свету и приоде
Чтоб приобщиться к мудрости в народе»

[1, с.6]

Знания как сила и потребность. Такой огромный научный и педагогический опыт был бы невозможен без потребности познания и потребности делиться полученными знаниями. Валерий Павлович отмечал, что потребность в познании есть у незначительной части людей. К сожалению, по природе своей далеко не каждому человеку интересно, как устроен мир. И, тем более, ещё меньшему количеству людей интересно тратить своё время и усилия для нахождения в этом мире неизвестных до этого закономерностей. Интерес к знаниям искренний и его сложно подделать. Поэтому те, кому познание интересно, видны сразу. И таких людей Валерий Павлович всегда уважал и поддерживал.

Знание в философии Валерия Павловича – уникальный для каждого человека продукт его жизненного опыта, определяющий не только поведение, но и саму способность воспринять, заметить и даже увидеть. Валерий Павлович понимал последовательность обработки информации в человеческом мозге, где видение – это сознательный процесс, а значит требует предварительной когнитивной обработки, сравнения видимого образа с представляемым и его идентификации. Поскольку знание – продукт опыта, а опыт у всех разный, то и знание, а следовательно – видение у разных людей будет существенно отличаться:

«*Казалось, видят все одно и то же*
Но все, увы, не так! Похоже,
Что каждый видит то, что может
Или таким, как кто-нибудь поможет
***
Все в механизмах зрения дробится
И если диво в страхе нам явится
То видим мы лишь то, что знаем
Или хотя бы как-то представляем»

[1, с.23]

И вот это ключевое «поможет» - о том, что возможности собственного опыта ограничены, поскольку ограничено жизненное время человека. Достичь

большого количества знаний, руководствуясь исключительно личным опытом, невозможно, поэтому необходимо взаимодействовать, помогать друг другу обретать знания и понятийный аппарат. Для этого существуют книги и целая система образования. И стоит понять, что мы смотрим на мир преимущественно чужими глазами, ведь большая часть имеющихся знаний и понятий – это заслуга других людей, от которых мы этому научились.

Валерий Павлович видел силу знаний в способности осознанного принятия решений. Человек боится того, чего не знает. Из чего и выходит рецепт – чтобы победить страх, необходимо увидеть, узнать и сознательно воспринять окружающую действительность. Только так можно принять решения и реагировать:

«*Страх – вечные мотивы в мире странном*
Страх кажется нам темным и спонтанным
Он быстро силы в теле собирает
И реагировать на тени позволяет
***
Все в знаниях и зрении сокрыто
И тайной восприятия покрыто
Ведь если всё вокруг для нас понятно
То страха нет, поскольку жить приятно»

[1, с.23]

Универсальность философии знаний. Валерий Павлович интересовался различными сферами знаний. Помимо основных научных направлений, таких как защита информации, моделирование и искусственный интеллект, он всегда интересовался вопросами организации живого мира и особенно – человека. Изучение поведения последнего упрощалось необходимостью взаимодействия с большим количеством людей – через Валерия Павловича за 50 лет преподавательской деятельности прошли тысячи студентов и аспирантов. А большие объёмы информации всегда способствуют её анализу и поиску зависимостей и закономерностей. Знания о разных сферах бытия неизбежно складываются в единую систему, в единую философию знаний, где уже нет границ между предметами и темами. Анализ большого количества информации неизбежно приводит к нахождению общих закономерностей, характерных технике, законам природы и поведению человека и других разумных существ.

К чему в итоге приводит пересечение подобных научных интересов? К новым идеям, требующих глубоких познаний в разных вопросах. К примеру, работа над созданием средств защиты компьютерных систем в сочетании с познаниями в организации иммунитета живыми организмами приводит к созданию сложных биологически инспирированных систем защиты информации. А работа в области искусственного интеллекта и моделирования в сочетании с познаниями в области мышления человека и вовсе позволяет представить последнего, как суперкомпьютер. Именно это сравнение положено в основу многих работ, в частности – книги «**Слово об интеллекте. Концептуальные основы системной психологии**».

Интеллект, справедливости ради, присутствует и в более раннем творчестве Валерия Павловича, в том же Zodiac+, где подана попытка объяснить суть интеллекта в доступной среднему читателю форме:

«*Кто волю не сменит виденьем*
И удивит нас всех решеньем
В том точно сложен интеллект
Как всех способностей комплект»

[1,с.21]

Валерий Павлович использует техническую терминологию и визуальные схематические принципы при описании процессов, казалось бы, совершенно не технических. Поэтому для учёных, не имеющих схожего опыта познания и сфер компетентности, подобный подход кажется эклектикой, и не всегда встречает понимание. Что, опять таки, описано в Zodiac+ - «*видим мы лишь то, что знаем*». Но нужно отметить, что именно такой подход позволяет людям с техническим мышлением воспринять универсальность философии знаний, понять, что человек устроен, как суперкомпьютер, со своими сенсорными и аналитическими центрами, со сложнейшими мыслительными и когнитивными процессами, лежащими в основе столь сложных понятий, как эмоции и речь.

Актуальность подобного подхода подтверждается существованием международных конференций, в которых Валерий Павлович принимал участие, в частности, направления BICA – биологически инспирированных когнитивных архитектур, где проложена прямая зависимость между анализом алгоритмов мышления живых существ и создаваемым человеком искусственным интеллектом.

Доказанной целью человеческой памяти является систематизация знаний с целью возможности прогнозирования будущего. И актуально будет вспомнить, что все исследования Валерия Павловича направлены именно в будущее. А это значит, что в ближайшем будущем мы станем свидетелями развития совершенно новых интеллектуальных компьютерных систем, оперирующих понятиями эмоций и доверия.

Литература

1. В.П. Широчин. *ZodiaC+. О вере, старданиях и любви (стихи)*.1993, Киев, Диасофт, 64с.

ВЕДОМОСТИ ОБ АВТОРЕ

Широчин Семён Станиславович – кандидат технических наук.
Email: semenstsh@gmail.com

SemenShyrochyn

PHILOSOPHY OF KNOWLEDGE IN SCIENTIFIC AND PEDAGOGICAL ACTIVITY OF V.P. SHYROCHYN

Holding a conference of Valery Shyrochyn's memory emphasizes importance of researching of his approach to knowledge and their formation.

In Valery Shyrochyn's works knowledge is the most important tool for conscious perception of the world around us. Knowledge basecreatesthe platform for the conscious analysis of information from sensory systems, for understanding imagesand recognizing patterns. Knowledge base creates the power of the ability of conscious decision-making. Knowledge allows you to get rid of fear, because a person is afraid of what he does not know. The accumulation and transfer of knowledge is the most important task of any man.

The set of knowledge for each person is unique, as well as his life experience, so it determines not only the behavior, but also the very ability to perceive, notice and distinguish information from sensory systems.

The accumulation of a large amount of information boosts opportunities of its analysis and search for dependencies and regularities. The knowledge in different spheres of life is transformed to the unifying system, where boundaries between disciplines disappear. Analysis of a large amount of information inevitably leads to finding common patterns, characteristic techniques, laws of nature and human behavior.

In his works, Valery Shyrochyn simulates the work of the human intellect, presenting it in the form of a supercomputer design.That leads to models of the work of various components of human thinking, presented in technical and engineering terms.Future trends in the development of intellectual systems will inevitably borrow the mechanisms existing in living organisms.The relevance of this approach is confirmed by the existence of international conferences on this topic.

Keywords: intellect, pedagogy, psychology, cognition

УДК 004.032.26

Новотарський М.,
Гулько А.

ГЕНЕРАТИВНІ ЗМАГАЛЬНІ МЕРЕЖІ ТА ЇХ ЗАСТОСУВАННЯ

У роботі розглянуто основні тенденції розвитку сучасних генеративних змагальних мереж (ГЗМ) та напрями їх практичного застосування. Наведено деякі особливості створення таких мереж для покращення якості зображень, які отримують при рентгеноскопії та комп'ютерній томографії. За рахунок модифікації функції втрат, яка враховує змагальну складову, вирішується задача покращення чіткості зображення за умови збереження важливих деталей.

Ключові слова: генеративна змагальна нейронна мережа, дискримінатор, генератор, функція втрат.
Рис.: 3. Бібл.: 11.

The paper is devoted to consideration of the main tendencies of development of modern generative adversarial networks (GAN) and directions of their practical applications. Some featuresof creating such networks for improving the quality of images obtained by X-ray and computer tomography are presented. Due to the modification of the loss function due to the consideration of the competition component, the problem of improving the clarity of the image is preserved with the preservation of important details.

Keywords: Generative Adversarial Neural Network, Discriminator, Generator, Loss Function.
Fig.: 3. Bibl.: 11.

Актуальність теми дослідження. Принцип змагальності, який використовують угенеративних змагальних нейронних мережах, дозволив досягти відчутного прогресу в сфері застосування штучних нейронних мереж. Актуальною темою зокрема є покращення чіткості зображень у медицині, оскільки дозволяє отримати вичерпну інформацію про стан організму при зменшенні шкідливого фізичного впливу на нього.

Постановка проблеми. Основна проблема при покращенні чіткості зображень полягає у тому, що для вирішення поставленої задачі в зображення необхідно додати певну інформацію без істотного спотворення початкового зображення. Вирішення такої проблеми потребує розробки механізмів генерування інформації з високим рівнем правдоподібності.

Аналіз останніх досліджень та публікацій. Сучасні генеративні змагальні мережі є важливою складовою досліджень у сфері генеративних моделей. В рамках даної роботи представлені моделі, на вхід яких подають тренувальний набір, сформований із зразків зі щільністю правдоподібності p_dта які тренуються з метою визначення оцінки правдоподібності результату p_m, одержаного відповідно до постановки задачі [1]. Сучасні ГЗМ, які використовуються в задачах класифікації, мають на меті безпосереднє визначення оцінки p_m. Особливо в останні роки розроблено велику кількість ГЗМ, які здатні безпосередньо генерувати зображення з використанням тренувальних наборів або без них.

Актуальність застосування ГЗМ зростає з кожним днем, що відображається у вибуховій кількості публікацій та повідомлень про нові розробки в даній сфері [2, 3, 4].

ГЗМ застосовують для передбачення подій у майбутньому шляхом врахування поточного стану навколишнього середовища та можливих активних дій об'єктів [5]. Особливістю ГЗМ є те, що вони можуть навчатися передбаченню за відсутності вхідних даних або при істотно меншій їх кількості, що є перевагою у порівнянні з мережами глибокого навчання [6, 7]. Важливою сферою застосування ГЗМ є також задачі, мультимодальними результатами у яких є велика кількість правильних відповідей. Прикладом такого застосування може бути мережа для передбачення наступного кадру на відео [8].Нарешті, багато задач вимагають реального створення зразків з деякого розподілу. Серед них, наприклад, задачі перетворення фотографії місцевості в карту [9] та перетворення ескізу в реалістичне зображення [10]. В даній роботі наведенозастосування ГЗМ до підвищення роздільної здатності зображень, яка також використовує властивість даного класу мереж до створення нової інформації з високим рівнем правдоподібності.

Структурні характеристики генеративних змагальних мереж. Всі ГЗМ включають два основних вузли: генератор G та дискримінаторD.

Дії цих вузлів нагадують сценарій поведінки шахрая та поліцейського. Шахрай намагається ввести в оману поліцейського, а завдання поліцейського полягає у викритті дій шахрая. На рис. 1 показана максимально узагальнена структура ГЗМ.

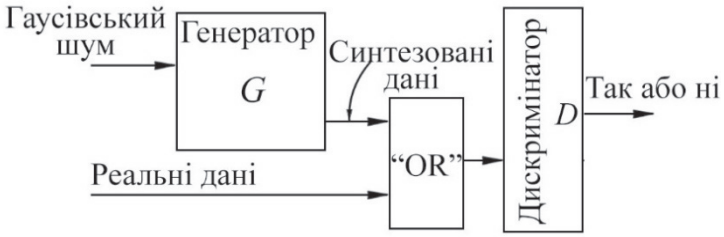


Рис. 1. Базова структура генеративної змагальної мережі

Вузли G та D – це багатошарові мережі, що можуть містити згорткові шари та повністю з'єднані шари зі зсувами. За умови розгляду генераторної мережі G як відображення з деякого простору, представленого прихованими шарами, у вихідні дані, її роботу можна формально задати функцією $G: G(z) \rightarrow R^{|x|}$, де $z \in R^{|z|}$– вектор гаусівського шуму у $|z|$-вимірному просторі прихованого шару, $R^{|x|}$–$|x|$-вимірний простір вхідних даних.

Мережу дискримінатора D також можна характеризувати як функцію, яка відображає дані про зображення на ймовірність того, що зображення походить з реального розподілу даних, а не з розподілу генератора:$D: D(x) \rightarrow (0,1)$. При заданих параметрах генератора G дискримінатор D тренують на розпізнавання реальних зображень (результат близький до 1) або синтезованих зображень (результат близький до 0). Після виконання етапу навчання дискримінатор вважають оптимальним, а генератор G може продовжувати навчатися, щоб знизити точність дискримінатора. Чим більше розподіл генератора відповідатиме

реальному розподілу даних, тим дискримінатору буде важче відрізнити реальний розподіл від синтезованого, і його вихідна оцінка наближатиметься до 0.5 для всіх вхідних даних.

Принципи навчання генеративних змагальних мереж. Тренування ГЗМ включає в себе як пошук параметрів дискримінатора, що максимізує його точність класифікації, так і пошук параметрів генератора, який мінімізує точність дискримінатора. Цей навчальний процес показано на рис. 2.

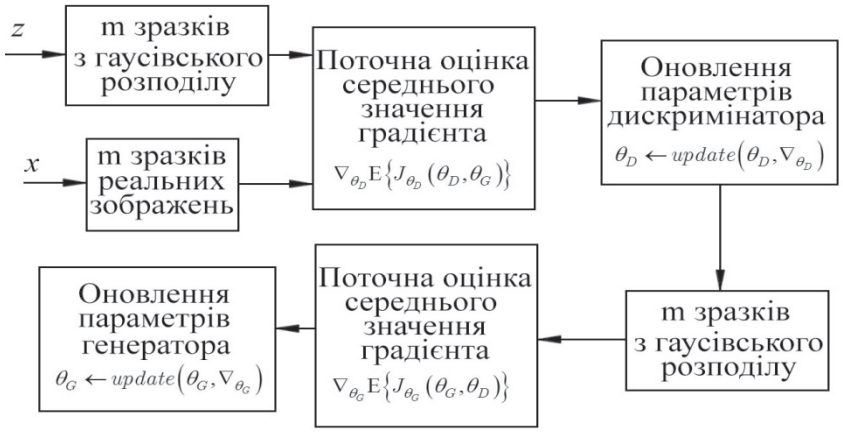


Рис. 2. Блок-схема процесу навчання ГЗМ

Для оцінювання успішності навчання використовують різні метрики, які у загальному випадку можна представити функцією значення $V(G,D)$, модифікація якої відбувається відповідно до виразу: $\max_D \min_G V(G,D)$, де

$$V(G,D) = \mathrm{E}_{p_d(x)} \log\big(D(x)\big) + \mathrm{E}_{p_g(x)} \log\big(1 - D(x)\big). \qquad (1)$$

Під час тренування параметри одного з вузлів оновлюються, а параметри іншого фіксуються. В [11] показано, що для фіксованого генератора існує єдиний оптимальний дискримінатор, значення якого дорівнює:

$$D^*(x) = \frac{p_d(x)}{p_d(x)+p_g(x)}. \qquad (2)$$

В ідеалі для одержання оптимального дискримінатора потрібно було б тренувати дискримінатор після кожного тренування генератора. Але на практиці дискримінатор тренують не так часто, в той час як генератор підлягає тренуванню після кожної відповіді дискримінатора.

При наближенні дискримінатора до рівноважного стану може виникнути проблема втрати генератором його градієнта. Для вирішення цієї проблеми використовують різні евристичні підходи для штучного підвищення градієнта, а також застосовують обчислення $\max_G \log D\big(G(z)\big)$ частіше ніж $\max_G \log\Big(1 - D\big(G(z)\big)\Big)$.

Застосування ГЗМ для підвищення чіткості зображень. Метою запропонованої в даній роботі ГЗМ є одержання оцінки роздільної здатності α вхідного зображення з низькою роздільною здатністю λ за умови, що λ – це версія зображення з низькою роздільною здатністю, яка є аналогом «невідомого» зображення з високою роздільною здатністю γ .

Аутентичні зображення з високою роздільною здатністю γ доступні лише під час тренувань. При тренуванні генератора отримуємо розподіл для λ, застосовуючи гаусівський фільтр до γ . Нашою кінцевою метою є навчання

генераторної функції G, яка створює оцінку для вхідного зображення, відповідну до аутентичного зображення γ . Мети досягають шляхом оптимізації спеціальної функції втрат l_α. Для тренувальних зображень $\gamma_n, n = 1,2,\dots,N$ з відповідними $\lambda_n, n = 1,2,\dots,N$ виконуємо наступні обчислення:

$$\hat{\theta}_G = \arg\min_{\theta_G} \frac{1}{N}\sum_{n=1}^{N} l_\alpha\left(G_{\theta_G}(\lambda_n), \gamma_n\right) \quad (3)$$

Використовуючи підхід, запропонований в [11], визначаємо дискримінаторну мережу D_{θ_D}, яку оптимізуємо з певною періодичністю відносно G_{θ_G} для вирішення змагальної задачі типу мін-макс:

$$\min_{\theta_G}\max_{\theta_D} \mathrm{E}_{\gamma\sim p_{train}(\gamma)} \log D_{\theta_D}(\gamma) + \\ + \mathrm{E}_{\lambda\sim p_G(\lambda)} \log\left(1 - D_{\theta_D}\left(G_{\theta_G}(\lambda)\right)\right) \quad (4)$$

Функція втрат l^α є зваженою сумою змістовної втрати l_x^α та змагальної втрати l_G^α:

$$l^\alpha = l_x^\alpha + 10^{-3} l_G^\alpha, \quad (5)$$

l_x^α – визначають з урахуванням структури нейронної мережі дискримінатора.

Змагальну втрату l_G^α визначаємо з виразу $l_G^\alpha = \sum_{n=1}^{N} -\log D_{\theta_D}\left(G_{\theta_G}(\lambda)\right)$

$$l_G^\alpha = \sum_{n-1}^{N}\left(-\log D_{\theta_D}\left(G_{\theta_G}(\lambda)\right)\right), \quad (6)$$

де λ $D_{\theta_D}\left(G_{\theta_G}(\lambda)\right)$ імовірність того, що реконструйоване зображення $G_{\theta_G}(\lambda)$ відповідає аутентичному зображенню α .

Результати досліджень. Дослідження проводилися з нейронними мережами дискримінатора та генератора, які складались з послідовно розташованих згорткових шарів та шарів з повним набором міжшарових зв'язків. Тренування проводились з використанням чорно-білих зображень, одержаних в результаті проведення рентгеноскопії та комп'ютерної томографії. На рис. 3 показано відтворення чіткого зображення при різних рівнях оптимальності дискримінатора.

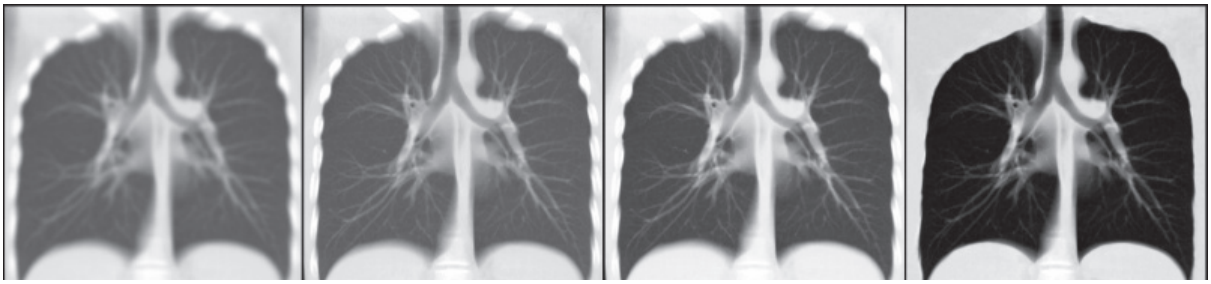


Рис. 3. Відтворення чіткого зображення при комп'ютерній томографії

Висновки. Дослідження в області нових архітектур та способів навчання генеративних змагальних мереж є надзвичайно актуальними, оскільки дозволяють істотно покращити ті результати, що були отримані при застосуванні традиційних штучних нейронних мереж та мереж з глибоким навчанням. В роботі наведено основні напрями сучасного застосування генеративних змагальних мереж, розглянуто їх структурні особливості та принципи навчання. Запропоновано підхід до створення згаданого типу мереж для вирішення задач покращення чіткості зображень, які одержані при рентгеноскопії та комп'ютерній томографії.

Литература

1. Dinh L., Sohl-Dickstein J., Bengio, S. (2016). *Density estimation using real nvp*. arXiv preprint arXiv:1605.08803.

2. Brock A., Lim T., Ritchie JM., Weston N. (2016)*Neural photo editing with introspective adversarial networks*. arXiv preprint arXiv:1609.07093.

3. Arjovsky M., Bottou L. (2017). *Towards principled methods for training generative adversarial networks*. arXiv preprint arXiv:1701.04862.

4. Che T., Li Y., Jacob AP., Bengio Y., Li W. (2016). *Mode regularized generative adversarial networks*. arXiv preprint arXiv:1612.02136, 2016.

5. Finn C. , Levine S. (2016). *Deep visual foresight for planning robot motion*. arXiv preprint arXiv:1610.00696 .

6. Donahue J., Krahenbuhl P., Darrell T. (2016). *Adversarial feature learning*. arXiv preprint arXiv:1605.09782 .

7. Dumoulin V., Belghazi I., Poole B., Lamb A., Arjovsky M., Mastropietro O., Courville A. (2016). *Adversarially learned inference*. arXiv preprint arXiv:1606.00704.

8. Lotter W., Kreiman G., Cox D. (2015). *Unsupervised learning of visual structure using predictive generative networks*. arXiv preprint arXiv:1511.06380.

9. Isola P., Zhu J.-Y., Zhou T., Efros A. A. (2016). *Image-to-image translationwith conditional adversarial networks.* arXiv preprint arXiv:1611.07004.

10. Zhu J.-Y., Krahenbuhl P., Shechtman, E., and Efros A. A. (2016). *Generative visual manipulation on the natural image manifold*. In European Conference on Computer Vision, Springer, (p. 597-613).

11. Goodfellow I. Pouget-Abadie J. Mirza M. Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. (2014) Generative adversarial nets. In Advances in Neural Information Processing Systems, (p.2672-2680).

ДОВІДКА ПРО АВТОРІВ

Новотарський Михайло Анатолійович – професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Novotarskyi Mykhailo – professor Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". E-mail: novot@ukr.net

Гулько Анна Вадимівна – студентка, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Gulko Anna Vadymivna – student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute". E-mail: anuta9612@icloud.com

Михайло Новотарський,
Анна Гулько

ГЕНЕРАТИВНІ ЗМАГАЛЬНІ МЕРЕЖІ ТА ЇХ ЗАСТОСУВАННЯ

Актуальність теми дослідження. Принцип змагальності, який використовують угенеративних змагальних нейронних мережах, дозволив досягти відчутного прогресу в сфері застосування штучних нейронних мереж. Актуальною темою зокрема є покращення чіткості зображань у медицині, оскільки дозволяє отримати вичерпну інформацію про стан організму при зменшенні шкідливого фізичного впливу на нього.

Постановка проблеми. Основна проблема при покращенні чіткості зображень полягає у тому, що для вирішення поставленої задачі в зображення необхідно додати певну інформацію без істотного спотворення початкового зображення. Вирішення такої проблеми потребує розробки механізмів генерування інформації з високим рівнем правдоподібності.

Аналіз останніх досліджень та публікацій. Сучасні генеративні змагальні мережі є важливою складовою досліджень у сфері генеративних моделей. В останні роки розроблено велику кількість генеративних змагальних мереж (ГЗМ), які здатні безпосередньо генерувати зображення з використанням тренувальних наборів або без них.

ГЗМ застосовують для передбачення подій у майбутньому шляхом врахування поточного стану навколишнього середовища та можливих активних дій об’єктів. Важливою сферою застосування ГЗМ є також задачі, мультимодальними результатами у яких є велика кількість правильних відповідей.

Структурні характеристики генеративних змагальних мереж. Всі ГЗМ включають два основних вузли: генератор G та дискримінаторD. Нейронна мережа генератора навчається за результатами роботи нейронної мережі дискримінатора. Мета нейронної мережі дискримінатора полягає в знаходженні відмінності між сигналами від генератора та тими сигналами, що надходять із зовнішнього середовища.

Принципи навчання генеративних змагальних мереж. Тренування ГЗМ включає як пошук параметрів дискримінатора, які максимізують точність, з якою він класифікує об’єкти, так і пошук параметрів генератора, які забезпечують зниження точності дискримінатора.

Результати досліджень. Дослідження проводилися з нейронними мережами дискримінатора та генератора, які складались з послідовно розташованих згорткових шарів та шарів з повним набором міжшарових зв’язків. Тренування проводились з використанням чорно-білих зображень, одержаних в результаті проведення рентгеноскопії та комп’ютерної томографії.

Висновки. В роботі наведено основні напрями сучасного застосування генеративних змагальних мереж, розглянуто їх структурні особливості та принципи навчання. Запропоновано підхід до створення згаданого типу мереж для вирішення задач покращення чіткості зображень, одержаних при рентгеноскопії та комп’ютерній томографії.

Обозний Дмитро,
Поштацька Катерина,
Артем Волокита,
Георгій Луцький

СТУДЕНТСЬКА ЕЛЕКТРОННА ПОШТА ТА СХОВИЩЕ ДАНИХ НА ОСНОВІ ХМАРНИХ ТЕХНОЛОГІЙ

STUDENT E-MAIL AND DATA STORAGE BASED ON CLOUD TECHNOLOGIES

У статі розглядається опис впровадження технологій хмарного сервісу MicrosoftOffice 365 на кафедрі обчислювальної техніки факультету інформатики та обчислювальної техніки Національного технічного університету України «Київського політехнічного інституту імені Ігоря Сікорського».

Ключові слова: MicrosoftOffice 365, хмарний сервіс, інформаційна безпека.

Актуальність теми дослідження.Провідні наукові центри світу мають широкий спектр можливостей для комунікації між студентами, науковими працівниками та персоналом. Впровадження нових стандартів якості зв'язку у всьому світі дає змогу перевірити поштову скриньку, або поділитись файлом за допомогою мобільного телефону з будь-якої точки планети. Зручний інтерфейс та доступ з будь-якого пристрою є також важливою частиною для сервісу, що забезпечує таку комунікацію. Ось чому це питання є досить актуальним сьогодні.

Постановка проблеми. На базі університету було розгорнуто декілька систем, які забезпечували різні потреби, але були відокремлені один від одного (загально університетська пошта та електронний кампус КПІ). Також була відсутня зручна реєстрація для користувача. Усе це сприяло пошуку рішень, які змогли б бути корисними студенту та надавати доступ до інформації в одному місці.

Аналіз останніх досліджень і публікацій. Пошуком оптимального сервісу для забезпечення потреб студентів займались університети по всьому світу. Технології хмарного сервісу MicrosoftOffice 365 були впроваджені в різних університетах закордоном (MIT, Harvard, МГУ ім. М. В. Ломоносова та ін.) та в Україні (КНЕУ, ДонНУ ім. В. Стуса, НТУ «ХПІ» та ін.). Даний сервіс був доступний за ліцензією Microsoft A1.

Виділення недосліджених частин загальної проблеми. Незважаючина велику розповсюдженістьвпровадженняхмарних технологійMicrosoftOffice 365 серед університетів по всьому світу, проблема зручного та зрозумілого пересічному користувачу інтерфейсу не є достатньо дослідженою.

Постановка завдання. Основною метою впровадження технологій від Microsoft є забезпечення максимально зручним та зрозумілим сервісом для обміну даними студентів та викладачів кафедри.

Викладення основного матеріалу.

Ліцензії отримуються шляхом реєстрації випробувальної версії Office 365 для навчальних закладів та внесенням відповідних DNS записів на DNS сервер. Після верифікації компанією Microsoft навчального закладу стають доступними ліцензії Microsoft A1. Ці ліцензії дають змогу скористатися онлайновою версією Office з

електронною поштою, відеоконференціями, інтегрованою голосовою поштою, спеціальним центром Microsoft Teams для співпраці в класі, інструментами для забезпечення відповідності та засобами захисту інформації, хмарним сховищем даних розміром до 5 ТБ для кожного користувача сервісом. Кількість користувачів необмежена, що задовольнило потреби факультету. На рисунку 1 зображений інтерфейс системи, що демонструє доступні для користувача програми та служби.

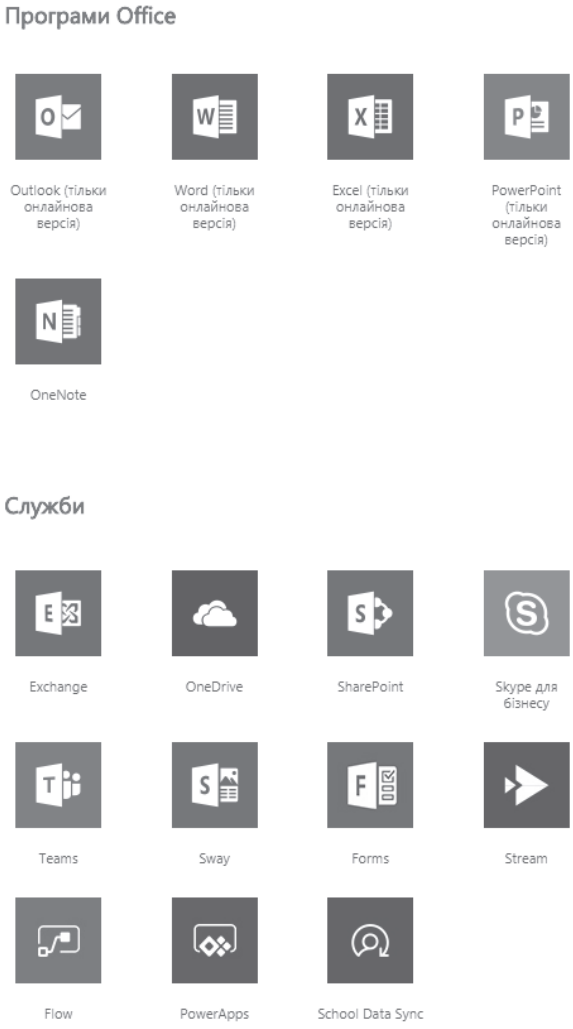


Рис.1. Доступні сервіси за ліцензією A1.

Для забезпечення унікальності формат логінів був обраний наступним чином: «Перша буква імені Прізвище Група@office365.fiot.kpi.ua.Наприклад: DOboznyiIO51@office365.fiot.kpi.ua».

Була створена групова поштова скринька support@office365.fiot.kpi.ua та бот @Office365_Support_bot у месенджері Telegram, що зробило сервіс більш зручним та зрозумілим. Групова скринька забезпечує швидку комунікацію між користувачами та адміністраторами сервісу. Для вирішення глобальних технічних проблем компанія Microsoft надає цілодобову підтримку.

Реєстрація користувачів відбувається шляхом відправлення на поштову скриньку підтримки або боту таблиці з даними студентів групи(рис.2). Реєстрацію можна проводити вручну на сайті або в застосунку для мобільних

пристроїв, або за допомогою скриптів PowerShell. Назад відповідальному за реєстрацію надсилається таблиця з логінами та тимчасовими паролями. Також для кожної групи створюється група поштова скринька, яка виконує роль масової розсилки повідомлень серед учасників групи.

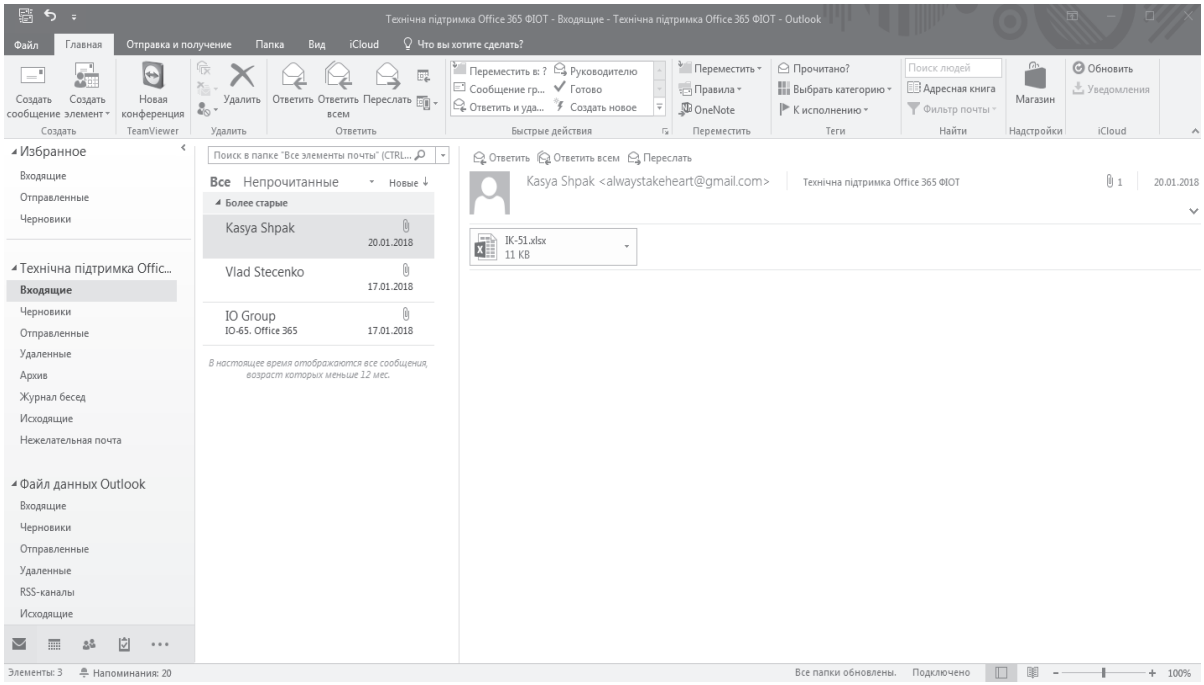


Рис. 2. Реєстрація користувача в системі.

Для забезпечення відповідності стандартам безпеки використовуються політики розпізнавання спам та фішинг повідомлень з можливістю гнучкого налаштування правил (рис 3).

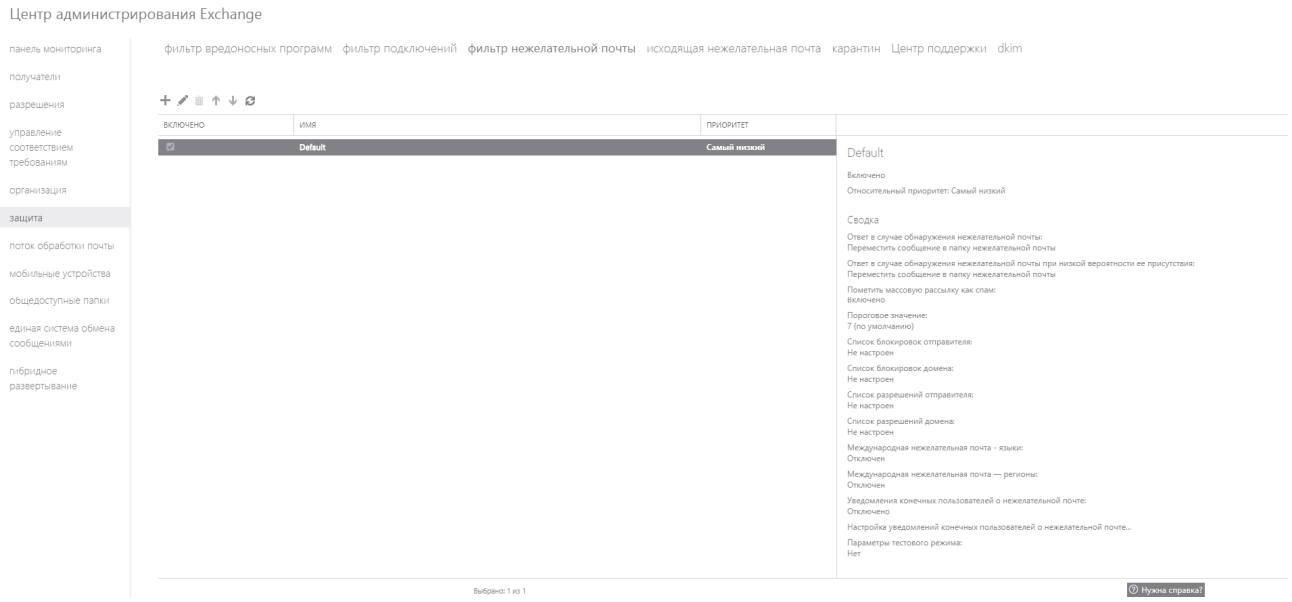


Рис. 3. Політики безпеки.

Висновки. У роботі продемонстрований процес впровадження хмарних технологій для обміну даними серед студентів та викладачів кафедри. Видно, що сервісзадовольняє потреби студентів та викладачів. Незважаючи на те, що проект знаходиться в тестовому режимі, ним користуються 789 користувачів, серед яких близько 100 активних. Загальний об'єм використаного хмарного сховища даних понад 1ТБ. В подальшому планується інтеграція даного застосунку з іншими сервісами університету для поліпшення зручності у використанні даними ресурсами. Також впровадження даного сервісу дає змогу використовувати послуги MicrosoftDreamspark, а також дає доступ до курсів, які надала компанія Microsoft.

Литература

1. Microsoft.Corp. Підтримка користувачів [Електронний ресурс] / Microsoft.Corp // Microsoft. – 2018. – Режим доступу до ресурсу: https://support. office.com/.

ДОВІДКА ПРО АВТОРІВ

Обозний Дмитро Миколайович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені ІгоряСікорського».

Obozniy Dmytro –student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: DObozniy@gmail.com

Поштацька Катерина Володимирівна – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені ІгоряСікорського».

Poshtatska Kateryna–student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: poshtatskayakatirina@gmail.com

Волокита Артем Миколайович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені ІгоряСікорського».

Volokyta Artem – associate professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: artem.volokita@kpi.ua

Луцький Георгій Михайлович – професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Loutskii Heorhii – professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

D. M. Obozniy,
K. V. Poshtatskaya,
A. M. Volokita,
H. M. Loutskii

STUDENT E-MAIL AND DATA STORAGE BASED ON CLOUD TECHNOLOGIES

Relevance of research topic. The world's leading scientific centers have a wide range of communication opportunities between students, researchers and staff. The introduction of new quality standards around the world allows you to check your mailbox, or share a file with a mobile phone from anywhere on the planet. A convenient interface and access from any device is also an important part of the service that provides such communication.

Target setting. Several systems were deployed based on the university that provided different needs, but were separated from each other (general university post and KPI electronic campus). In addition, there was no convenient registration for the user. All this contributed to finding solutions that could be useful to the student and provide access to information in one place.

Actual scientific researches and issues analysis. Universities around the world were looking for the best service to meet the needs of students. Microsoft Office 365 cloud service technologies have been implemented at various universities abroad (MIT, Harvard, Moscow State University, etc.) and in Ukraine (KNEU, DonNU V. Stus, NTU "KPI", etc.). . This service was available under the Microsoft A1 license.

Uninvestigated parts of general matters defining. Despite the widespread use of Microsoft Office 365 cloud technologies among universities around the world, the problem of user-friendly and understandable user interface is not sufficiently explored.

The research objective. The main goal of Microsoft technology implementation is to provide the most convenient and understandable service for the exchange of data of students and faculty members of the department.

Conclusions. The process of introduction of cloud technologies for data exchange among students and faculty of the department is demonstrated in the work. It is evident that the service satisfies the needs of students and teachers. Despite the fact that the project is in test mode, it uses 789 users, among which about 100 active ones. The total amount of cloud data storage used is over 1TB.

Keywords: Microsoft Office 365, cloud service, information security.

Секція 1.SEC
(Безпека комп'ютерних систем та мереж. Відмовостійкі розподілені обчислення..)

Section 1. SEC
(Securityofcomputersystemsandnetworks. Fault-tolerantdistributedcomputing.)

Марковський Олександр Петрович,
Сербін Олександр Дмитрович

АЛГОРИТМ ЦИФРОВОГО ПІДПИСУ З ВИКОРИСТАННЯМ ПОЛІНОМІАЛЬНОЇ АРИФМЕТИКИ НА ПОЛЯХ ГАЛУА

ADIGITAL SIGNATURE ALGORITHM BASEDON POLINOMIAL ARIFMETIC SOFGALO ISFIELDS

Based on the cyclic properties of the exponential operation on the Galois fields, a new method for forming a digital signature is proposed. The constructivity of the proposed method is proved. It is theoretically and experimentally proved that the proposed method allows to accelerate the formation and verification of a digital signature in comparison with known methods.

Key words: digital signature, Galois fields, data protection, irreversible transformations.

На основі циклічних властивостей операції експоненціювання на полях Галуа запропоновано новий метод формування цифрового підпису. Доведена конструктивність запропонованого методу. Теоретично та експериментально доведено, що запропонований метод дозволяє прискорити формування та перевірку цифрового підпису у порівнянні з відомими методами.

Ключові слова: цифровий підпис, поля Галуа, захист даних, незворотні перетворення.

Актуальність теми дослідження. Екстенсивний розвиток і зростаюча доступність Інтернету в останні десятиліття призводять до значного поглиблення взаємодії між віддаленими користувачами. У багатьох сферах людської діяльності, зокрема в банківській сфері, відбувається перехід від паперової форми документів до цифрової, що значно підвищує оперативність їх обробки. Необхідною умовою якісної взаємодії сторін за участю цифрових документів є застосування ефективних механізмів контролю їх цілісності та авторства. Ключову роль серед цих механізмів займають схеми формування та перевірки цифрового підпису. З плином часу та розвитком інформаційних технологій змінюються також способи взаємодії віддалених користувачів, що розширює область застосування цифрового підпису, наприклад для контролю автентичності повідомлень у сервісах для обміну повідомленнями, та, посилює потребу

в ефективних механізмах захисту. З появою і розвитком хмарних технологій широкому колу користувачів стають доступними значні за обсягом обчислювальні ресурси, що можуть бути використані зловмисниками для зламу існуючих алгоритмів цифрового підпису. Потенціал хмарних технологій опосередковано впливає на зниження криптографічної стійкості існуючих алгоритмів та порушує баланс сил у світі інформаційної безпеки; це вимагає пошуку адекватних рішень щодо вдосконалення існуючих криптографічних методів, в тому числі і технології цифрового підпису.

Таким чином, проблема підвищення ефективності і захищеності алгоритмів формування цифрового підпису є актуальною з позиції сучасного етапу розвитку комп'ютерних технологій.

Постановка проблеми. Як і для будь-якого механізму криптографічного захисту даних, для оцінки ефективності схеми формування цифрового підпису використовуються два основних критерії[4]:

1. Рівень захищеності або об'єм затрат, необхідний для зламу засобів захисту.

2. Об'єм витрат, необхідних для реалізації цільового рівня захисту.

Будь-який механізм захисту інформації – це компроміс між цими двома критеріями. Формування цифрового підпису – це допоміжна операція, тому її обчислювальна складність має бути невеликою.

В сценарії формування цифрового підпису задіяні наступні сторони. Відправник – сторона, яка створює, підписує та відправляє документ електронний документ по каналу передачі даних отримувачу. Отримувач - сторона, що отримує електронний документ і перевіряє його на автентичність та цілісність шляхом верифікації підпису. Зловмисник – третя сторона, що має доступ до каналу передачі даних і має на меті підробку або спотворення електронного документу[4].

В більшості випадків підписується не сам документ, а його хеш-сигнатура, завдяки чому зменшується об'єм обчислень і вирішуються проблеми сумісності та цілісності[4].

Переважна більшість існуючих схем цифрового підпису є асиметричними, що базуються на обчислювально складних задачах. Перед схемами цифрового підпису ставляться дві головні вимоги:

1. Верифікація підпису повинна проводитися відкритим ключем, відповідним саме тому закритому ключу, який використовувався під час підписання.

2. Створення легітимного цифрового підпису без наявності закритого ключа має бути задачею, складність якої виходить за рамки технічних можливостей сучасних комп'ютерів.

Основу асиметричних схем складають задачі дискретного логарифмування або факторизації чисел. Схема формування цифрового підпису передбачає три етапи:

1. *Генерація ключової пари*. За допомогою алгоритму генерації ключа з набору можливих значень закритих ключів обчислюється закритий ключ та відповідний йому відкритий ключ.

2. *Формування підпису*. Для заданого електронного документа чи повідомлення за допомогою закритого ключа обчислюється підпис.

3. *Верифікація підпису*. Для даних документа та підпису за допомогою відкритого ключа визначається дійсність підпису.

Аналіз останніх досліджень і публікацій. З моменту виникнення першої публікації про алгоритм цифрового підпису було розроблено багато варіантів. Найбільш поширеними на даний момент є алгоритми на основі RSA, Ель-Гамаля та DSA. Вони є також найбільш дослідженими та вважаються достатньо криптостійкими. Ці алгоритми активно використовуються і прийняті за державні стандарти в багатьох країнах світу.

Алгоритм формування цифрового підписуна основі RSA полягає в наступному. Вибираються два великі прості числа p і g Обчислюється їх добуток $n = p \cdot g$. Далі обчислюється функція Ейлера $\varphi(n) = (p - 1) \cdot (g - 1)$. Вибирається ціле число q таке, що $1 < q < \varphi(n)$ та q взаємно просте з $\varphi(n)$. Далі за допомогою Евклідового алгоритму знаходиться число d таке, що $e \cdot d \bmod \varphi(n) = 1 \bmod \varphi(n)$, тобто $d = e^{-1}$. Числа e, nстановлять відкритий ключ, а число d – закритий. Щоб підписати документ або повідомлення m користувач має обчислити хеш-сигнатуру повідомлення $H(m)$ і значення $s = H(m)^d \bmod n$. Перевірка підпису передбачає обчислення $a = s^e \bmod n$. Якщо $a = H(m)$, підпис вважається дійсним[5].

Алгоритм Ель-Гамаля базується на обчислювальній складності операції дискретного логарифмування на скінчених полях. Підпис передбачає наступну послідовність дій. Обчислюється хеш-функція повідомлення mяк$H(m)$. Далі відбувається генерація ключів: генерується випадкове просте число p, підбирається таке g, що є первісним корнем p, вибирається випадкове ціле число x, таке, що $1 < x < p - 1$. Наступним кроком є обчислення $y = g^x \bmod p$. Далі вибирається випадкове числоk таке, що $1 < k <$ (p -1) і обчислюється $r = g^k \bmod p$. Обчислюється число $s = k^{-1} (H(m) - x \cdot r) \bmod (p - 1)$. Пара (r, s) становить підпис повідомлення.

В свою чергу, щоб перевірити підпис треба спочатку перевірити виконання умов $0 < r < p$ і $0 < s < p - 1$. Далі треба обчислити хеш-функцію повідомлення$H(m)$ і перевірити, чи $y^r \cdot r^s = g^{H(m)} \bmod p$. Якщо умова виконується, то підпис вірний[6].

Алгоритм DSA є варіантом алгоритму Ель-Гамаля. В ньому також використовується складність обчислення логарифма в кінцевих полях. Порядок підписування документа виглядає наступним чином. Спочатку вибираються хеш-функція $H(m)$ та число q, що мають однакову бітову розрядність. Далі вибирається просте число p, таке, що $(p - 1)$ ділиться на q. Вибирається таке число g, що його мультиплікативний порядок по модулю p дорівнює q.

Процедура підписування передбачає таку послідовність дій. Вибирається випадкове число k{\displaystyle k\in (0,q)} таке, що $0 < k < q$, вибирається випадкове ціле число x, таке, що $1 < x < p - 1$.Обчислюється {\displaystyle r=(g^{k}\mod p)\mod q}$(r = g^k \bmod p) \bmod q$. Якщо {\displaystyle r=0}$r = 0$, вибирається інше k. Обчислюється {\displaystyles=k^{-1}(H(m)+x\cdotr)\modq}$s = k^{-1}(H(m) + xr) \bmod q$. Якщо {\displaystyle r=0}$s = 0$, вибирається інше k. Підпис становить пара (r, s) загальної довжинизззз{\displaystyle 2N} загальної довжини 2N.

Перевірка підпису відбувається наступним чином. Обчислюється $w = s^{-1} \bmod q$. Обчислюється $u_1 = H(m) \cdot w \bmod q$. Обчислюється $u_2 = r \cdot w \bmod q$. Обчислюється $v = (g^{u1} \cdot y^{u2} \bmod p) \bmod q$. Підпис вірний, якщо $v = r$.

Виділення недосліджених частин загальної проблеми. З плином часу і прогресом комп’ютерних технологій зростає необхідність в підвищенні рівня

захищеності алгоритмів. Рівень захищеності для цього класу алгоритмів може бути підвищений лише за рахунок збільшення розрядності операндів. Разом з тим, добре відомо[5], що збільшення розрядності в двоє призводить до 8-микратного росту обчислювальної складності. Значне підвищення часу реалізації порушує часові нормативи роботи протоколів мережевого захисту.

Головним недоліком алгоритму RSA є обчислювальна складність операції експоненціювання. Також, алгоритм вразливий до різного роду атак при неграмотно підібраних параметрах та при повторному використанні секретних ключів[3].

Використання алгоритму Ель-Гамаля також пов'язане з певними труднощами. Число k необхідно змінювати кожного разу як підписується документ, інакше знаходження закритого ключа стає легкою задачею [6]. Основні обчислювальні затрати йдуть на обчислення експоненти $r = g^k mod p$ та знаходження k^{-1} за алгоритмом Евкліда. Якщо порівнювати алгоритм Ель-Гамаля з RSA, то при тому ж рівні стійкості він оперує з цілими числами на 25% коротше, ніж RSA, але довжина підпису виходить в 1,5 рази більше, що збільшує час її обчислення і посилює вимоги до надійності каналу передачі даних.

У DSA піднесення до степені по модулю та знаходження обернених значень k^{-1} та s^{-1} є, як і у алгоритмі Ель-Гамаля, найбільш ресурсомісткими операціями. Криптостійкість алгоритма сильно так само залежить від якості підбору випадкового параметру k. Повторення параметру k для двох повідомлень робить можливим обчислення ключа третьою стороною [4].

Таким чином, основний недолік існуючих алгоритмів цифрового підпису полягає в тому, що в сучасних умовах їх обчислювальна складність стає на заваді дотримання часових рамок протоколів мережевого захисту інформації малопотужними термінальними пристроями систем комп'ютерного моніторингу та управління.

Постановка завдання. Метою статті є підвищення швидкості формування цифрового підпису за рахунок використання альтернативного математичного базису - скінчених полів Галуа.

Викладення основного матеріалу. Для досягнення поставленої мети – підвищення швидкості формування цифрового підпису, пропонується замість модулярної арифметики використовувати поліноміальну арифметику полів Галуа, в якій такі базові операції, як експоненціювання і множення, потребують значно менших обсягів обчислень ніж у класичній модулярній арифметиці. Так, піднесення числа до квадрату зводиться до операції вставки нулів між розрядами числа, а відсутність переносів, в свою чергу, зменшує обчислювальну складність операцій множення і додавання. Важливою особливістю операцій на скінчених полях є незалежна обробка розрядів, що дає змогу ефективно організувати розпаралелювання обчислювального процесу.

Пропонується метод формування цифрового підпису, що базується на використанні циклічних властивостей операції експоненціювання на полях Галуа та знання періоду повторення певних чисел в полях Галуа, якщо відомі поліноми, на які розкладається породжуючий поліном поля.

Циклічні властивості операції експоненціювання чисел в полях Галуа були досліджені в [1].

Нехай дано поле *GF(M(x))*. Для будь-якого числа *A*, що співвідноситься з поліномом *a(x)* з поля *GF(M(x))*, існує таке найменше ціле додатне *T*, що$A|^T remM = A$. Число*T* називається періодом повторення або порядком числа *A*. Числа, що належать до одного поля і мають однаковий період називаються такими, що належать до однієї циклічної групи.

Нехай дано два простих поліноми *g(x)* і *p(x)*, ступенів відповідно *d* і *v*. Нехай існують $h = 2^d-1$, $l = 2^v-1$. Тодіh іlможна представити добутком простих множників відповідно як $h = \alpha_1^{k1} \alpha_2^{k2} \dots \alpha_n^{kn}$і$l = \beta_1^{z1} \beta_2^{z2} \dots \beta_m^{zm}$. Числа в полі, породженому добутком *g(x)* і *p(x)*, мають періоди повторення, значення яких відповідають усім можливим комбінаціям добутків простих множниківg(x) і *p(x)* і їх степенів. Кількість періодичних груп обчислюється за формулами комбінаторики для сполучень з *n* елементів по *m*. При кількості простих множників у степенях ≤ 3 (наприклад, якщо канонічний розклад *g* і *p* виглядає так: 5^2, 7^2, 11^3, 31) кількість груп чисел, що мають однаковий період можна обчислити за формулою: ф

$$S = \sum_{m=1}^{n} C_m^n + (\sum_{i=1}^{p} k_i - p) \sum_{m=1}^{n-1} C_m^{n-1} + (C_2^{\sum k_i - p} - \sum_{i=1}^{p} k_i - (p+1)) + \prod_{i=1}^{p} (k_i - 1) \sum_{m=1}^{n-p} C_m^{n-p} + \sum_{i=1}^{p} (k_i - 1)$$

Наприклад, якщо в якості *g(x)* і *p(x)* взяти 11 і 55 та обчислити породжуючий поліном поля *M(x)* = *p(x)*⊗*g*(*x*), числа будуть згруповані за наступними періодами: 1, 7, 31, 217.

Група чисел з періодом T = 1 містить наступні числа: 1,*a* = *u*⊗*p*та*b* = *w*⊗*g* такі, що *a*⊕*b* = *1*.

Згідно з [2], $(u \otimes p)|^{h+1} remm = u \otimes p$.

Якщо *h* та *l* - прості числа, то групи чисел з періодами *T* = *h*+*1* або*T* = *l*+*1* містять добутки $\varphi_1 \otimes p$, $\varphi_2 \otimes p$, ... $\varphi_{h-1} \otimes p$та $\eta_1 \otimes g$, $\eta_2 \otimes g$ ... $\eta_{l-1} \otimes g$відповідно.

В групі чисел з *T* =*h*· *l* знаходяться усі інші числа.

Базуючись на вказаних вище властивостях, було розроблено метод цифрового підпису, що передбачає використання періоду повторення чисел поля як закритого ключа.

З написаного вище стає очевидним, що період повторення чисел в полі Галуа визначається множниками породжуючого полінома *M(x)*, які є відомими лише відправнику. Для зловмисника обчислення періоду числа з поля *GF(M(x))* шляхом простого експоненціювання виявилося б занадто ресурсомістким, приймаючи до уваги, що в реальній практиці розрядності операндів становлять 2048 та 4096. З цієї причини в методі цифрового підпису було б доцільним використовувати період *T* в якості закритого ключа.

Для методу, що пропонується, інтерес представляють лише числа з періодом *T* = *h*· *l*. Пошук числа *A*, що належить до цієї групи виконується наступним чином. Вибирається випадкове число $1 < A < 2^{dv} - 1$, що задовольняє наступним умовам:

1. *A remg*≠ *0*
2. *A remg*≠ *1* (1)

3. $A\ remp \neq 0$

4. $A\ remp \neq 1$

Якщо всі умови виконані, можна вважати, що число A не належить до перших трьох груп, а значить належить до останньої групи з $T = h \cdot l$. Якщо умови не виконані вибирається нове випадкове число число.

Процедура формування цифрового підпису виглядає наступним чином:

1. Відправник обчислює хеш-функцію документа m як $H(m)$, що в двійковому представленні має розрядність l. При використанні алгоритму SHA256, $l = 256$.

2. Відправник довільнимчином вибирає пару простих поліномів $p(x)$ та $g(x)$ з різними степенями: $p(x) = x^v + p_{v-1} \cdot x^v + ... + + p_1 \cdot x + p_0$ степені v та $g(x) = x^d + g_{d-1} \cdot x^{d-1} + + ... + g_1 \cdot x + g_0$ степені d, де $p_0, p_1, ..., p_{v-1} \in \{0,1\}$, $g_0, g_1, ..., g_{d-1} \in \{0,1\}$, причому $d > v$ і $(d + v) > 258$.

3. Відправник формує поліном $M(x)$ у вигляді поліноміального добутку вибраних двох поліномів $p(x)$та $g(x)$: $M(x) = p(x) \otimes g(x)$.

4. Відправник вибирає число A, що задовольняє умові (1) і співвідноситься з поліномом $a(x)$ в полі та період повторення якого відомий тільки відправнику і становить T. $T > 2^{257}$.

5. Відправник вибирає випадкове число e таке, що $1 < e < 2^{dl}$.

6. Відправникобчислює число d, що є мультиплікативно оберненим до e, тобто для якого виконується рівність $d \otimes e\ remM = 1$.Число d становить відкритий ключ. Якщо такого числа не існує, відправник повертається до пункту 5.

7. Відправник вибирає випадкове число k: $0 < k < 2^{255}$і обчислює $r = T - H(m) - k$. Далі відправник обчислює $F = A|^k \otimes A|^r remM$.

8. Відправник обчислює$W = F \otimes e\ remM$і додає отримане число до цифрового підпису. Цифровий підпис становлять наступні числа: $<W, A, M, d>$.

Запропонована процедура перевірки цифрового підпису отримувачем:

1. Отримувач отримує документ m.

2. Отримувач обчислює хеш-функцію документа m як $H(m)$.

3. Отримувач обчислює $C = A|^{H(m)} remM$.

4. Отримувач обчислює$D = W \otimes d \otimes Crem M$. Якщо $D = A$, то вважається, що верифікація пройшла успішно і підпис вірний.

Аналіз ефективності. Як і для будь-якого механізму криптографічного захисту даних, ефективність запропонованого методу ідентифікації віддалених користувачів оцінюється за двома базовими критеріями: рівнем захисту та часовими характеристиками реалізації.

Оцінку рівня захищеності доцільно виконувати з позиції третьої сторони, яка може мати доступ до документу та цифрового підпису, але не знає множників, на які розкладається утворюючий поліном $M(x)$ поля Галуа.

Для зловмисника, що має доступ до каналу передачі даних, задача відтворення періоду числа A зводиться до задачі розкладання відомого поліному $M(x)$ на два простих множника $p(x)$ та $g(x)$ з різними ступенями. Вирішення такої задачі шляхом перебору для реальних степенів поліномів виходить далеко на межі технічних можливостей сучасних комп'ютерних систем.

Основна перевага запропонованого способу формування цифрового підпису полягає в тому, що використання експоненціювання в полях Галуа на

відміну від модулярного експоненціювання дозволяє значно прискорити час виконання програм та спростити апаратну реалізацію. Нижче наведені основні чинники, які дозволяють прискорити програмну реалізацію незворотних перетворень на полях Галуа в порівнянні з мультиплікативними перетвореннями модулярної арифметики, що лежать в основі існуючих методів:

- операція піднесення числа до квадрату, питома вага якої складає 75% об'єму обчислень, на полях Галуа зводиться до вставки нулів між бітами числа, тобто не потребує ніяких обчислювальних операцій. В той же час операція піднесення до квадрату n-розрядного числа в традиційній алгебрі потребує виконання $(\frac{n}{m})^2/2$ операцій процесорного множення;

- при виконанні операцій на полях Галуа кожен розряд оброблюється незалежно від інших: це дає змогу ефективно організувати розпаралелювання обчислювального процесу, особливо при використанні апаратних засобів;

- операції на полях Галуа не використовують переносу, що значно прискорює виконання в порівнянні з традиційною алгеброю, для яких формування міжрозрядних переносів для розрядностей 2018 і 4096, що здебільшого використовуються на практиці, потребує помітних затрат часових та апаратних ресурсів.

В роботі [1] показано, що експоненціювання в полях Галуа виконується на 2-3 порядки швидше у порівнянні з модулярним експоненціюванням, яке є базовою процедурою для існуючих методів DSA [4], RSA [4] та Ель-Гамаля[4] реалізації цифрового підпису.

Експериментальні дослідження показали, що при апаратній реалізації алгоритму на програмованих матрицях досягається значне прискорення процедур формування цифрового підпису в порівнянні з методом Ель-Гамаля. Зокрема, моделювання засобами VHDL показало зростання швидкості формування цифрового підпису на 1-2 порядки.

Висновки. В роботі досліджено можливості використання незворотних перетворень на полях Галуа для створення ефективних механізмів формування цифрового підпису. Зокрема, теоретично досліджені властивості циклічності операції експоненціювання на полях Галуа спеціальних класів.

На основі отриманих теоретичних результатів запропоновано метод формування цифрового підпису.

Внаслідок того, що ці операції виконуються на 1-2 порядки швидше за мультиплікативні операції модулярної арифметики, що лежать в основі відомих методів контролю цілісності та автентичності даних, використання запропонованого методу дозволяє суттєво прискорити процес формування цифрового підпису.

Список літератури

1. Николайчук Я. Коди поля Галуа: теорія та застосування./Монографія/ Николайчук Я. - Тернопіль: ТзОВ "Терно граф". - 2012. – С. 392.

2. Марковський О. Використання алгебри полів Галуа для реалізації концепції «нульових знань» при ідентифікації та автентифікації віддалених користувачів. / Марковський О., Захаріудакіс Лефтеріс, Максимук В. - Електронне моделювання. - 2017. - Т.39. - № 6. - С. 32 – 46.

3. BonehD. TwentyYearsofAttacksontheRSACryptosystem./ BonehD. - NoticesoftheAmericanMathematicalSociety (AMS). – 1999. - Vol. 46. - No. 2. - PP. 203-213.

4. Schneier B. Applied Cryptography. Protocols. Algorithms and Source codes in C./ Schneier B. - Ed.John Wiley. – 1996. – P. 784.

5. Rivest R.L. A Method for Obtaining Digital Signatures and Public-Key Cryptosystems./ Rivest R.L., Shamir A., Adleman L. - Computers and Mathematics with Applications. - 1979.– Vol. 5. – No. 3. - PP. 169-178.

6. ElGamal T. A public key cryptosystem and a signature scheme based on discrete logarithms./ ElGamal T.A. - IEEE Trans Inf Theory. – 1985. – Vol. 31. – No. 3. - PP. 469–472.

ДОВІДКА ПРО АВТОРІВ

Марковський Олександр Петрович – доцент, кандидат технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Markovskyi Oleksandr – associate professor, candidate of engineering sciences, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: markovskyy@i.ua

Сербін Олександр Дмитрович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Serbin Oleksandr – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: s0meparry@gmail.com

O. Markovskyi, O. Serbin

A DIGITAL SIGNATURE ALGORITHM BASED ON POLINOMIAL ARIFMETICS OF GALOIS FIELDS

Relevance of the topic. In many spheres of human activity, in particular in the banking sector, there is a transition from the paper form of documents to digital, which greatly increases the efficiency of their processing. A prerequisite for the qualitative interaction of the parties with the participation of digital documents is the application of effective mechanisms for controlling their integrity and authorship. A key role among these mechanisms is taken by the schemes for the formation and verification of digital signatures.

Target setting. With the advent and development of cloud technologies, a wide range of users gain access to large volumes of computing resources that can be exploited by malicious people to break the existing algorithms of digital signature. The potential of cloud technologies indirectly affects the reduction of cryptographic stability of existing algorithms and violates the balance of power in the world of information security; this requires finding adequate solutions for improving existing cryptographic methods, including digital signature technology.

Actual scientific researches and issues analysis. Since the first publication of the digital signature algorithm, many options have been developed. The most commonly used algorithms are based on RSA, El Gamal and DSA. They are also the most researched and are considered to be cryptographic resistant. These algorithms are widely used and adopted as state standards in many countries around the world.

Uninvestigated parts of general matters defining. With the passage of time and the advancement of computer technology, the need to improve the level of security of algorithms is also increasing. The level of security for this class of algorithms can only be improved by increasing the length of the operands. This leads to exponential growth of computing complexity and, in turn, to significant increase in the implementation time, that violates the time standards of network protection protocols.

The research objective. The objective is to increase the speed of the formation of a digital signature by using an alternative mathematical basis - finite Galois fields.

The statement of basic materials. A digital signature scheme is proposed, which is based on the use of cyclic properties of the exponential operation in Galois fields and knowledge of the period of repetition of certain numbers in Galois fields, if known polynomials on which the generating polynomial of a field is decomposed. The main advantage of the proposed scheme of a digital signature is that the use of exponentiation in the fields of Galois, in contrast to the modular exponentiation, can significantly accelerate the execution time of the programs and simplify the hardware implementation.

Conclusion. The possibilities of using irreversible transformations in Galois fields to create effective mechanisms for the formation of a digital signature are explored. Based on the obtained theoretical results the method of digital signature creation is proposed. Due to the fact that these operations are performed 1-2 orders faster than multiplicative operations of modular arithmetic, which are the basis of known methods for controlling the integrity and authenticity of data, the use of the proposed method allows to significantly accelerate the process of forming a digital signature.

Key words: digital signature, Galois fields, data protection, irreversible transformations.

УДК-004.7

Марич Володимир, Волокита Артем

СПОСІБ ПЛАНУВАННЯ ПЕРЕВІРКИ РЕЗУЛЬТАТІВ ВОЛОНТЕРСЬКИХ ОБЧИСЛЕНЬ ДЛЯ ЗАДАЧ РЕАЛЬНОГО ЧАСУ

METHODS OF SCHEDULING FOR CHECKING THE RESULTS OF VOLUNTEER COMPUTING FOR REAL-TIME TASKS.

У статті представлено спосіб планування перевірки результатів волонтерських розподілених обчислень, який за рахунок оцінки обсягу перевірок на основі скорингової моделі з алгоритму планування earliestdeadlinefirst, дозволяє використовувати волонтерські обчислення для задач реального часу.

Ключові слова: волонтерські обчислення, планувальник задач, скоринг, грід система, реальний час.

Рис.: 1. Бібл.: 10.

The article presents a method for planning the verification of the results of volunteer distributed computing, which, by estimating the amount of rechecks based on a scoring model with a planning algorithm earliest deadline first, it allows you to use volunteer calculations for real-time tasks.

Key words:volunteer computing, task scheduler, scoring, grid system, real-time.

Fig.: 1. Bibl.: 10.

Актуальність теми дослідження. З кожним роком обсяг і складність обчислювальних задач монотонно зростають, можливостей персональних комп’ютерів не вистачає, і тому багато хто звертається до компаній, що володіють суперкомп’ютерами, або обраховують свої задачі в хмарах. Дані послуги є дуже дорогими, тому в останній час велику популярність набирають системи добровільних розподілених обчислень. Переважна більшість цих систем безкоштовні. В якості обчислювальної одиниці виступають персональні комп’ютери, або ноутбуки будь-якої людини, що добровільно надала доступ до своїх ресурсів. Це особливо турбує дослідників, чиї погляди розвиваються швидше, ніж їх комп'ютери. Томудослідження протікають дуже і дуже повільно, через брак ресурсів або коштів, необхідних для виконання масивних обчислювальних проектів.

Дані системи побудовані на мультиагентній структурі. Така структура дозволяє за невеликий проміжок часу організувати розподіленні обчислення на основі звичайної комп’ютерної мережі. Яскравим прикладом такої програмної системи є BOINC [1].

Постановка проблеми. При організації розподілених обчислень на основі звичайних мережевих комп'ютерів можуть виникати загрози безпеки обчислювальних процесів, а саме відмова обладнання або несанкціоновані дії зловмисників, що спрямовані на блокування вузлів обчислювальної системи або дискредитацію результатів роботи розподіленого обчислювального процесу.

Перша загроза не є надто суттєвою, бо у разі втрати доступу до ресурсів однієї, або декількох обчислювальних машин, система перерозподілить задачі

між іншими. Тому основною проблемою виступає – недостовірність отриманих результатів обчислень.

Є лише єдина можливість впевнитись в правильності отриманої інформації – це повторний обрахунок задачі на “перевірених” комп’ютерах. Вони називаються агентами системи. Однак ресурс агентів і час обчислення задачі обмежені. Тому виникає питання оцінки обсягу перевірки результатів.

Аналіз останніх досліджень і публікацій. У статті [2] розглядається проблема оптимізації вибіркових перевірок для забезпечення надійності грід систем, в яких деякі ненадійні вузли (диверсанти) можуть виконувати атаки на системи.

У статті [3] автори узагальнюють вибіркову перевірку шляхом введення ідеї перевірки дефектів. Це узагальнення дозволяє гарантувати коректність обчислень в ситуації, коли вибіркова перевірка не є повністю надійною.

У статті [4] описується грід система, яка дозволяє здійснювати обчислення на віддаленій, а також потенційно небезпечній, системі для перевірки прогресу та правильності виконання обчислень.

Виділення недосліджених частин загальної проблеми. Виходячи із недоліків структури добровільних розподілених обчислень, а саме можливості дискредитації результатів з боку зловмисників, та обмеженості ресурсів агентів, виникає питання створення способу планування перевірки результатів.

Постановка завдання. Створити спосіб планування перевірки результатів волонтерських обчислень в реальному часі, що включає визначення обсягу перевірки та їх порядок.

Викладення основного матеріалу. Мультиагентна система [5] являє собою множину агентів Aу вигляді перевірених надійних комп’ютерних модулів агентів $\{a_1, a_2, ..., a_n\} \subseteq A$. Множина Aнакладається на множину $\{p_1, p_2, ..., p_j\} \subseteq P$мережевих комп'ютерів ($P=A$) так, що агент a_i закріплюються за відповідним p_i комп'ютері мережі. Агенти взаємодіють через сервер.

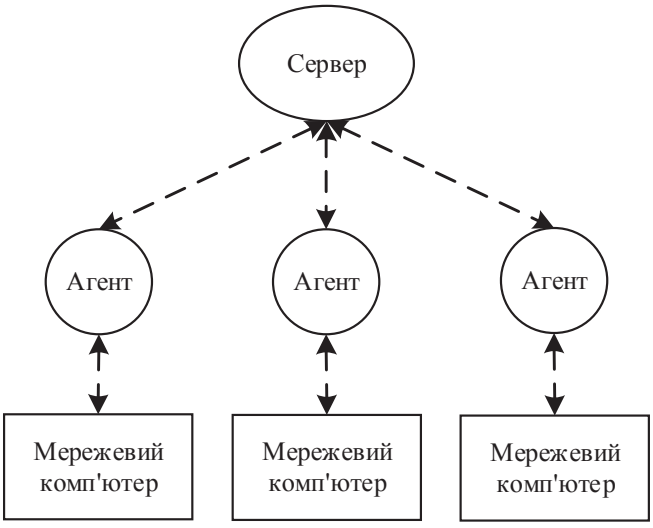


Рис.1 Структура мультиагентної системи

Вся мультиагентна система, керуючи комп'ютерами p_i, організовує систему розподілених обчислень для вирішення всієї множини T завдань $\{t_1, t_2, \ldots, t_n\} \subseteq T$.

Кожен модуль агента керує ресурсами комп'ютера і стежить за виконанням завдання. У свою чергу, кожен комп'ютер працює незалежно від інших комп'ютерів.

Алгоритм агента функціонує наступним чином. На початку організації розподілених обчислень в комп'ютерній мережі $\{p_1, p_2, \ldots, p_n\} \subseteq P$ знаходяться, керуючі їх роботою, агенти $\{a_1, a_2, \ldots, a_n\} \subseteq A$. На першому етапі агент a_i, отримує основну інформацію для організації розподілених обчислень. Вона включає в себе обчислювальну задачу t_iі її частину, яку агент повинен виконати. Після отримання агентом задачі і загальної інформації про систему він ініціює на своєму комп'ютері обчислювальний процес для виконання t_i. Після виконання кожного рішення агент розсилає результат усім іншим агентам.

Очевидно, що використання окремого агента для контролю лише одного обчислювального модуля є не раціональним. Тому кожен агент має в своєму розпорядженні декілька модулів.

Оскільки система складається з добровільно наданих комп’ютерів p_i - існує ймовірність втручання зловмисників, для дискредитації результатів. Тому одним із найголовніших пунктів, що входять в обов’язки роботи агента, це перевірка валідності обчислень.

Кожна мультиагентна система, відповідно до своїх ресурсів, може вирішувати декілька задач. Для кожної задачі система вираховує кінцевий час вирішення обчислення. Для підтримки цілісності системи, результати цих задач повинні бути надані вчасно. Цю проблему також повинні вирішувати агенти у реальному часі.

Виходячи з цих потреб, та беручи до уваги, що ресурси агентного модуля обмеженні, постає питання оцінки обсягу результату який потрібно перевірити. Для кожного результату, наданими окремим обчислювальним модулем, його частина для повторного обчислення, та позиція в черзі на перевірку в агенті буде різною.

Для впевненості у достовірності отриманої інформації слід враховувати велику кількість факторів, які на пряму впливають на обсяг перевірки результату та на пріоритет в черзі. І з основних можна виділити:

- кінцевий час виконання задачі;
- складність та важливість задачі;
- обсяг отриманого результату;
- час витрачений на розв’язок;
- довіра до обчислюваного модуля.

Перші два фактори напряму задаються системою. Третій і четвертий отримуємо після закінчення виконання обчислень. Останній включає в себе комплекс факторів, і теж вираховується системою.

Високий рівень довіри може мати об'єкт, який не є джерелом деструктивних впливів. Низьким рівнем довіри володіють об'єкти, які щойно з'явилися в системі, раніше невідомі об'єкти, тобто об'єкти, які є або можуть завдавати шкоди системі.

Тобто рівень довіри до об'єкта – це ймовірність того, що об'єкт не є джерелом деструктивного впливу. Рівень довіри в мультиагентної системи обчислюється різними методами, де основною ідеєю виступає сумарна оцінка для кожного обчислювального модуля від кожного агента [6-8].

Всі ці фактори є різнотипними, тому для оцінки їх впливу пропонується використовувати скорингову модель.

ЇЇ суть [9] полягає в тому, що кожному параметру, що характеризує об’єкт, надається реальна оцінка в балах. У спрощеному вигляді скорингова модель – це зважена сума визначених характеристик. Така методика є універсальною і може застосовуватись для будь-яких об’єктів. За допомогою скорингу на основі виділених факторів можна визначити, наскільки велика ймовірність того, що конкретний об’єкт може вчинити ті чи інші дії. Данна модель вже давно використовується у кредитній системі. Загальний вигляд скорингової формули:

$$S = \sum_i^n k_i * x_i \quad (1)$$

S – сумарний бал;
k_i – коефіцієнт;
x_i – бал окремого фактору.

Для кожного окремого фактора x_i є своя верхня оцінка. Якщо ні, то слід використовувати монотонно зростаючі обмежені зверху і знизу функції $f(x_i)$. Відповідно до характеру прояву фактору можна використовувати ряд наступних функції: функція експоненціального розподілу; функція Гамма-розподілу; функція розподілу Вейбула. Всі вони монотонно зростають від 0 до 1.

Завдяки тому, що під час отримання локальних максимумів окремих факторів, існує загальний максимум, можна точно визначити обсяг отриманих даних який потрібно перевірити. Формула має наступний вигляд:

$$n = (1 - {}^{S}/_{S_{max}}) * N \quad (2)$$

S_{max} – максимальний можливий бал;
N – розмір отриманого результату;
n – розмір результату який потребує перевірки.

Найголовнішими факторами для обчислення обсягу перевірки є довіра та складність задачі. Якщо складність для всіх обчислювальних модулів однакова, бо вони вирішують одну задачу, то ступінь довіри до кожного – індивідуальна. Тому при правильному виборі коефіцієнтів та формул, даний підхід надає змогу швидко та з високою точність перевіряти результати отриманні від добровільних обчислювальних модулів.

Частини результатів, що потребують повторного обчислення, для комп’ютерів з високим балом довіри будуть менші, і навпаки. Також можна додатково ввести обмеження, що розмір і важливість задач, які надаються для обробки, напряму залежить від ступеню довіри. Це дозволить запобігти втручанню зловмисників, оскільки для дискредитації серйозних розмірів результату всіє задачі, вони потребують підвищення рівню довірі, що в свою чергу залежить від кількості правильно обчислених раніше задач.

Такі фактори як: кінцевий час виконання, складність і важливість задачі, та час витрачений на розв'язок, використовуються під час обчислення пріоритету задачі в черзі. Для його підрахування можна також використовувати скорингову модель, за формулою (1). Чим більший бал, тим ближчий номер задачі в черзі агента, на виконання повторного обчислення.

Процес вибору наступного результату за пріоритетом, подібний до планування процесів на виконання у комп'ютерній операційній системі. Існує два основних підходи статичний або динамічний. Статичне планування обчислює виконання замовлень перед початком виконання. Це вимагає знання характеристик задач, і створює невеликі накладні витрати. Динамічне планування приймає рішення під час виконання процесів. Це дозволяє створювати більш гнучку систему, але потребує додаткових витрат.

Мультиагентна система відноситься до систем, що працюють в режимі реального часу, тому доцільніше використання алгоритмів з динамічним планування. Враховуючи характеристики системи слід розглядати earliestdeadlinefirst (EDF) [10]. Він працює наступним чином: при настанні дії планування (задача завершена, встановлена нова задача і т. д.), планувальник шукає найближчий до крайнього часу її виконання завдання, і цей процес призначається для виконання.

Планування за найближчим терміном завершення є оптимальним переважно для одно-процесорних систем, в тому відношенні, що якщо існує можливість виконати набір незалежних завдань, кожен з яких характеризується часом настання, вимогою виконання та крайнім терміном завершення, так, щоб гарантувати виконання всіх задач до крайнього терміну.

В даній моделі під час надання пріоритетів агент враховує фактори, які впливають на організацію черги в EDF-алгоритмі. Пріоритети слід обраховувати після кожного вибору задачі із черги, для підтримання динамічності системи.

Висновки. Проаналізувавши вразливість добровільних обчислень, було розроблено спосіб планування перевірки результатів обчислень в реальному часі. В якості оцінки обсягу використовується скорингова модель. Планування базується на алгоритмі EDF.

В подальших дослідженнях авторами плануються розглянути різні типи задача реального часу, в тому числі розпізнавання образів і block-chain обчислення., та запропонувати більш ефективні та захищенні способи перевірок для таких задача.

Список використаних джерел

1. Anderson, David P. Boinc: A system for public-resource computing and storage. //5th IEEE/ACM International Workshop on Grid Computing – IEEE Computer Society – 2004.

2. WatanabeK. etal. OptimalSpot-CheckingforDelayedAttackonDesktopGrid Systems //ComputerModellingandSimulation (UKSim), 2013 UKSim 15th International Conference on. – IEEE, 2013. – С. 600-605.

3. Fukushi M, Watanabe K. Generalized spot-checking for reliable volunteer computing //IEICE TRANSACTIONS on Information and Systems. – 2010. – Т. 93. – №. 12. – С. 3164-3172.

4. Yang S. et al. Trust but verify: monitoring remotely executing programs for progress and correctness //Proceedings of the tenth ACM SIGPLAN symposium on Principles and practice of parallel programming. – ACM, 2005. – С. 196-205.

5. Ховансков С. А., Литвиненко В. А., Хованскова В. С. Алгоритм организации безопасных распределенных вычислений на основе многоагентной системы //Известия Южного федерального университета. Технические науки. – 2016. – №. 10 (183)

6. Губанов Д. А. Обзор онлайновых систем репутации/доверия. – 2009.

7. Абрамов Е. С., Басан Е. С. Разработка модели защищенной кластерной беспроводной сенсорной сети //Известия Южного федерального университета. Технические науки. – 2013. – №. 12 (149).

8. Зикратов И. А. и др. Построение модели доверия и репутации к объектам мультиагентных робототехнических систем с децентрализованным управлением //Научно-технический вестник информационных технологий, механики и оптики. – 2014. – №. 3 (91).

9. Волик, Н. Г. Скоринг як експертний метод оцінювання кредитного ризику комерційного банку при споживчому кредитуванні //Вісник Запорізького національного університету: Економічні науки.–Запоріжжя – 2008. – ЗНУ 1 (40-44).

10. NAGER S. K., GILL N. S. Comparative Study of RM and EDF Scheduling Algorithm in Real Time Multiprocessor Environment. – 2017.

ДОВІДКА ПРО АВТОРІВ

Марич Володимир Васильович – студент 4-го курсу, Факультету інформатики та обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Marych Volodymyr – student, Faculty of Informatics and Computer Science, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: 777Titan111@gmail.com

Волокита Артем Миколайович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Volokyta Artem – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: artem.volokita@kpi.ua

Marych Volodymyr,
Volokyta Artem

METHODS OF SCHEDULING FOR CHECKING THE RESULTS OF VOLUNTEER COMPUTING FOR REAL-TIME TASKS.

Relevance of research topic. Every year, the amount and complexity of computing tasks monotonously increases, the capabilities of personal computers are not enough, and so many people turn to companies that have supercomputers, or calculate their tasks in the clouds. These services are very expensive. Therefore, in recent times, voluntary distributed computing systems are becoming increasingly popular. A good example of such a software system is BOINC.

Target setting. In the organization of distributed computing on the basis of original network computers, there may be security threats to performing computational processes, namely, hardware failure or unauthorized intruder actions. Resource agents and time calculations are limited. Therefore, there is a question of assessing the extent of verifying the results.

Actual scientific researches and issues analysis. Over the past years there are more articles devoted to it problems optimize sample checks to ensure the reliability of grid systems, in which some unreliable nodes (saboteurs) can perform attacks on systems.

Uninvestigated parts of general matters defining. Based on the shortcomings of voluntary distributed computing structure, namely the possibilities of discrediting the results from intruders, and agents of limited resources, the question arises about how to plan the results check.

The research objective. Create a method of planning real-time volunteer results checking, which includes determining amount of rechecks and their order.

The statement of basic materials. The analysis of the multiagent system is carried out. The vulnerable aspects of this model were identified. A method for scheduling a results check based on the scorecard model and EDF algorithm is described. The main factors that influence the amount of the result check and the calculation of the priority are considered.

Conclusions. After analyzing the vulnerability of voluntary computing, a method for planning the verification of the results of computations in real time was developed.

УДК004.056

Інна Стеценко, Дмитро Огородников

МОДЕЛЮВАННЯ СЦЕНАРІЮ КІБЕРАТАКИ НА ОСНОВІ АТАКИ «PETYA», «NOTPETYA»

У статті описаний сценарій кібератак на основі вірусів шифрувальників. За результатом аналізу сценарію можна класифікувати більшість кібератак на інформаційні системи.

Ключові слова: кібератака, інформаційні системи, комп'ютерний вірус, життєвий цикл, шифрувальник.

Рис.: 6. Табл.: 1. Бібл.: 6.

The article describes a cyberattack script based on encrypted viruses. As a result of the scenario, most cyberattacks can be classified into information systems.

Keywords:cyberattack, information systems, computer virus, life cycle, encryptor.

Fig.: 6. Tabl: 1.Bibl.: 6.

Актуальність теми дослідження. Повідомлення «Виявлено поширення банківського вірусу трояна», «На державних пристроях виявлена програма для кібер-шпигунства», «Виявлений вірус, який вже заразив п'ять мільйонів пристроїв», «Ботнет навчився перетворювати пристрої на проксі-сервери» стали звичайними новинами тижня. Від шкідливих програм потерпають все частіше державні та фінансові установи, а кібертероризм порівнюють зі зброєю масового знищення. Загрози, які несуть злочини в інформаційному просторі, є небезпечними для транспортних, енергетичних, виробничих систем. Розуміння процесів, які відбуваються при запуску шкідливого програмного забезпечення, є необхідним кроком для розробки систем інформаційного захисту. Тому дослідження кібератак з метою визначення шляхів їх поширення є актуальною задачею.

Постановка проблеми. Для реалізації своїх цілей кіберзлочинці використовують спеціальне програмне забезпечення. Для доступу до особистих інформаційних ресурсів жертви злочинці використовують комп'ютерні віруси, ботнети або мережеві хробаки, які здійснюють копіювання, модифікацію та знищення інформації. Основними цілями кібератак є:

– отримання доступу до державних і військових секретів, а також банківської та особистої інформації (наприклад, вірус Red October);

– нанесення збитків окремим фізичним елементам інформаційного середовища (наприклад, мережевий хробак і ботнет Mirai);

– крадіжка чи знищення інформації методами обходу систем захисту і впровадження стороннього програмного забезпечення (наприклад, вірус Wanna Cry);

– захоплення каналів розповсюдження інформації;

– створення перебоїв у мережах комунікацій.

Виявлення та протидія кіберзагрозам в сучасних інформаційних системах спираються на вже відомі атаки. Проте кіберзлочинці постійно вишукують нові інструменти та розробляють нові програми, яким не можуть протидіяти системи

інформаційного захисту. Систематизація та узагальнення інформації про шкідливі програми та способи їх використання в кібератаках є важливим ресурсом для розробки нових систем інформаційного захисту.

Аналіз останніх досліджень і публікацій. На сьогоднішній день існує достатньо багато наукових досліджень, що стосуються моделювання кібератак та їх сценаріїв. Моделювання DoS/DDoS/DRDoS атак з урахуванням структури мережі та вагових коефіцієнтів, які оцінюють рівень загрози атаки, розглянуто в роботі «Атаки на відмову в обслуговуванні комп'ютерних мереж»[4]. Результати моделювання пропонується враховувати під час проектування систем інформаційного захисту. В роботі «Impact Analysis of Faults and Attacks in Large-Scale Networks»[5] для аналізу безпеки на основі глобальних показників після існуючих ситуацій з ураженням використовують графи і дерева атак. Дослідження слабких місць, політик безпеки, конфігурацій, списків програмного і апаратного забезпечення, встановлених на кожній платформі, події і міри, які можна використати для запобігання, представлено в публікації Common Vulnerabilities and Exposures [6].

Виділення недосліджених частин загальної проблеми. У схемі не розглянуті проблеми, що описує життєвий цикл атак від спроби заподіяння шкоди успішному ураженню до маніпуляціям даних.

Постановка завдання. Для передбачення дій кібернападника необхідні методи, які можуть аналізувати його попередні дії та на їх основі визначати (з певною достовірністю) кінцеву ціль атаки та наступні його дії. Аналіз сценарію кібератаки поділяється на три етапи: 1) класифікація атаки; 2) визначення основних кроків атаки; 3) моделювання та аналіз результатів. За результатом аналізу сценарію можуть бути розроблені інструментальні засоби виявлення відомих атак вже на перших їх кроках.

Викладення основного матеріалу.Класифікацію атак можна здійснювати, конкретизуючи ознаку, за якою виконується класифікація. Виділимо такі ознаки: характер взаємодії, цілі взаємодії, умова початку здійснення атаки, наявність зворотного зв’язку з об’єктом атаки, розташування відносно об’єкту атаки.

За характером взаємодії розрізняють атаки пасивні та активні. При виконанні активних атак, порушник виконує активні дії, які пов’язані зі зміною потоків даних. При пасивних атаках шкідливий програмний код розміщується, наприклад, у просторі інтернету, і порушник сподівається, що жертва самостійно через неуважність (атаки методом фішингу) виконає запуск шкідливого коду.

За цілями взаємодії розрізняють атаки, які спрямовані на порушення конфіденційності, цілісності або працездатності інформації.

За умовою початку здійснення вирізняють атаки за запитом від жертви, безумовну атаку та атаку з урахуванням умов. Атаку за запитом від жертви можна вважати продовженням пасивної атаки. Після переходу на шкідливий ресурс відбувається «добровільний» запуск шкідливого програмного коду. Безумовною є атака без дослідження систем захисту жертви. Атака з урахуванням умов виконується на основі визначення коефіцієнту успішності проведення атаки.

За наявністю зворотного зв'язку з об'єктом атаки атаки можна поділити на атаки зі зворотнім зв'язком та однонаправлені атаки.

Одним із основних факторів атаки є місце розташування суб'єкту атаки, тому розрізняють зовнішню та внутрішню атаки. При внутрішній атаці порушник може знаходитися одразу в системі, наприклад, як співробітник компанії, або мати доступ до внутрішньої мережі від самого початку. При зовнішній, - порушнику потрібно спочатку оминути зовнішній захист.

Схема класифікації атак представлена на рисунку 1.



Рис. 1. Класифікація кібератак.

Атаки можна розділити на вісім класів за уразливими місцями операційної системи:

1) «соціальна інженерія» - атака, яка направлена на введення жертви в оману;

2) «запозичення прав» - атака, яка має на меті захоплення прав авторизованих користувачів;

3) «використання» - використання вразливостей програмного забезпечення чи операційної системи;

4) «відносна (оманлива) довіра» - використання довіри до мереж чи сайтів;

5) «атаки управлінням даними» - троянські програми, віруси, хробаки;

6) «інфраструктурна вразливість» - використання особливостей стандартів, специфікацій;

7) «відмова в обслуговуванні» - атака, яка направлена на перешкоджання використанню системи;

8) «чародійство» - невідома атака, яка ще не зафіксована або не розшифрована.

Характер кібератаки, у більшості випадків, можна визначити за класом атаки та її спрямованістю. Часто одна атака може бути відволіканням уваги від іншої атаки. Найбільш розповсюдженими атаками в мережі інтернет, а також найбільш простими та, водночас, найбільш ефективними серед усіх методів зараження пристроїв є атаки розповсюдження стороннього програмного забезпечення [3], які використовують, наприклад, для майнінгу криптовалюти, створення ботнета, інформаційного збирання (особистих та банківських даних, в тому числі, паролів).

Планування і реалізація кібератаки складається з кількох етапів. На етапі збору інформації (рис. 2) здійснюється втручання в мережу методами шкідливого коду або використанням навичок, які мають вузький профіль (використання певних мов програмування, вразливих місць специфікацій і стандартів). На наступному етапі відбувається ураження пристрою, який є ціллю кібератаки (рис. 3). Після вдалого враження пристрою відбувається етап маніпуляції з даними та виконання функцій розповсюдження атаки (рис. 4, таблиця 1).

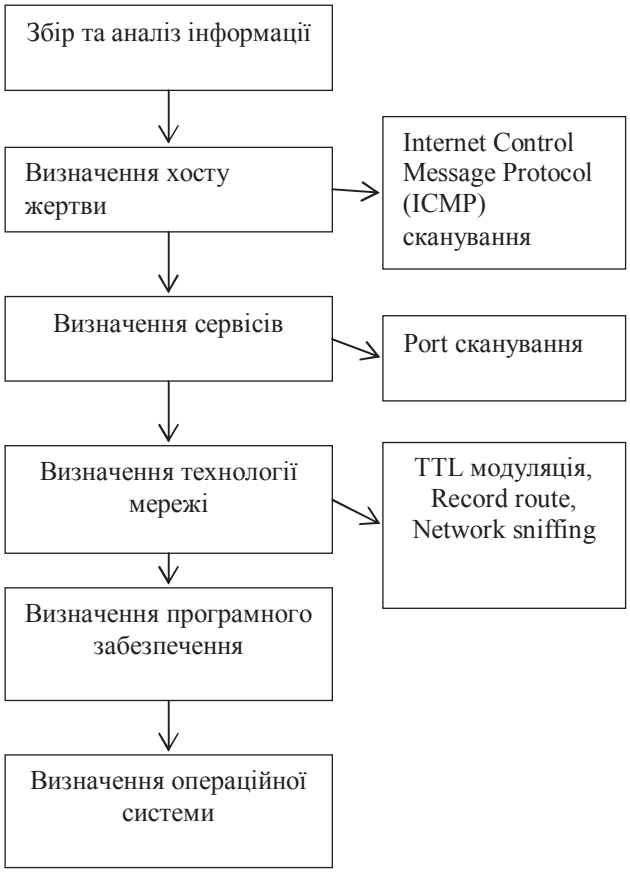


Рис. 2. Етап розвідки. Збір і аналіз інформації.

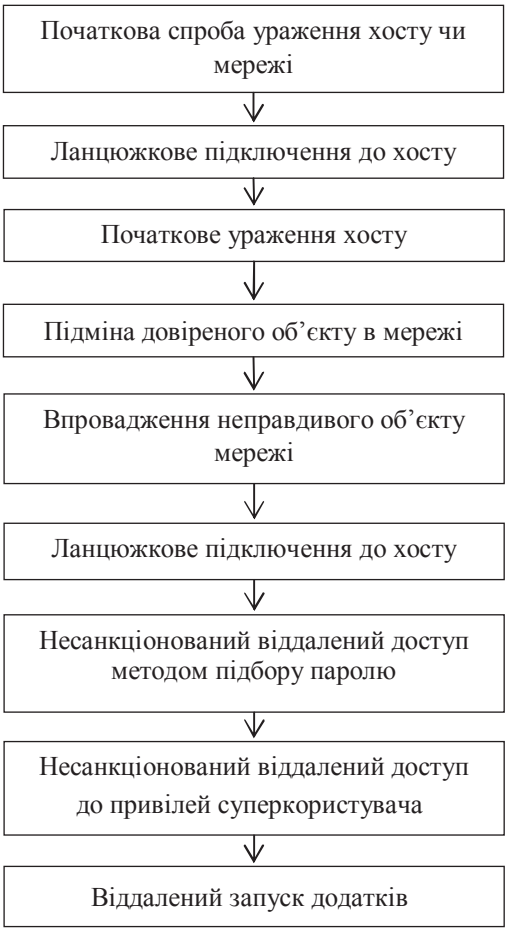


Рис. 3. Етап ураження пристрою чи мережі



Рис. 4. Етап розповсюдження атаки

Таблиця 1

Послідовність дій в ураженій системі

Функція	**Технологія і методи**
Видалення початкових слідів зламу	Самописні методи
Створення відкритого каналу з джерелом атаки	Дешифрування паролів. Встановлення захищеного прихованого каналу
Створення пріоритету процесів	Зараження системи, створення багатьох процесів у ОС
Несанкціоновані дії. Копіювання, видалення, змінення інформації	Використання шифрувальних технологій
Розповсюдження атаки	Дешифрування паролів, атаки через раніше встановлені/відкриті канали

Опис сценарію кібератаки «Petya», «NotPetya». Зразок вірусу має функцію крипто-шифрувальника [2]. Він шифрує файли з конкретним розширенням, потім перезаписує місце MBR (master boot record), очищує журнали (системний каталог), виконує перезавантаження системи. Після цього виводить повідомлення про умови дешифрування, в даному випадку, викуп.

Етап 1 «Зараження». Основним методом розповсюдження вірусу стали масові посилання, котрі розповсюджувались за допомогою електронної пошти [1]. Посилання атаки «фішингу»: користувач відкриває отриманий файл-вкладення на пошті, і виконується інфікування. Наступним кроком розповсюдження є локальний метод технології LAN, спрямований на ураження пристрою-сусіда.

Вірус отримує дані аутентифікації, використовуючи функцію CredEnumerate та допоміжне функціонування mimikatz. Далі вірус розповсюджується по мережі,

запускаючи функції admin $ (рис. 5) та PsExec.exe для віддаленого доступу до виконання команд, і функцію wmic.exe для доступу до інструментарію управління Windows [2]. Оскільки вказані функції доступні лише в сімействі операційних систем Windows, то на інших операційних системах запуск вірусу не результативний. Вірус також намагається отримати доступ до «SMB EternalBlue» і «EternalRomance», перевіряючи вразливі для зараження місця через TCP-порти.

```
wsprintfW(&Name, L"\\\\%s\\admin$", a1);
NetResource.dwScope = 0;
memset(&NetResource.dwType, 0, 0x1Cu);
NetResource.lpRemoteName = &Name;
NetResource.dwType = 1;
sub_10008B70((int)&v23);
wsprintfW(&FileName, L"\\\\%ws\\admin$\\%ws", a1, &v23);
while ( 1 )
{
  pszPath = 0;
  v11 = v4;
  v10 = WNetAddConnection2W(&NetResource, lpPassword, lpUserName, 0);
```

Рис. 5. Виклик функції ОС Windows admin $

При завантаженні вірус перевіряє наявність у системі файлу «C:\Windows\perfc». Якщо файл існує, то вірус завершує виконання (для зупинки зараження вже інфікованого пристрою). Вірус має також систему маскування. Комп'ютер сканується на активні процеси в операційній системі за хеш-кодуванням:

- 0x2E214B44 - «avp.exe» – it's Kaspersky AntiVirus - Kaspersky Internet Security;
- 0x6403527E - «ccSvcHst.exe» – Symantec Service Framework;
- 0x651B3005 - «NS.exe» – Norton Security.

При виконанні цих процесів вірус припиняє своє функціонування. Це виконується через загрозу програми бути ідентифікованою як вірус і видаленою з системи.

Етап 2 «Шифрування». Для шифрування вірус отримує випадкові дані через функцію CryptGenRandom, яка вважається криптографічною стійкою функцією.

Вірус записує у початок диску код-виклику функції, яка запускає ланцюжок шифрування лише після перезавантаження пристрою. Після перезавантаження функція шифрує MFT(головну файлову таблицю) системи, - це база даних, в якій зберігається інформація про вміст дискового розділу з файловою системою NTFS. Приблизний алгоритм шифрування MFT описується за процедурою:

- Зчитується сектор 0x20.
- Встановлюється позначення шифрування MFT.
- EncryptionKey копіюється у тимчасовий буфер.
- Поле з EncryptionKey записується нульовими байтами.
- Сектор 0x20 записується на диск.
- Зчитується сектор 0x21.
- Його вміст шифрується EncryptionKey + нулі.
- Змінений сектор 0x21 перезаписується на диск.

Коли всі файли зашифровані, пристрій знову перезавантажується і виконує текст вимоги викупу з полем для дешифрування. Шифрувальних шифрує файли з такими розширеннями [2]: .3ds .7z .accdb .ai .asp .aspx .avhd .back .bak .c .cfg .conf .cpp .cs .ctl .dbf .disk .djvu .doc .docx .dwg .eml .fdb .gz .h .hdd .kdbx .mail .mdb .msg .nrg .ora .ost .ova .ovf .pdf .php .pmf .ppt .pptx .pst .pvi .py .pyc .rar .rtf .sln .sql .tar .vbox .vbs .vcb .vdi .vfd .vmc .vmdk .vmsd .vmx .vsdx .vsv .work .xls .xlsx .xvd .zip.

Відповідний фрагмент коду представлений на рисунку 6.

```
if ( !(FindFileData.dwFileAttributes & 0x10) || FindFileData.dwFileAttributes & 0x400 )
{
  v5 = (struct _WIN32_FIND_DATAW *)PathFindExtensionW(FindFileData.cFileName);
  if ( (WCHAR *)v5 != &FindFileData.cFileName[wcslen(FindFileData.cFileName)] )
  {
    wsprintfW(&v10, L"%ws.", v5);
    if ( StrStrIW(
           L".3ds.7z.accdb.ai.asp.aspx.avhd.back.bak.c.cfg.conf.cpp.cs.ctl.dbf.disk.djvu.doc.docx.dwg.eml.fdb."
            "gz.h.hdd.kdbx.mail.mdb.msg.nrg.ora.ost.ova.ovf.pdf.php.pmf.ppt.pptx.pst.pvi.py.pyc.rar.rtf.sln.s"
            "ql.tar.vbox.vbs.vcb.vdi.vfd.vmc.vmdk.vmsd.vmx.vsdx.vsv.work.xls.xlsx.xvd.zip.",
           &v10) )
    {
      encrypt_file__sub_1000189A(&FileName, a3);
    }
  }
}
else if ( !StrStrIW(L"C:\\Windows;", &FileName) )
{
  encrypt_files_in_directory__sub_10001973(&FileName, a2 - 1, a3);
}
```

Рис. 6. Фрагмент коду віруса-шифрувальника Not-Petya
(вказано розширення файлів для шифрування)

Алгоритм шифрування схожий на алгоритм шифрувальника Salsa20, котрий виконує операції шифрування конкатенації 32-ти бітних чисел, побітову конкатенацію по модулю(XOR) і побітові зсуви, але був модифікований.

Ключ дешифрування повинен складатися з набору латинських символів і чисел розміром у 32 символи. Хеш-функція, наприклад, функція (SPONGENT), отримує на вхід деяку кількість байт і видає значення у розмірі 32 байта. Після цього алгоритм повторюється 128 разів і результат отримується як EncryptionKey.

Етап 3 «Розповсюдження». Розповсюдження вірусу виконується через виклик вірусом системних функцій «GetExtendedTcpTable», «GetIpNetTable», «NetServerEnum», «WNetEnumResource», «DhcpEnumSubnets», «DhcpEnumSubnetClients» для створення списку мережевих хостів.

Етап 4 «Видалення слідів». Після успішного шифрування вірус очищує системні журнали, викликавши команду WEVTUTIL, і перезавантажує пристрій. Перезавантаження виконується через виклик функцій «NtRaiseHardError», «InitiateSystemShutdownExW» та «ExitWindowsEx».

Висновки. В цілому, сценарій кібератаки складається з таких етапів:

- Атака на пристрій. Спроба стороннього програмного забезпечення отримати доступ до об’єкту атаки. У випадку успіху виконується наступний етап. У випадку невдачі, атака повторюється іншим методом або на інший пристрій.

- Зараження. Безперешкодний запуск шкідливого скрипта чи програми на пристрої. Ураження мережі та виконання команд для отримання прав суперкористувача.
- Приховування слідів зараження (залежить від типів атак). Приховування слідів ураження пристрою для успішного виконання шкідливого коду. Приховування коду від програм захисту/антивірусів.
- Виконання шкідливого коду (маніпуляція з даними). Запуск скрипта і виконання поставленої мети атаки. Шифрування даних, копіювання, видалення. Встановлення допоміжних каналів для розповсюдження зараження.
- Видалення слідів перебування (залежить від типів атак). Отримання результату від атаки, видалення логів та інформації з системних журналів, видалення підозрілих файлів.
- Розповсюдження (залежить від типів атак). Розповсюдження атаки у мережі через встановлені канали передачі. Створення ботнету. Розповсюдження спам-листів і повідомлень з посиланням на файл вірусу.

Виявлення шкідливих програм на комп'ютерному пристрої не дає можливості визначити кінцеву мету атакуючого. Моделі сценаріїв кібератак представляють не тільки вразливості інформаційної системи, які використовують кіберзлочинці, але й визначають послідовність дій, спрямовану на досягнення кінцевої мети. В подальшому моделі сценаріїв кібератак мають бути реалізовані з урахуванням перебігу часу, що надасть можливість оцінювати час розповсюдження кібератаки в розподіленій інформаційній системі.

Список використаних джерел

1. Огородников Д.В. Методи зовнішнього втручання у технології передачі даних Інтернет-речей // Інтернет речей: проблеми правового регулювання та впровадження : Матеріали науково-практичної конференції. 24 жовтня 2017 р., м. Київ. / Упоряд. : В. М.Фурашев, С. Ю. Петряєв. – Київ : Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського» Вид-во «Політехніка». 2017. - С. 101-105.

2. ISSP REPORT "PETYA" – "NOTPETYA" REVERSE ANALYSIS [Електронний ресурс] (2017) – Режим доступу: https://issp.ua/issp_system_images/Petya-NotPetya-Reverse-by-ISSP-Labs.pdf

3. THE WEB APPLICATION SECURITY CONSORTIUM THREAT CLASSIFICATION 01.01.2010 [Електронний ресурс] - Режим доступу: http://projects.webappsec.org/f/WASC-TC-v2_0.pdf

4. Карпінський М. П. Атаки на відмову в обслуговуванні комп'ютерних мереж / М. П. Карпінський, У. О. Яциковська, А. В. Балик, М. Александер // Вісник Національного університету "Львівська політехніка". Комп'ютерні системи та мережі. – 2014. – № 806. – С. 94-99.

5. Hariri S. Impact Analysis of Faults and Attacks in Large-Scale Networks / S. Hariri, G. Qu, T. Dharmagadda, M. Ramkishore, C.S. Raghavendra // IEEE Security and Privacy. - Vol. 1, 2003. - P. 49-54.

6. Common Vulnerabilities and Exposures (CVE) [Електронний ресурс] – Режим доступу: http://cve.mitre.org/.

ДОВІДКА ПРО АВТОРІВ

Стеценко Інна Вячеславівна – доктор технічних наук, професор кафедри автоматизованих систем обробки інформації і управління, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Inna Stetsenko – professor, Departament of Automated Systems of Information Management and Management, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: i.stetsenko@kpi.ua

Огородников Дмитро Володимирович – студент, кафедра автоматизованих систем обробки інформації і управління, Національний технічний університет України «Київський політехнічнийінститут імені Ігоря Сікорського».

Dmytro Ohorodnykov – student, Department of Computer-Aided Management and Data Processing System, ASOIU,National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail:swarty16@gmail.com

Inna Stetsenko, Dmytro Ohorodnykov

MODELING SCENARIO OF CYBERATTACK BASED ON “PETYA”, “NOTPETYA” ATTACK

Target settings. Illegal actions aimed at violating the safety of an individual, a society or a state, methods used to relate material objects-obtaining unauthorized access, altering or breaking the integrity of information to gain an advantage in solving political economic or social problems are now defined as cyberterrorism. In consequence, the actual task to prevent attacks is to simulate the steps of actions that are performed at the attack.

Actual scientific researches and issues analysis. With the tendency of increasing the amount of information in the Internet network, the protection of information from unauthorized persons became a very topical issue. There are many articles that describe the specific attacks that vulnerabilities use, what they do with the information after a successful access to it, but the overall picture of the attack scenarios is not described.

Uninvestigated parts of general matters defining. There are no issues in the schema that describes the life cycle of attacks from attempted damage to successful lesion and data manipulation.

The research objective. The purpose of the study is to create a cyberattack script based on the popular Petya and NotPetya encryption viruses.

The statement of basic materials. The life cycle of the cyberattack's "Notpetya" was studied, on the basis of which a cyberattack script was created, by type of virus encryption-extortionists.

Conclusions. Was created the classification scheme of cyber attacks, attacks were divided into classes, simulated cyberattack life cycle, created cyberattack scene.

The script for cyberattacks consists of the following steps:

• Attack on the device. An attempt by third-party software to access an attack object. In case of success, the next stage is fulfilled. In the event of a failure, the attempt is repeated by another method or another device.

• Infection. Unhindered launch of a malicious script or program on your device. Defeat the network and execute commands to obtain superuser privileges.

• Hiding traces of infection (depending on the types of attacks). Hiding traces of device damage to successfully execute malicious code. Hiding code from security / antivirus software.

• Fulfilling the malicious code (manipulation with data). Run the script and execute the target attack target. Data encryption, copying, deletion. Installation of auxiliary channels for the spread of infection.

• Removing traces of stay. (depends on the types of attacks). Obtaining the result from an attack, deleting logs and information from system logs, removing suspicious files.

• Distribution. (depends on the types of attacks). Distribution of attack on the network through established channels of transmission. Create a botnet network. Distribution of spam emails and messages with a link to a virus file.

Detection of malware on a computer device does not provide an opportunity to determine the ultimate goal of the attacker. The models of cyberattack scenarios represent not only the vulnerabilities of the information system that use cybercriminals, but also determine the sequence of actions aimed at achieving the ultimate goal.

In the subsequent model scenario, cyberattacks should be implemented withtaking into account the current time, which will give an opportunity to estimate the time. Distribution of cyberattacks in distributed information system.

Keywords:cyberattack, information systems, computer virus, life cycle, encryptor.

УДК 004.056

Одинець Т. А.,
Остапченко К. Б.

ЗАГРОЗИ БЕЗПЕКИ В ХМАРНИХ СИСТЕМАХ ТА ШЛЯХИ ЇХ УНИКНЕННЯ

У роботі розглядається проблема загроз безпеки даних, якими оперують додатки та сервіси, що працюють в хмарних системах, а також даних, які безпосередньо зберігаються в хмарних сховищах даних. Виділяються основні види загроз та наводяться рекомендації щодо їх мінімізації.

Ключові слова: хмарні системи, компрометація даних, IT-інфраструктура, кібератака.

Бібл.: 3.

The article deals with the problem of data security threats that is managed by applications and services running on cloud systems, as well as data that is directly stored in cloud storage repositories. In the scope of this article the main types of threats and recommendations for their minimization are described.

Key words: cloud systems, data compromise, IT-infrastructure, cyberattack.

Bibl.: 3.

Актуальність теми дослідження. З огляду на те, як хмарні системи з кожним днем стають все більш поширеними і все більше людей користуються такими сервісами як Google Drive, Dropbox, Amazon Drive, Microsoft OneDrive, постає питання безпеки їх використання. Адже з кожним днем з'являється все більше інформації про прецеденти крадіжки персональних даних та файлів, що зберігались в хмарних системах, одним з яких є атака на акаунт Tesla в хмарі Amazon Web Services. Такі випадки змушують сумніватися в безпеці використання хмарних систем.

Постановка проблеми. Хмарні провайдери, серед яких є такі світові IT-гіганти як Google, Microsoft, Amazon, створюють свої системи дуже доступними як для звичайного користувача, так і для транснаціональної корпорації. А величезна кількість даних, які розміщуються в хмарних системах, робить їх привабливою цілю для хакерів. У цій роботі розглядаються п'ять найбільш поширених загроз безпеки, з якими кінцевий користувач може зіткнутися при зберіганні даних на хмарі, або ж хостингу на ній програмних додатків.

Аналіз останніх досліджень і публікацій. В зв'язку з ростом кількості випадків втрати даних в хмарних системах, з'являється все більше досліджень в цій області та наукових публікацій. Зокрема, провідну роль у таких дослідженнях приймає Cloud Security Alliance (CSA) – некомерційна організація, що займається просуванням найкращих практичних засобів та рекомендацій в розробці систем безпеки на всіх рівнях IT-інфраструктури.

Виділення недосліджених частин загальної проблеми. Дана робота присвячена вивченню та аналізу найбільш поширених загроз безпеки в хмарних системах. Дослідження сфокусовано на вивченні загроз та способів їх уникнення.

Постановка завдання.Визначити найбільш розповсюджені види загроз безпеки в хмарних системах та сформулюватирекомендації щодо захисту в них даних.

Викладення основного матеріалу. Розглянемо п’ять найбільш розповсюджених загроз безпеки в хмарах та наведемо методи їх усунення.

Загроза № 1: Вразливість API хмарних сервісів.

Більшість хмарних сервісів і додатків використовують API для взаємодії з іншими хмарними службами. В результаті безпека API-інтерфейсів безпосередньо впливає на безпеку хмарних сервісів. Більш того, можливість злому збільшується, коли компанії надають стороннім користувачам доступ до API.

На сьогоднішній день вже стали стандартами технології для безпечної передачі даних, такі як OAuth, OpenID Connect, SAML та інші. Розробники фреймворків активно впроваджують ці стандарти в свої продукти, усуваючи необхідність для розробників API самим налаштовувати свої системи на відповідність цим стандартам. Проте, незважаючи на це, досі кожен з розроблених API можуть бути вузьким місцем в питаннях доступності, конфіденційності, цілісності та безпеки. У гіршому випадку це може привести до того, що бізнес втратить конфіденційну інформацію, пов'язану з їх клієнтами і іншими сторонами.

Згідно CSA, кращий спосіб захистити себе від такої загрози - це впровадження додатків і систем моделювання загроз в життєвий цикл розробки. Також CSA рекомендує організувати контроль доступу, використовувати інструменти захисту і раннього виявлення загроз [1].

Загроза №2: Атаки на відмову в обслуговуванні (DDoS).

Атаки DDoS загрожують комп'ютерним мережам вже досить тривалий період часу, однак хмарні технології зробили їх більш поширеними. Ці атаки стимулюють велику потужність обробки і впливають на доступність, продуктивність і швидкодію хмарних систем. Найгірше те, що під час такої атаки неможливо нічого зробити для її усунення. Тому, як тільки це станеться, нічого не залишається робити, за винятком того, щоб сидіти і чекати. Звичайно, також доведеться заплатити за додаткове навантаження, викликану атакою, яка, в залежності від рівня серйозності, може призвести до значних фінансових втрат [2].

Більшість сучасних хмарних сервісів мають системи захисту від DDoS-атак, проте вони не дають абсолютної гарантії. Тому, кращим способом запобігти такій атаці є проведення регулярних перевірок безпеки для виявлення вразливостей до DDoS-атак, а також налаштування ефективної системи обміну критичними ресурсами з адміністраторами для своєчасного сповіщення про загрозу.

Загроза №3: Компрометація облікових записів і обхід аутентифікації.

Така загроза обумовлюється ненадійністю паролів з боку користувачів, або ненадійністю системи аутентифікації з боку провайдера хмарного сервісу. Друга причина є більш серйозною проблемою, адже якщо управління ключами шифрування і сертифікатами відбувається неналежним чином, то під загрозу витоку персональних даних та конфіденційної інформації потрапляють всі користувачі сервісу. Крім того, на корпоративному рівні організації нерідко стикаються з проблемами управління правами та дозволами, коли кінцевим користувачам призначаються значно більші повноваження, ніж в дійсності необхідно. Проблема зустрічається і тоді, коли користувач переходить на іншу посаду в рамках компанії або звільняється. Як правило, мало хто поспішає актуалізувати права на доступ згідно з новими ролями користувача. В результаті виникає неконтрольоване вузьке місце в безпеці.

Відомо, що найбільш збитковим витоком даних за останні 5 років стала атака з метою викрадення ідентифікаційних даних клієнтів Anthem Insurance. Атаку оцінили в 10 балів за Індексом критичності, де було скомпрометовано понад 80 мільйонів облікових записів.

Щодо загрози компрометації даних, CSA рекомендує використовувати:

- механізми багатофакторної аутентифікації;
- одноразові паролі;
- токени;
- смарт-карти;
- USB-ключі.

Це дозволить захистити хмарні сервіси, оскільки застосування вище наведених методів ускладнює процес компрометації паролів.

Загроза №4: Цільові кібератаки.

На сьогоднішній день хмарні системи є досить стійкими до такого роду атак, адже за даними досліджень Cisco, кіберзлочинці ще не володіють достатнім набором інструментів для обходу сучасних систем шифрування, які використовуються провайдерами, або ж для їх обходу потрібні колосальні ресурси. Тим не менш, загроза цільової кібератаки не виключена навіть для найбільш сучасних систем, серед яких найбільш вразливими є хмарні SaaS-сервіси (Software as a Service), де користувач не контролює нічого, окрім власних даних. Більш безпечними є IaaS-сервіси (Infrastructure as a Service), де користувач має можливість самостійно слідкувати за рівнем безпеки даних та встановлювати засоби їх захисту.

Варто зазначити, що у випадку, якщо зловмиснику вдалось закріпити власну присутність в цільовій інфраструктурі, його виявлення стає доволі серйозним та важким завданням.

CSA рекомендує проводити спеціалізоване навчання співробітників по розпізнаванню методів зловмисника, використовувати розширені інструменти безпеки, вміти правильно управляти процесами, знати про планові відповідні дії на інциденти, застосовувати профілактичні методи, що підвищують рівень безпеки інфраструктури [1].

Загроза №5: Ризики, пов’язані зі зловживанням хмарними сервісами.

Хмари можуть використовуватися легітимними і нелегітимними організаціями. Мета останніх - використовувати хмарні ресурси для здійснення зловмисних дій: запуску DDoS-атак, відправки спаму, поширення шкідливого контенту та інше. Постачальникам послуг вкрай важливо вміти розпізнавати таких учасників, для чого рекомендується детально вивчати трафік і використовувати інструменти моніторингу хмарних середовищ.

Ще однією причиною для таких загроз є повна міграція даних корпоративної системи до хмарних сервісів. Часто організації ігнорують ризики, які можуть бути завдані через вразливості хмарних систем, не створюючи власних систем резервування. Як результат, виникає недотримання концепції сегментування мережі, що може призвести до повної втрати даних при першій же помилці роботи хмарної системи або атаці на неї [3].

CSA рекомендує використовувати стратегію «безпеки в глибину», впроваджувати механізми багатофакторної аутентифікації, системи виявлення вторгне-

нь, дотримуватися концепції сегментування мережі і принципу надання найменших привілеїв.

Висновки. Сучасні хмарні системи є новим етапом розвитку IT-інфраструктури і вже сьогодні поступово витісняють з ринку старші покоління систем зберігання даних та хостингу додатків за рахунок надання користувачам миттєвого доступу до широкого спектру ресурсів і додатків. Проте, важливо пам'ятати, що всі хмарні провайдери мають вразливі місця та несуть потенційні загрози користувачам. Однак, ці загрози можна мінімізувати, дотримуючись вище наведених рекомендацій.

Список використаних джерел

1. Wei J. Cloud Security Alliance. Security as a Service [Електронний ресурс] / електрон. наук. вид. Security Guidance for Critical Areas of Focus in Cloud. – 2015. – № 4. – С.102–144. – Режим доступу до журн. : https://downloads.cloudsecurityalliance.org/assets/research/security-guidance/security-guidance-v4-FINAL.pdf.

2. Ali M. Security in cloud computing: Opportunities and challenges [Електронний ресурс] : електронневидання Elsevier / Khan S. U., Vasilakos A. V. // An International Journal. – 2017. – № 305. – С. 359–373. – Режимдоступудожурн. : https://pdfs.semanticscholar.org/40e7/adc40b609cd13f9b84a5cc9ae780ffba6ec0.pdf.

3. Zissis D. Addressing cloud computing security issues [Електронний ресурс] / Lekkas D. // Future Generation Computer Systems. – 2012. – № 28. – С.15–43. – Режим доступу до журн. : https://doi.org/10.1016/j.future.2010.12.006.

ДОВІДКА ПРО АВТОРІВ

Остапченко Костянтин Борисович – к.т.н., доцент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Ostapchenko Konstantin – doctor of technical sciences, associate professor, Department of Tehnical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: okb2003@ukr.net

Одинець Тетяна Анатоліївна – студентка, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Odynets Tetiana – student, Department of Tehnical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: todynets@gmail.com

Odynets Tetiana,
Ostapchenko Konstantin

SECURITY THREATS IN CLOUD SYSTEMS AND WAYS OF THEIR DESTRUCTION

Relevance of research topic. Considering how cloud systems become more common every day, the data security issues that is stored on them become more actual. More and more people use services such as Google Drive, Dropbox, Amazon Drive, Microsoft OneDrive and others to save and access their files in an easy way. However, more and more information about the precedents of theft of personal data and files stored in cloud-based systems appears every day, one of which is the attack on the Tesla account in the Amazon Web Services cloud. Such cases make it doubt the safety of cloud systems.

Formulation of the problem. Cloud providers, among which there are such global IT giants as Google, Microsoft and Amazon, create their systems very affordable for both the ordinary user and the multinational corporation. And the huge amount of data stored in cloud systems makes them an attractive target for hackers. This article discusses the five most common security threats that an end-user may face when storing data on a cloud, or host software applications on it.

Analysis of recent research and publications. Due to the growing number of cases of data loss in cloud systems, more research in this area and scientific publications are emerging. In particular, the Cloud Security Alliance (CSA), a nonprofit organization dedicated to promoting best practices in developing security systems at all levels of the IT infrastructure, takes on the leading role in such research.

Selection of unexplored parts of the general problem. This article is devoted to the study and analysis of the most common security threats in cloud systems. The research focuses on the study of threats and ways to avoid them.

Setting objectives. Identify the most common types of threats and describe how to protect your data.

Presentation of the main material. The analysis and definition of the 5 most common security threats in the clouds were conducted, as well as methods and tools for their elimination.

Conclusions. Modern cloud systems are a new stage in the development of IT infrastructure and are now gradually replacing older generations of storage and hosting systems from the market by providing users with instant access to a wide range of resources and applications. But it's important to remember that all cloud providers have vulnerabilities and potential threats to users. However, these threats can be minimized by following the above recommendations.

Key words: cloud systems, data compromise, IT-infrastructure, cyberattack.

УДК 004.056

Троценко В. В. ,
Остапченко К. Б.

ЗАХОДИ БЕЗПЕКИ В СИСТЕМАХ З ІНТЕРНЕТУ РЕЧЕЙ

У даній роботі розглядається проблема безпеки даних, якими оперують пристрої в IoT-системах, а також пропонуються базові принципи щодо проектування, розробки та інтеграції IoT-продуктів для розробників, вендорів та кінцевих користувачів таких продуктів.

Ключові слова: інтернет речей, багатофакторна аутентифікація, шифрування даних, динамічне тестування.

Бібл.: 3.

The article describes the problem of data security used by devices in IoT systems and offers the basic principles for designing, development and integrating IoT-products for developers, vendors and end-users of such products.

Key words: Internet of things, multifactor authentication, data encryption, dynamic testing.

Bibl.: 3.

Актуальність теми дослідження. Останнім часом все більше систем Інтернету речей інтегрується в середовище нашого побуту, що передбачає обмін даними через пристрої цих систем. З ростом обсягу даних, що передаються та обробляються в таких системах постає питання безпеки доступу до них. Метою даного дослідження є представлення набору правил розробки систем Інтернету речей для досягнення високого рівня безпеки таких розробок, які можуть використовуватися як основи майбутніх стандартів та сертифікатів. Більшість з цих правил можуть бути застосованими до будь-якого пристрою з інтернет-з'єднанням; однак, дана робота зосереджена на безпеці та заходах конфіденційності для IoT [1].

Постановка проблеми. Деякі виробники випускають та продають пристрої IoT, які не містять достатнього набору функцій безпеки. Це призводить до серйозних збитків, як економічних, так іморальних, конкретним особам та організаціям, що використовують IoT системи. Останній приклад цього - відеореєстратори та IP-камери, які зараз виготовляються компанією XiongMai Technologies. Оскільки пристрої IoT набувають великої популярності, якщо не вжити певних заходів для захисту цих пристроїв, то масштаб завданих збитків в майбутньому може бути ще більшим [3].

Аналіз останніх досліджень і публікацій. Стрімкий розвиток та розповсюдження IoT-технологій та систем, розроблених на їх базі, спричинив багато запитань до безпеки функціонування таких систем. В зв’язку з цим, останнім часом організації, що розробляють світові стандарти якості програмного та апаратного забезпечень, зокрема IEEE, проводять багато досліджень у цій галузі та публікують їх результати.

Виділення недосліджених частин загальної проблеми. В даній роботі пропонуються загальні принципи до проектування та розробки IoT-систем та пристроїв з високим рівнем безпеки і захищеності даних, обмін якими відбувається в таких системах.

Постановка завдання. На даний час, оскільки різні запобіжні заходи безпеки є обумовленими специфікою області застосування та визначаються в різних умовах по-різному, неможливо визначити набір універсальних правил для безпеки IoT. Однак важливо сформулювати набір загальних принципів для заходів безпеки та конфіденційності при розробці та інтеграції IoT систем.

Викладення основного матеріалу. В рамках даної роботи пропонується набір рекомендацій щодо забезпечення високого рівня безпеки IoT систем, згрупованих за призначенням:

– Безпека пристроїв;
– Безпека мереж;
– Безпека IoT-систем в цілому.

Безпека пристроїв.

1. Своєчасне встановлення оновлень прошивки

На сьогоднішній день постачальники пристроїв та виробники не мають достатніх фінансових стимулів для забезпечення IoT-пристроїв постійними оновленнями, на відміну від розробників програмного забезпечення, оскільки дохід надходить виключно від продажу пристрою, а не від технічного обслуговування. Проте, такі оновлення мають місце у разі виявлення того чи іншого недоліку в програмному забезпеченні пристрою, тому що вони є закріпленими в гарантійному документі на законодавчому рівні. Саме тому будь-яке оновлення прошивки пристроїв має бути обов'язково встановленим.

2. Динамічне тестування

Важливо, щоб пристрої IoT пройшли ретельне тестування та встановили мінімальну базу метрик, що відображають достатній рівень безпеки. Статичне тестування IoT пристроїв не призначене та не використовується для виявлення вразливостей, які існують в компонентах, що не входять до складу апаратної частини, що була інтегрована в пристрій, таких як процесори та пам'ять. Тобто, за допомогою статичного тестування перевіряються тільки апаратні компоненти пристрою, що відповідають за передачу даних, а не ефективність їх обробки [2].

В той же час, динамічне тестування здатне виявляти як слабкі сторони коду, так і будь-які дефекти чи уразливості, що виникають внаслідок апаратного забезпечення, і які можуть бути невидимими для статичного аналізу. Динамічне тестування може виявити вразливості, які створюються, коли новий код використовується на старих процесорах. Саме тому, рекомендується виробникам, які купують апаратне і програмне забезпечення від інших, здійснювати динамічне тестування, щоб забезпечити безпеку елементів.

3. Розробка процедури для безпечного захисту даних при утилізації пристроїв

Врешті-решт, коли пристрої стають застарілими, користувачі відмовляються від подальшого їх використання з огляду на техніку безпеки. В такому випадку слід провести процес безпечного видалення особистих даних. Це питання конфіденційності, оскільки, якщо пристрій залишається в експлуатації або використовується неправильно, він може слугувати для крадіжки особистої інформації про користувача в екосистемі IoT. Така ж проблема існує і для пристроїв IoT, які продаються другим власникам або стають стандартним обладнанням у будинках та передаються після продажу будинку. Для мінімізації

ризиків, пов'язаних з такою проблемою, виробникам необхідно підготувати рекомендаціїдля користувачів, які описують процедуру безпечного видалення даних з застарілих пристроїв IoT. Промислова практика в корпоративних системах передбачає політику "скинути, переробити або знищити" (DRD) з періодичним переглядом плану, щоб визначити, які пристрої потребують утилізації та як розпоряджатися ними. Деякі виробники закликають користувачів проводити такі процедури безпосередньо через службу підтримки виробника [1].

Безпека мережі.

1. Використання складного процесу аутентифікації

У пристроях IoT не слід використовувати прості значення для імені користувача та пароля, такі як admin / admin. Пристрої не повинні використовувати облікові дані за замовчанням, які є інваріантними на кількох пристроях, і не повинні надавати змогу переходити до налагоджувального режиму, оскільки в такому випадку є можливість отримання доступу до інших пристроїв системи, через інтерфейси інших .

Рекомендується використовувати двофакторну аутентифікацію (2FA), яка вимагає від користувача вживати як пароль, так і іншу форму аутентифікації, яка не залежить від знань користувача, наприклад, випадковий код, що генерується за допомогою SMS-повідомлень. Для IoT-додатків особливо рекомендується використання контекстної аутентифікації (CAA), також відомої як адаптивна аутентифікація, в якій контекстна інформація та алгоритми машинного навчання постійно оцінюють ризик зламування. Якщо ризик є високим, то система не дозволить вхід без багатофакторного токену [1].

2. Використання складного шифрування та безпечних протоколів

Навіть якщо паролі пристроїв є безпечними, зв'язок між пристроями може бути вразливою. У IoT існує багато протоколів, включаючи Bluetooth, Zigbee, Z Wave, 6LoWPAN, Thread, Wi-Fi, мобільний зв'язок, NFC, Sigfox, Neul і LoRaWAN. Залежно від протоколу та доступних обчислювальних ресурсів, пристрій може бути більш-менш здатним використовувати високорівневе шифрування. При розробці сучасних IoT систем рекомендується застосовувати найсильніші методи та алгоритми шифрування, переважно IPsec та / або TLS / SSL [2].

Проте, бувають випадки, коли шифрування не є бажаним, наприклад, в SAE J2735 Основні повідомлення про безпеку (BSM), автомобілі бездротового зв'язку можуть комунікувати один з одним, щоб уникнути зіткнень. У цих випадках повідомлення можуть бути відправлені відкритими, тобто, без шифрування, та перевірені за допомогою цифрових підписів. Проте слід враховувати наслідки пропуску шифрування. Якщо дані передаються незашифрованими та непідписаними, слід дотримуватись запобіжних заходів фільтрації отриманих даних, щоб переконатись, що помилково отримані дані не можуть нанести шкоди.

3. Зменшення пропускної здатності пристрою

Останнім часом в системах, які функціонують на базі IoT-пристроїв атаки DDoS стали досить частими прецедентами. Більшість пристроїв IoT складаються з товарних компонентів, які значно перевищують можливості мережі для функції, яку вони повинні виконувати, що спричиняє затори з запитів в домашніх мережах і потенційно може призвести до величезних втрат даних. Для прикладу припустимо, що кожен IoT-пристрій здатний генерувати трафік на рівні ліній, еквівалент-

ний Gigabit Ethernet (81,274 - 1,488,096 кадра в секунду), наприклад, система ARM9 на одному чіпі (SoC) має два таких вбудованих з'єднання. Таким чином, встановивши понад 10 IoT-пристроїв в систему та ввімкнувши їх в роботу на повну потужність, ми отримаємо 80% вірогідність створення DDoS-атаки всередині самої системи.

Не виключаються також і DDoS-атаки, спричинені зловмисниками. Враховуючи значну пропускну здатність IoT-пристроїв, вони легко можуть пропускати таку кількість трафіку, серед якого шкідливий в режимі реального часу виявити є дуже складним завданням.

Розробникам рекомендується обмежувати кількість мережевого трафіку, що можуть пропускати пристрої IoT на програмному рівні. Незважаючи на високи запас потужності обробки пакетів даних, якими комунікують IoT-пристрої в системах, доцільно знижувати та регулювати його з метою забезпечення ефективної та стабільної їх роботи. Також розробникам рекомендується використовувати обмеження пропускної здатності на апаратному рівні та рівні ядра процесору, щоб збільшити швидкість передачі мережі до стабільного рівня. Такі обмеження ускладнюють для зловмисників використання пристрою під час атаки DDoS, навіть якщо він повністю скомпрометував це. Окрім того, пристрої слід запрограмувати для самоконтролю при незвичній поведінці та відновлювати себе до заводських налаштувань при виявленні тривожної поведінки. Якщо перезавантаження пристроїв до заводських налаштувань неможливо, пристрої потрібно принаймні перезавантажити, щоб потенційно очистити код, який зловмисник запускає в пам'яті [3].

Безпека IoT-системи в цілому.

1. Захист конфіденційної інформації

Основна ідея IoT полягає в тому, щоб з'єднувати побутові об'єкти через Інтернет або спеціальну внутрішню мережу. Пристрої IoT надають інтерфейси для комунікації, які можуть бути виявленими іншими пристроями IoT. Більшість протоколів є вразливими та можуть призвести до витоку особистої інформації, такої як ім'я власника або інші персональні дані. Не виключено, що дані, що були викрадені можуть бути пов'язані з іншими джерелами інформації для цільових атак. Рекомендується використовувати сервісні механізми та захищені протоколи аутентифікації й передачі даних, щоб пристрій могли виявити виключно авторизовані клієнти [1].

2. Інститут сертифікації IoT

Через поширені проблеми з безпекою та конфіденційністю, які вже спричинені пристроями IoT, поступово створюється технічна база, на основі якої розробники беруть на себе відповідальність за свої розробки. IEEE або деяка міжнародна організація надають професійну програму сертифікації для проектувальників, розробників та провайдерів нових технологій IoT, які зобов'язуються дотримуватися визнаних правил створення нових пристроїв.

Проте, на сьогоднішній день цих заходів недостатньо і орган сертифікації повинен перевіряти принаймні наступні елементи продукту постачальника:

- Використовувані або рекомендовані протоколи, за якими було розроблено продукт, не мають містити інформації про користувачів, окрім необхідних для функціонування пристрою;
- Коли виникають проблеми з конфіденційністю, сертифікований постачальник має негайно реагувати на проблеми;

- Аутентифікація має бути достатньо комплексною та відповідати перевіреним протоколам;
- Пристрої не мають перевантажуватись та повинні бути захищеними відповідно до нормативних документів;
- Пристрої повинні мати ідентифікаційну мітку, яку неможливо легко підробити, і яка містить веб-посилання, на якому клієнти можуть шукати статус сертифікації пристрою разом із описом пристрою (модель та серійний номер тощо).

Висновки. В даній роботі було розглянуто набір загальних принципів, на основі яких мають будуватись сучасні IoT-пристрої та системи. Наведено загальні рекомендацію для розробки та експлуатації таких пристроїв та систем як на апаратному, так і на програмному рівнях. Дотримуючись даних рекомендацій, розробники та постачальники IoT-продуктів, мінімізують як економічні, так і репутаційні ризики, пов'язані із безпекою своїх продуктів.

Список використаних джерел

1. Stankovic J. A. Research Directions for the Internet of Things [Електронний ресурс] / IEEE INTERNET OF THINGS JOURNAL. – 2014. – №1. – С.1–7. – Режимдоступудожурн. : https://ru.scribd.com/document/354442742/IEEE-Research-Directions-for-the-Internet-of-Things.

2. Hua-Dong Ma. Internet of Things: Objectives and Scientific Challenges [Електронний ресурс] / Journal of Computer Science and Technology. – 2011. – №26. – С.919–920. – Режим доступу до журн. : https://link.springer.com/article/10.1007%2Fs11390-011-1189-5.

3. Baker F. Broadband Internet Technical Advisory Group. [Електронний ресурс] : Internet of Things (IoT) Security and Privacy Recommendations. – 2016. – №1. – С.17–41. – Режим доступу до журн. : https://www.bitag.org/documents/BITAG_Report_-_Internet_of_Things_(IoT)_Security_and_Privacy_Recommendations.pdf.

ДОВІДКА ПРО АВТОРІВ

Остапченко Костянтин Борисович – к.т.н., доцент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Ostapchenko Konstantin – doctor of technical sciences, associate professor, Department of Tehnical Cybernetics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: okb2003@ukr.net

Троценко Владислав Вікторович – студент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Trotsenko Vladyslav – student, Department of Tehnical Cybernetics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: vtrotsenko1@gmail.com

Trotsenko Vladyslav, Ostapchenko Konstantin

SECURITY MEASURES IN IOT SYSTEMS

Relevance of research topic. Recently, more and more systems of the Internet are integrated into the environment of our everyday life, which involves the exchange of data through the devices of these systems. With the increase in the amount of data transmitted and processed in such a system raises the issue of security access to them. The purpose of this study is to present a set of Internet safety principles (IoT), rules for developing Internet things to achieve a high level of security and best practices for such developments that can be used as the basis for future standards and certificates. Most of these rules can be applied to any device with an Internet connection; However, this article focuses on security and privacy practices for IT.

Formulation of the problem. Some manufacturers produce and sell IoT devices that do not have a sufficient set of security features. This leads to serious damage, both economic and moral, to specific individuals and organizations that use the IoT system. The latest example is video recorders and IP cameras now mentioned by XiongMai Technologies. Because IoT devices are becoming popular, if you do not take certain measures to protect these devices, the scale of the damage done in the future may be even greater.

Analysis of recent research and publications. The rapid development and spread of IoT technologies and systems developed on their basis has raised many questions about the security of the functioning of such systems. In this regard, organizations that develop world-class standards for the quality of software and hardware, such as IEEE, have been conducting a lot of research in this area and publishing them.

Selection of unexplored parts of the general problem. This article provides general principles for the design and development of IoT systems and devices with a high level of security and security of data exchange that occurs in such systems.

Setting objectives. At the moment, since various precautionary measures are due to the specifics of the scope and are defined in different conditions in different ways, it is impossible to define a set of universal rules for the security of the IoT. However, it is important to describe a set of general principles and best practices for security and privacy measures in the development and integration of IoT systems.

Presentation of the main material. Within the framework of this article a set of recommendations for ensuring a high level of safety of IoT systems grouped by appointment is proposed.

Conclusions. In this paper, a set of general principles on which modern IoT devices and systems should be built are considered.. Following these recommendations, developers and vendors of IoT products minimize both the economic and reputational risks associated with the safety of their products.

Key words: Internet of things, multifactor authentication, data encryption, dynamic testing.

УДК 004.4

Подтьопа Сергій

ПОМИЛКИ ПРИ РОБОТІ З WEB-СИСТЕМАМИ АВТЕНТИФІКАЦІЇ

У статті розглядаються основні помилки починаючих та досвідчених web-розробників при роботі з системами автентифікації. Розглянуті помилки актуальні для всіх мов програмування, хоча стаття зосереджена над мовою JavaScript і її бібліотеками.

Ключові слова: автентифікація, web-додаток, JavaScript, захист даних.
Бібл.: 13.

The paper deals with the issues the major errors of beginning and experienced web-developers when working with authentication systems. The bugs discussed are relevant to all programming languages, although the paper deals focuses on the language of JavaScript and its libraries.

Key words: authentication, web-application, JavaScript, data protection
Bibl.: 13.

Актуальність теми дослідження. У зв'язку з активним розвитком web-індустрії, зростає потреба в якісних системах автентифікації для web-ресурсів. На сам перед дана стаття буде актуальна для JS-розробників, але помилки вказані в ній часто зустрічаються і в інших платформах.

Постанова проблеми. Часто при роботі з системами автентифікації виникають помилки, що призводить до незахищеності даних користувачів.

Аналіз останніх досліджень і публікацій. Незважаючи на значну кількість робіт, присвячених питанню автентифікації та захисту даних в цілому, проблема роботи з web-системами автентифікації залишається мало дослідженою.

Мета дослідження. Метою даного дослідження є виділити основні помилки при роботі з web-системами автентифікації. Стаття буде зосереджена на найбільш розповсюджених помилках та їхньому уникненні при створенні власної системи автентифікації чи використанні існуючої.

Викладення основного матеріалу. На сьогоднішній деньіснує безліч навчальних посібників, які спрямовані на допомогу в установці та налагодженні системи автентифікації для web-додатків. Практично всі вони містять ті чи інші помилки. Жодне з них не дозволяє створити повномасштабне рішення, необхідне для працюючого веб-додатку.

Сховище даних користувачів. Запис і читання облікових даних є доволі звичним завданням для систем по управлінню автентифікацією, традиційним способом вирішення цих завдань є використання власної бази даних. Більшість існуючих систем автентифікації є лише проміжним програмним забезпеченням, що зазвичай просто повідомляє нашому додатку, що користувач пройшов або не пройшов перевірку, вимагаючи відповіді для роботи зі сховищем паролів в локальній базі даних.

Звернувшись до системи Passport.js [1], а саме до посібника, можна побачити, що в прикладах використання, не виявилося підсистеми підтримки бази даних. Приклад мав на увазі просте використання деякого набору акаунтів.

На перший погляд все доволі логічно. Нам наведений в приклад звичайний web-додаток. Але розробник, який скористається ним, через недосвідченість чи брак часу не буде нічого серйозно поліпшувати. Зауваживши, що паролі в прикладі не шифруються, а зберігаються у вигляді тексту поруч з кодом логіки валідації. Про сховище облікових даних в прикладі навіть не згадується. Дуже цікава вийде система автентифікації.

Далі наведений ще один посібник по Passport.js , написаний в 2015 році [2]. Даний посібник використовує Mongoose ODM і читає облікові дані з бази даних. Тут є все, зокрема інтеграційні тести, а також ще один шаблон, що можна використовувати. Проте, Mongoose ODM зберігає паролі, використовуючи тип даних String, як і в попередній інструкції, у вигляді звичайного тексту, тільки на цей раз в екземплярі MongoDB. Добре відомо, що екземпляри MongoDB зазвичай дуже добре захищені.

Звернемось тепер до всіма улюбленого пошукового сервісу. Верхнім результатом є керівництво від scotch.io [3], даний посібник є кращим, в ньому використовується bcrypt з коефіцієнтом трудомісткості рівним 8 для шифрування паролів. Але 8 – це мало, навіть дуже мало. Більшість сучасних bcrypt-бібліотек використовують 12. Коефіцієнт трудомісткості 8 був хороший для адміністративних облікових записів вісімнадцять років тому, відразу після випуску першої специфікації bcrypt.

Якщо навіть не говорити про зберігання секретних даних, жодний з посібників не показує реалізацію механізму скидання паролів. Найважливіша частина системи автентифікації, де багато власних помилок, була залишена в якості завдання для розробника.

Система скидання паролів. Схожа проблема в області безпеки – скидання пароля. Жодний з посібників, що знаходяться у верхній частині пошукового сервісу, абсолютно нічого не говорить про те, як робити це з використанням Passport.

При реалізації системи скидання паролів є тисячі способів все зіпсувати. Ось найпоширеніші помилки в рішенні цієї задачі, що можна зустріти:

1. ***Передбачувані токени.*** Гарним прикладом є токени, засновані на поточному часі. Токен, що побудований на базі поганого генератора псевдорандомних чисел, хоча і виглядає краще, проблему не вирішують теж.

2. ***Погане сховище даних***. Зберігання незашифрованих токенів скидання пароля в базі даних означає, що якщо база буде зламана, ці токени – фактично паролі, що зберігаються у вигляді відкритого тексту. Створення довгих токенів за допомогою криптографічно стійкого генератора псевдовипадкових чисел дозволяє запобігти грубих атак на токени скидання паролів, але не захищає від локальних атак. Токени скидання паролів є обліковими даними і їх слід розглядати як такі.

3. ***Токени, без обмеженого терміну дії.*** Якщо термін дії токенів не закінчується, то атакуючий має час, щоб скористатися тимчасовим вікном скидання пароля.

4. ***Відсутність додаткових перевірок***. Додаткові питання при скиданні пароля – це стандарт верифікації даних де-факто. Звичайно, це працює як треба лише в тому випадку, якщо розробники вибирають хороші питання. У подібних питань є власні проблеми. Тут варто сказати і про використання електронної пошти для відновлення пароля, хоча міркування про це можуть здатися зайвої перестраховкою. Адреса електронної пошти – це те, що у вас є, а не те, що ви

знаєте. Адреса об'єднує різні чинники автентифікації. Як результат, електронна пошти стає ключем до будь-якого облікового запису, яка просто відправляє на нього токен скидання пароля.

Наголосивши загальні питання, перейдемо до конкретики.

Можна звернути увагу на бібліотеку для скидання паролів п'ятирічної давності від чудового видавця substack [4]. З огляду на швидкість розвитку web-технологій, цей пакет нагадує дідуся, розглянувши деталі, можна сказати, що функція Math.random() доволі передбачувана, тому її не слід використовувати для створення токенів, тому йдемо далі.

Отож, повертаємося до пошукового сервісу. Виникає відчуття, що тема, котра цікавить нас, розкрита лише в одному матеріалі. Візьмемо перший результат, знайдений за запитом [5]. Знову зустрічаємо bcrypt, з меншим коефіцієнтом трудомісткості, ніж потрібно в сучасних умовах, рівним 5.

Посібник виглядає досить-таки цілісним в порівнянні з іншими, так як використовує crypto.randomBytes для створення по-справжньому випадкових токенів, термін дії яких закінчується, якщо вони не були використані. Але пункти 2 і 4 з вищенаведеного списку помилок при скиданні пароля в посібнику не враховані. Токени зберігаються ненадійно - згадуємо першу помилку керівництв з автентифікації, пов'язану зі зберіганням облікових даних.

Позитивним є те, що викрадені з такої системи токени мають обмежений термін дії. Проте, ці токени особливо приємні, якщо у атакуючого є доступ до об'єктів користувачів в базі даних через BSON, або може отримати вільний доступ до Mongo через невірну конфігурацію. Атакуючий може просто запустити процес скидання пароля для кожного користувача, прочитати незашифровані токени з бази даних і створити власні паролі для облікових записів користувачів, замість ресурсномістких атак по словнику bcrypt, використовуючи потужний комп'ютер.

Токени API. Токени API є обліковими даними. Вони так само важливі, як паролі та токени скидання паролів. Практично всі розробники знають це і намагаються дуже надійно зберігати свої ключі та коди, однак, програм, над якими вони працюють, це доволі часто не стосується.

Справи складаються так, що в пошуковому сервісі важко знайти вартий уваги матеріал, а з популярних посібників [6] можна відмітити деякі помилки в зберіганні облікових даних:

1. Ключі токенів API зберігаються у вигляді звичайного тексту в наведеній репозиторії.

2. Для зберігання паролів використовується симетричний шифр. Це означає, хтось може заволодіти ключем шифрування і розшифрувати всі паролі. До того ж, тут спостерігаються неправильні взаємини між ключем шифрування і секретним ключем токенів API.

3. Для шифрування даних в сховищі паролів, використовується алгоритм AES-256-CTR. Даний алгоритм взагалі не варто використовувати, і вказаний його варіант нічого не змінює [7].

На жаль більшість посібників [6, 8, 9] мають схожі помилки з захисту облікових даних користувачів та в них спостерігається та ж уразливість, пов'язана з розкриттям інформації.

Обмеження кількості спроб автентифікації. В жодному серед наведених вище посібників зі створення систем автентифікації не знайшлося згадок про обмеження числа спроб автентифікації або про блокування облікового запису.

Без обмеження кількості спроб автентифікації, атакуючий може виконувати онлайн атаку по словнику, в надії отримати доступ до облікового запису зі слабким паролем. Блокування облікового запису так само допомагає у вирішенні цієї проблеми завдяки вимогам введення додаткової інформації при наступних спробах входу в систему.

Слід пам'ятати про те, що обмеження числа спроб автентифікації сприяє доступності сервісу. Так, використання bcrypt створює серйозне навантаження на процесор. Без обмежень, функції, в яких викликають bcrypt, стають вектором відмови в обслуговуванні рівня додатки, особливо при використанні високих значень коефіцієнта трудомісткості. В результаті обробка безлічі запитів на реєстрацію користувачів або на вхід в систему з перевіркою пароля створює високе навантаження на сервер.

Хоча відповідного навчального керівництва на цю тему майже немає, є безліч допоміжних бібліотек для обмеження числа запитів [10, 11, 12]. Важко сказати про рівень безпеки цих модулів, але свою функцію мають цілком виконати.

Висновки. Автентифікація – важка задача. Багато хто може сказати, що посібники – це лише пояснення основ, але недосвідчені читачі вірять словам авторів посібників, у яких набагато більше досвіду, ніж у тих, хто читає керівництва.

Якщо ви – розробник початківець – не довіряйте навчальним посібникам. Копіювання коду з таких матеріалів, часто може, привести до проблем у сфері web-автентифікації. Якщо вам дійсно потрібні надійні, готові до використання в коді бібліотеки для автентифікації, погляньте на щось, чим вам буде зручно користуватися, на те, що володіє більшою стабільністю і краще випробувано часом.

Список використаних джерел

1.Passport.Simple, unobtrusiveauthenticationforNode.js. [Електронний ресурс]. — Режим доступу: http://www.passportjs.org/docs/

2. Michael Herman. User Authentication with Passport and Express 4. [Електронний ресурс]. — Режим доступу:http://mherman.org/blog/2015/01/31/local-authentication-with-passport-and-express-4/

3. Chris Sevilleja. Easy Node Authentication: Setup and Local. [Електронний ресурс]. — Режим доступу:https://scotch.io/tutorials/easy-node-authentication-setup-and-local

4. Substack. Middleware for managing password reset emails. [Електроннийресурс]. — Режим доступу:https://npmjs.com/package/password-reset/

5. Sahat Yalkabov. How To Implement Password Reset In Node.js. [Електронний ресурс]. — Режим доступу:http://sahatyalkabov.com/how-to-implement-password-reset-in-nodejs/

6. Soni Pandey. User Authentication using JWT (JSON Web Token) in Node.js. [Електронний ресурс]. — Режим доступу:https://medium.com/@pandeysoni/user-authentication-using-jwt-json-web-token-in-node-js-using-express-framework-543151a38ea1/

7. Is regular CTR mode vulnerable to any attacks? [Електронний ресурс]. — Режим доступу: https://crypto.stackexchange.com/questions/33846/is-regular-ctr-mode-vulnerable-to-any-attacks/33861#33861

8. JonathanMH. About Express, Passport and JSON Web Token Authentication for Beginners. [Електронний ресурс]. — Режим доступу:https://jonathanmh.com/express-passport-json-web-token-jwt-authentication-beginners/

9. Joshua Slate. Creating a Simple Node/Express API Authentication System with Passport and JWT. [Електронний ресурс]. — Режим доступу:http://blog.slatepeak.com/creating-a-simple-node-express-api-authentication-system-with-passport-and-jwt/

10. ExpressRateLimit. [Електронний ресурс]. — Режим доступу:https://npmjs.com/package/express-rate-limit/

11. Expresslimiter. [Електронний ресурс]. — Режим доступу:https://npmjs.com/package/express-limiter/

12. Adam Pflug. Express-brute. [Електронний ресурс]. — Режим доступу:https://github.com/AdamPflug/express-brute

13. Your Node.js authentication tutorial is (probably) wrong. [Електронний ресурс]. — Режим доступу:https://hackernoon.com/your-node-js-authentication-tutorial-is-wrong-f1a3bf831a46/

ДОВІДКА ПРО АВТОРА

Подтьопа Сергій Андрійович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Podtopa Serhii –student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: seriy_gry@ukr.net

Podtopa Serhii

ERRORS IN WORK WITH WEB SYSTEMS AUTHENTICATION

Relevance of the research. Due to the active development of the web-industry, there is a growing need for high-quality authentication systems for web-resources. The article itself will be relevant to JS developers, but the errors mentioned in it often found on other platforms as well.

Target setting. The lack of comparative studies on selection of a parallelism degree of granularity to solve specific mathematical problems that require significant computation, and tools that implement parallelism.

Actual scientific researches and issues analysis. Despite the large number of works devoted to the issue of authentication and data protection in general, the problem of working with web-system authentication systems remains poorly researched.

The research objective. The purpose of this study is to identify the main errors in working with web-authentication systems. The article will focus on the most common mistakes and avoid them when creating their own authentication system or using an existing one.

The statement of basic materials. The four main errors that occur when working with the authentication system identified. Examples of the occurrence of these errors and the ways of avoiding them are given.

Conclusions. Authentication is a difficult task. If you are a beginner developer, do not trust the tutorials. Copying the code from such materials can often lead to problems in the field of web-authentication.

Key words: authentication, web-application, JavaScript, data protection.

Валерий Симоненко,
Александр Слюсаренко

СИСТЕМА ОБЕСПЕЧЕНИЯ БЕЗОПАСНОСТИ ВЫЧИСЛИТЕЛЬНОГО УЗЛА НА ОСНОВЕ РЕЖИМА АУДИТ

В этой статье рассматривается способ повышения безопасности вычислительного узла с помощью стандартных средств операционных систем – локальных политик безопасности, в частности политик аудита. В журнале безопасности фиксируются все события, определенные политиками аудита, которые задаются для каждого объекта отдельно.

Ключевые слова: локальные политики безопасности, журнал событий безопасности, аудит.

This article discusses how to increase the security of a computing node using standard operating system tools – local security policies, in particular audit policies. The security log records all events defined by audit policies that are specified for each object separately.

Keywords: local security policies, security events log, audit.

Постановка проблемы: Проблема заключается в том, что информация передаваемая через интернет, шифруется различными способами и необходимо отлавливать те вредоносные программы, команды или действия которые пропускают различными средства безопасности, а так же обеспечить защиту от внутренних угроз.

Актуальные научные исследования и анализ проблем: В связи с повышением актуальности задач безопасности вычислительных узлов необходимо решать ряд проблем, связанных с ней, которые не могут решить уже существующие средства.

Выделение неисследованных частей общей проблемы: Несмотря на значительный объем работ, посвященный тематике безопасности вычислительных узлов, многие вопросы остаются без должного рассмотрения. Например, динамический анализ передаваемых данных, поскольку данные шифруются различными способами, то стандартные средства не всегда могут отловить вредоносные программы и команды. Более того, ни одно из этих средств не обеспечивает внутреннюю безопасность узла.

Постановка задания: Задача заключается в том, чтобы повысить уровень безопасности компьютера, отлавливая те вредоносные программы и команды, которые пропускают другие средства безопасности. Например, межсетевые экраны выборочно пропускают трафик, сравнивая его параметры с заданными. Одной из задач, которыми занимаются межсетевые экраны, является защита сегментов сети или отдельных хостов от несанкционированного доступа с использованием уязвимых мест в протоколах сетевой модели OSI или в программном обеспечении, установленном на компьютерах сети. В модемах есть режим тестирования (другого модема). Под видом этого тестирования можно залезть в другой компьютер и считать любую информацию. Соответственно, если на объект, защиту которого осуществляет межсетевой экран, разрешается неограниченный модемный доступ, злоумышленники могут его обойти.

Для решения этих проблем можно использовать стандартные средства журналирования событий в операционных системах – локальные политики безопасности, в частности политику аудита.

Изложение основного материала: Технически, аудит безопасности в Windows реализуется через настройку политик аудита и настройку аудита объектов. Политика аудита определяет, какие события и для каких объектов будут генерироваться в журнал событий Безопасность. Регулярный анализ данных журнала безопасности относится к организационным мерам, для поддержки которых может применяться различное программное обеспечение. В самом простом случае можно обходиться приложением Просмотр событий.

Для автоматизации задач анализа событий безопасности могут применяться более продвинутые программы и системы управления событиями безопасности (Security Information and Event Management), обеспечивающие постоянный контроль журналов безопасности, обнаружение новых событий, их классификацию, оповещение специалистов при обнаружении критических событий. Журнал безопасности представляет собой базу данных или файл, в котором регистрируются события, связанные с безопасностью системы. Благодаря системе аудита, администратор может узнать, кто, каким образом и когда воспользовался (или пытался воспользоваться, но получил отказ в доступе) интересующими его ресурсами.

Настройка средств аудита позволяет выбрать типы событий, подлежащих регистрации, и определить, какие именно параметры будут регистрироваться.

Наиболее общими типами событий для аудита безопасности являются:

1) доступ к файлам и каталогам
2) управление учетными записями пользователей и групп
3) вход пользователей в систему и выход из неё

Как правило, фиксируются следующие параметры, касающиеся действий, совершаемых пользователями:

1) выполненное действие
2) идентификаторы пользователей и групп, выполнивших действие
3) дата и время выполнения

Аудит приводит к дополнительной нагрузке на систему, поэтому необходимо регистрировать лишь события, действительно представляющие значение с точки зрения безопасности.

Предусмотрены два уровня аудита безопасности:

- Уровень авторизации службы, на котором производится авторизация вызывающего абонента.
- Уровень сообщений, на котором WCF проверяет допустимость сообщений и проверяет подлинность вызывающего абонента.

Можно проверить результаты аудита обеих уровней (успех или ошибка), что называется *поведением аудита*.

В Windows 7 Максимальная локальная политика безопасносит предусматривает управление следующими политиками аудита:

- Аудит входа в систему
- Аудит доступа к объектам
- Аудит доступа к службе каталогов

- Аудит изменения политики
- Аудит использования привилегий
- Аудит отслеживания процессов
- Аудит системных событий
- Аудит событий входа в систему
- Аудит управления учетными записями

Аудит отслеживания процессов – этот параметр безопасности определяет, будет ли операционная система выполнять аудит событий, связанных с процессами, такими как создание процесса, завершение, обработка дублирований, а также непрямой доступ к объектам. Если этот параметр политики определен, администратор может задать аудит только успехов, только неудач, успехов и неудач либо отключить аудит этих событий совсем. Если включен аудит успехов, запись аудита создается при каждом успешном отслеживании операционной системой действий, связанных с процессами. Если включен аудит неудач, запись аудита создается каждый раз, когда операционная система не может выполнить одно из этих действий. Поведение аудита можно задавать либо путем программирования, либо через конфигурацию.

Выводы: Межсетевой экран, не может обеспечить защиту от внутренних угроз. Хотя firewall можно разработать так, чтобы предотвратить получение конфиденциальных данных внешними нарушителями, он все равно не запретит внутренним пользователям копировать данные. Из этого следует, что было бы заблуждением думать, что наличие firewall обеспечит защиту от внутренних атак. Firewall не защищают от загрузки пользователями зараженных вирусами программ из Интерента или от передачи таких программ по средством электронной почты. Кроме этого, потенциально узкое место представляет собой пропускная способность firewall, так как все соединения должны осуществляться только через него, а в некоторых случаях, кроме всего прочего, еще и подвергаться фильтрации.Аудит событий безопасности позволит повысить безопасность компьютера, поскольку операционная система может отслеживать вставки кода вредоносных программ из интернета, а с помощью журналирования событий позволит так же обезопаситься от внутренних угроз.

Список використаних джерел

1. Преимущества и недостатки использования брандмауера [Электронный ресурс http://clubwindows.ru/?p=51575], дата визита 05.12.2017
2. Межсетевой экран: недостатки и преимущества использования [Электронный ресурс https://v-bezopasnosti.ru/stati/mezhsetevoj-ekran-nedostatki-i-preimushhestva-ispolzovaniya/], дата визита 05.12.2017
3. Аудит событий безопасности [Электронный ресурс https://msdn.microsoft.com/ru-ru/library/ms731669(v=vs.110).aspx], дата визита 29.11.2017

ДОВІДКА ПРО АВТОРІВ

Сімоненко Валерій Павлович – професор, доктор технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут ім. ІгоряСікорського».

Simonenko Valery Pavlovich – Professor, Doctor of Technical Sciences, Department of Computer Engineering, National Technical University of Ukraine "Kiev Polytechnic Institute. Igor Sikorsky. "
E-mail: svp@comsys.kpi.ua

Слюсаренко Олександр Євгенович – студент 4 курсу кафедри обчислювальної техніки, НаціональнийтехнічнийуніверситетУкраїни «Київський політехнічний інститут ім. ІгоряСікорського».
Slyusarenko Alexander Evgenievich - 4th year student of the Department of Computer Science, National Technical University of Ukraine "Kyiv Polytechnic Institute. Igor Sikorsky. "
E-mail: pureanddivine@gmail.com

Сімоненко В. П.,
Слюсаренко А. Є.

СИСТЕМА ЗАБЕЗПЕЧЕННЯ БЕЗПЕКИ ОБЧИСЛЮВАЛЬНОГО УЗЛА НА ОСНОВІ РЕЖИМУ АУДИТ

Постановка проблеми: Проблема полягає в тому, що інформація передається через інтернет, шифрується різними способами і необхідно відловлювати ті шкідливі програми, команди або дії які пропускають різними засоби безпеки, а також забезпечити захист від внутрішніх загроз.

Актуальні наукові дослідження і аналіз проблем: У зв'язку з підвищенням актуальності завдань безпеки обчислювальних вузлів необхідно вирішувати ряд проблем, пов'язаних з нею, які не можуть вирішити вже існуючі засоби.

Виділення недосліджених частин загальної проблеми: Не дивлячись на значний обсяг робіт, присвячений тематиці безпеки обчислювальних вузлів, багато питань залишаються без належного розгляду. Наприклад, динамічний аналіз даних, що передаються, оскільки дані шифруються різними способами, то стандартні засоби не завжди можуть відловити шкідливі програми і команди. Більш того, жодна з цих коштів не забезпечує внутрішню безпеку вузла.

Постановка завдання: Завдання полягає в тому, щоб підвищити рівень безпеки комп'ютера, отлавливая ті шкідливі програми і команди, які пропускають інші засоби безпеки. Для вирішення цих проблем можна використовувати стандартні засоби журналирования подій в операційних системах – локальні політики безпеки, зокрема політику аудиту.

Изложение основного материала:Технічно, аудит безпеки в Windows реалізується через настройку політик аудиту і настройку аудиту об'єктів. Політика аудиту визначає, які події і для яких об'єктів будуть генеруватися в журнал подій Безпека. Регулярний аналіз даних журналу безпеки відноситься до організаційних заходів, для підтримки яких може застосовуватися різне програмне забезпечення. У найпростішому випадку можна обходитися додатком Перегляд подій.

Висновки: Мережеві екрани не захищають від завантаження користувачами заражених вірусами програм з інтерент або від передачі таких програм по засобом електронної пошти. Аудит подій безпеки дозволить підвищити безпеку комп'ютера, оскільки операційна система може відстежувати вставки коду шкідливих програм з інтернету, а за допомогою журналювання подій дозволить так само захиститися від внутрішніх загроз.

Ключові слова: локальні політики безпеки, журнал подій безпеки, аудит.

УДК 004.4

Тизунь Віталій

ВРАЗЛИВОСТІ БЕЗПЕКИ ТА КОНФІДЕНЦІЙНОСТІ ІНФОРМАЦІЇ ПРИ ВИКОРИСТАННІ ПОПУЛЯРНИХ JAVASCRIPT БІБЛІОТЕК

У статті розглядаються основні вразливості безпеки та конфіденційності інформації на web-сервісах, пов'язані з використанням популярних JavaScript бібліотек. Особливо увагу звернено на вразливості, що приводять до витоку користувацької інформації.

Ключові слова: захист даних, web-сервіс, JavaScript, вразливість безпеки.
Бібл.: 13.

The article discusses about the main vulnerabilities of security and privacy of information on web services associated with the use of popular JavaScript libraries. Particular attention is paid to vulnerabilities that lead to the leak of user information.

Key words: data protection, web-service, JavaScript, security vulnerability.
Bibl.: 13.

Актуальність теми дослідження. При розробці багатьох web-сервісів для полегшення самого завдання розробки часто використовуються сторонні бібліотеки, які не завжди гарантують те, що реалізовані в них методи будуть абсолютно безпечними для використання. Дана проблема є доволі актуальною, так як навіть в популярних бібліотеках, що використовуються багатьма ресурсами часто знаходять вразливості.

Постанова проблеми. При роботі з сторонніми бібліотеками розробник ресурсу може не знати про вразливості, що містять певні методи використовуваної бібліотеки.

Аналіз останніх досліджень і публікацій. При дослідженнях великої кількості web-ресурсів було виявлено, що більшість із них містять вразливості, пов'язані з використанням популярних бібліотек.

Мета дослідження. Метою даного дослідження є виділення найвідоміших вразливостей безпеки, бібліотек, використання яких може призвести до них та методи запобігання даних вразливостей.

Викладення основного матеріалу. В даний час розробка засобами JavaScript неможлива без використання принаймні однієї з тисяч доступних бібліотек JavaScript із відкритим кодом. Бібліотеки роблять програмування в JavaScript простішим та ефективнішим шляхом прийняття процесів, які вимагають написання кількох рядків коду, зазвичай шляхом створення єдиної функції, яку можна викликати.

Проте ці переваги піддаються ризикам. Бібліотеки JavaScript можуть мати дефекти безпеки, які можуть зробити вразливим веб-сайт, що використовує їх.

Аналіз веб-ресурсів на наявність вразливостей. Згідно з дослідженням, проведеним Північно-Східним університетом в Бостоні [1], більше 37% веб-сайтів використовують щонайменше одну версію бібліотеки з відомими вразливостями.

Загальні вразливі місця безпеки в JavaScript включають міжсторінкові сценарії, перехресні запити через перемички та переповнення буферів. Перехресні сценарії

дозволяють злодіям вставляти шкідливий код у довірені веб-сторінки, які потім передають цей шкідливий код користувачам, які переглядають сайт. Підробка міжсторінкового запиту дозволяє злодіям використовувати браузер користувача для відстеження його активності на інших сайтах. Переповнення буфера відбувається, коли зловмисники надсилають дані, що занадто великі для зберігання в буфері, внаслідок чого дані записуються за межі буфера та дозволяють зловмисникам вставляти шкідливий код, пошкоджені дані або аварійно завершити роботу програми.

Згідно з публікацією блогу npm [2] загальноприйнятим є те, що сучасний проект JavaScript залежить від 700 до 1200 сторонніх пакетів, і задача пошуку інформації про вразливості в інтернеті являється неможливою для розробника. Саме тоді інструменти для аналізу складу програм стають необхідними та надзвичайно корисними.

Інструменти аналізу складу аналізують ваш код та знаходять вразливі компоненти. Це прискорює процес виявлення вразливостей вашого сайту, а також зменшує ризик людської помилки. В більшості випадків такі інструменти активно аналізують проект та сповіщають розробника при знаходженні вразливостей.

У дослідженні Північно-Східного університету було встановлено, що в середньому веб-сайт у своєму наборі даних використовує версії бібліотек, які на 1177 днів старіші, ніж найновіший реліз. Перехід на нову версію вимагає певного часу, тому що потрібно виконати певні зміни програмного коду та їх тестування, щоб перевірити сумісність нових версій пакетів з ресурсом.

Постійне оновлення до більш нових версій бібліотек старих проектів може заважати розробникам приділяти свій час новим проектам, але не можна залишати поточні сайти вразливими.

Інше дослідження, проведене командою Snyk [3], говорить про те, що 77% із досліджених командою 433000 сайтів містять певні вразливості. При цьому 51% з них дозволили собі дві відомі вразливості, а 9.2% - більше чотирьох.

Командою був складений список, які бібліотеки, що використовувались даними ресурсами містять в собі вразливості. Найпопулярнішою в списку являється бібліотека jQuery. 92.5% з числа веб-ресурсів з вбудованою jQuery використовують старі версії бібліотек з великим числом вразливостей. Друге місце займає jQuery UI, що застосовується в 19.9% досліджених ресурсів.

Серед інших вразливих бібліотек знаходяться Modernizr (використовується в 15.1% досліджених ресурсів), Bootstrap (13.7% ресурсів), yepnope (9.9% ресурсів).

Вразливості в перерахованих бібліотеках у більшості випадків пов’язана з міжсторінковим скриптуванням і підстановкою контенту. Вони можуть бути використані для визначення вмісту куків браузера при відкритті користувачем спеціально оформленого посилання.

Також слід відзначити випуск нового масштабного звіту від проекту Open Web Application Security [4]. В ньому дослідники в області інформаційної безпеки спробували класифікувати відповідні ризики, а також запропонували рейтинг найістотніших і популярних проблем, з якими доводиться стикатися веб-спільноті.

Лідируючі позиції в цьому списку впевнено утримують вразливості, пов'язані з появою неперевірених даних в складі виконуваних команд або запитів (SQL, NoSQL, OS, LDAP Injection). За ними слідують вразливості від

некоректної роботи механізмів ідентифікації та автентифікації, що дозволяють отримати доступ без авторизації.

Все більш актуальними стають вразливості, що призводять до витоку користувацьких персональних даних, вони займають третє місце. Також стала популярною нова категорія вразливостей, пов'язаних з некоректною обробкою зовнішніх посилань у XML-документах (атака класу XXE). П'яте місце в новому рейтингу відведено проблемам, що з'являються в оброблювачах списків доступу. Вони дозволяють виконувати операції, не передбачені діючими повноваженнями.

Інструменти аналізу вразливостей. Одним із інструментів пошуку вразливостей являється Snyk. Цей інструмент представляє чудовий інтерфейс для пошуку та перегляду вразливостей системи безпеки. Він також забезпечує інтеграцію з GitHub. Інструменти командного рядка, надані Snyk, не тільки повідомлять про вразливості компонентів, але також запропонують виправити їх за допомогою спеціального інструмента. Переходячи з командного рядка на веб-сайт Snyk можна легко перевірити модулі NPM на вразливості, переглянути інформаційну панель, що відображає вразливості у вашому поточному проекті та налаштувати параметри Github.

Іншим хорошим засобом являється Retire.js, який теж являється чудовим сканером вразливостей для бібліотек JavaScript. Хоча Retire.js не надає функцій веб-додатків або інтеграції Github, таких як Snyk, він використовує інші засоби. Даний інструмент перевіряє на вразливості не лише модулі NPM, які використовуються в Node.js, а й перевіряє бібліотеки JavaScript.

Retire.js запускається насамперед за допомогою інструмента командного рядка, але його можна використовувати різними способами:

- як grunt плагін,
- як gulp завдання,
- як розширення Chrome,
- як розширення Firefox,
- як burp плагін,
- як плагін OWASP Zap.

Висновки. Пошук та виявлення вразливостей в web-сервісах, що використовують JavaScript являється складною задачею із-за того, що при розробці часто бувають задіяні сторонні бібліотеки. Також при розробці серверної сторони проекту з використанням Node.js та NPM модулей задача сильно ускладнюється тим, що NPM модулі часто мають безліч залежностей від інших модулів і дуже важко прослідкувати за актуальністю версій та відсутністю вразливостей у всіх залежностях.

Однак існують засоби, які дозволяють спростити та автоматизувати процес пошуку та виправлення вразливостей в проекті.

Список використаних джерел

1. Thou Shalt Not Depend on Me: Analysing the Use of Outdated JavaScript Libraries on the Web. [Електронний ресурс]. — Режим доступу: http://www.ccs.neu.edu/home/arshad/publications/ndss2017jslibs.pdf

2. The npm Blog. Why use SemVer? [Електронний ресурс]. — Режим доступу: https://blog.npmjs.org/post/162134793605/why-use-semver

3. Snyk Blog: 77% of 433,000 Sites Use Vulnerable JavaScript Libraries. [Електронний ресурс]. — Режим доступу: https://snyk.io/blog/77-percent-of-sites-still-vulnerable/

4. OWASP Top 10 - 2017 The Ten Most Critical Web Application Security Risks. [Електронний ресурс]. — Режим доступу: https://www.owasp.org/images/7/72/OWASP_Top_10-2017_%28en%29.pdf.pdf

ДОВІДКА ПРО АВТОРІВ

Тизунь Віталій Юрійович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Tyzun Vitalii Yuriyovych – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: vitaliy.tyzun@gmail.com

Tyzun Vitalii

VULNERABILITIES OF SAFETY AND CONFIDENTIALITY OF INFORMATION WHEN USING POPULAR JAVASCRIPT LIBRARIES

Relevance of the research. When developing many web-services use third-party libraries, which does not always guarantee that the methods implemented in them will be completely safe for use

Target setting. When working with third-party libraries, the resource developer may not be aware of vulnerabilities that contain certain methods of the library used.

Actual scientific researches and issues analysis. When investigating a large number of web-services, it has been found that most of them contain vulnerabilities associated with the use of popular libraries.

The research objective. The purpose of this study is to identify the most prominent security vulnerabilities, libraries that provide them, and methods for preventing these vulnerabilities.

The statement of basic materials. At the moment, the development of JavaScript tools is not possible without the use of at least one of the thousands of available open source JavaScript libraries.

However, these benefits are risky, because JavaScript libraries may have security vulnerabilities.

Conclusions. Finding and detecting vulnerabilities in web services using JavaScript is a daunting task because third-party libraries are often involved in development. However, there are tools that can simplify and automate the process of finding and correcting vulnerabilities in the project.

Key words: data protection, web-service, JavaScript, security vulnerability.

УДК 004.074.32

Олександр Марковський,
Аббасі Шагін,
Вікторія Максимук

ЕФЕКТИВНА ОРГАНІЗАЦІЯ БАГАТОРІВНЕВОЇ ХЕШ-ПАМ'ЯТІ

У статті пропонується підхід до підвищення ефективності організації багаторівневої хеш-пам'яті для пошуку за ключем. Верхній рівень такої хеш-пам'яті утворює пошуковий кеш-буфер, а нижній – основна пам'ять. Особливістю організації хеш-пошуку, що досліджується в статті, є квазістатичний характер ключів. Це означає, що інтенсивність операцій зміни ключів на порядки менша за інтенсивність операцій пошуку за ключем. Розроблена ймовірнісна математична модель хеш-пошуку у квазипостійних масивах ключів, які зберігаються в дворівневій хеш-пам'яті. На основі цієї моделі запропоновано підхід до оптимізації параметрів хеш-пам'яті. Теоретично та експериментально доведено, що запропонована організація потребує не більше одного звернення до повільної пам'яті нижнього рівня.

Ключові слова: хеш-пам'ять, хеш-пошук, колізії адресації.
Бібл.: 2.

In paper an approach for increasing effectiveness of organization of multi-level hash-memory for rapid retrieval by key. Upper level of such hash-memory is retrieval buffer in cash memory unit and low level is in main memory. The peculiarities of researched hash-memory organization cocsist of near-permanent characters of key array. It means that intensively of retrieval operation by few order is height in compare to operations connected with key change. The probabilistic model of hash-searching in near-permanent keys arrays which are storages in two-level memory has been developed. On the base of this model the effectiveness organization of hash-access to information has been proposed. It has been shown that proposed hash-access organization requires ensure no more then one access to slow low-level memory.

Key words: hash-memory, hash-retrieval, addressing collign.
Bibl.: 2

Актуальність теми дослідження. Пошук за ключем традиційно є однією з базових процедур обробки інформації. Для багатьох практично важливих застосувань питома вага операцій пошуку складає до 30-40% [1]. В останні роки багаторазово виросли обсяги інформаційних масивів, у яких виконується пошук, значно більш жорсткими стали вимоги щодо оперативності пошуку. Вказані чинники диктують необхідність створення нових підходів до організації пошуку інформації в сучасних та перспективних комп'ютерних системах.

Потенційно найбільшу швидкодію пошуку забезпечує хеш-адресація. До останнього часу суттєвою перешкодою на шляху широкого застосування хеш-пошуку було те, що ця технологія вимагає використання надлишкового об'єму пам'яті. Проте вражаючі успіхи інтегральної технології виготовлення мікросхем пам'яті призвели до їх суттєвого здешевлення, що створює технологічні передумови розширення використання хеш-пам'яті.

Таким чином, наукова задача вдосконалення та підвищення ефективності технології хеш-пошуку є актуальною на сучасному етапі розвитку інформаційних технологій.

Постановка проблеми. Сутність хеш-адресації полягає в тому, що ключі і пов'язана з ними інформація записуються за адресою, яка однозначно залежить від ключа. Суттєвим недоліком хеш-адресації є наявність колізій. Вони виникають коли два і більше ключів претендують на одну адресу. Технологічно проблема колізій вирішується за допомогою лінійного пробінгу. Його реалізація полягає в перевірці на вільність адреси, до якої звертається ключ: якщо ця комірка пам'яті зайнята, по порядку перевіряються наступні, і запис виконується в першу вільну комірку. Пошук здійснюється шляхом перевірки заданого ключа зі значеннями, записаними в хеш-памяті, за обчисленою хеш-адресою: коли вони співпадають, пошук завершено. В іншому випадку потрібно послідовно переглянути наступні комірки; знаходження порожньої вказує на те, що такого ключа не існує. Відомим недоліком використання лінійного пробінгу є вторинне групування записів, що призводить до сповільнення пошуку [1].

Аналіз останніх досліджень та публікацій. Одним з найважливіших підходів до формування хеш-перетворень, що не породжують колізій для заданого масиву ключів, є сегментування. Сегментування передбачає розбиття заданого масиву ключів на безліч піднаборів і формування хеш-перетворень, що не породжують колізії для кожного з піднаборів окремо.

Рекурсивний метод формування хеш-перетворення, не породжуючого колізій, запропонований в [2], заснований на використанні сімейств хеш-функцій, що не породжують колізій.

Ефективність рекурсивного формування функцій хеш-перетворення значною мірою залежить від ступеня заповнення хеш-пам'яті: при коефіцієнтах заповнення хеш-пам'яті більших за 0.75 час пошуку хеш-перетворення може стати критичним.

Виділення недосліджених частин загальної проблеми. Основними недоліками існуючих методів побудови хеш-перетворень для квазістатичних масивів ключів є те, що вони потребують дуже багато часу при коефіцієнті заповнення більше 0.75, а також те, що вони не враховують особливостей організації реальної пам'яті комп'ютерних систем, зокрема її багаторівневий характер.

Постановка завдання. Ціллю досліджень є підвищення ефективності хеш-адресації за рахунок оптимізації параметрів хеш-пам'яті при її багаторівневій організації.

Викладення основного матеріалу. Модель дворівневої хеш-пам'яті. При хеш-пошуку в сучасних обчислювальних системах, пам'ять яких має багаторівневу організацію, є ряд особливостей, які можуть принципово змінити підхід до швидкого пошуку даних.

Розглядається дворівнева хеш-пам'ять, що складається з повільнодіючої спільної пам'яті і швидкодіючої кеш-пам'яті, об'єм якої обмежений.

В основу розроблюваної моделі покладена концепція отримання хеш-перетворення $H(X)$, що забезпечує відображення заданої множини Ω із m ключів на s сторінках хеш-пам'яті таким чином, щоб кількість ключів, адресованих у кожну зі сторінок, не перевищувала $(\alpha+\delta)\cdot w$, де α- коефіцієнт завантаження хеш-пам'яті, δ - допустима варіація завантаження сторінки хеш-пам'яті,

$$\alpha+\delta\leq 1.$$

Коефіцієнт α завантаження хеш-пам'яті визначається відношенням кількості m записів, що зберігаються в ній до максимально можливої їх кількості $M = s\cdot w$, виходячи з обсягу хеш-пам'яті:

$$\alpha = \frac{m}{M} = \frac{m}{s \cdot w}. \quad (1)$$

Отримання хеш-перетворення $H(X)$, що задовольняє вказаній вище умові, може бути отримане шляхом підбору. В якості механізму підбору хеш-перетворення доцільним представляється використання стандартизованого шифроблоку типу DES або Rijndael, який здійснює криптографічне однозначне шифрування з використанням ключа K даних D у код C: $C = HK(D)$ [2]. Ключова умова для пошуку інформації X у такому варіанті використовується в якості вхідних даних шифроблоку, ключ K шифроблоку виступає в ролі настроювального коду і є, по суті, номером хеш-перетворення. Вихідний код C шифроблоку поділяється на дві частини: h-розрядний фрагмент використовується як хеш-адреса $A_K(X)$ сторінки, а розряди, що залишилися, представляють собою хеш-згортку $SK(X)$ ключа X пошуку. Відповідно, підбір хеш-перетворення $HK(X)$ виконується шляхом зміни ключа K шифроблоку.

Вважаючи, що на сторінці може розміститися w ключів, хеш-адреса $AK(X)$ сторінки є $h = log_2 s$ - розрядний код. Хеш-функція розподіляє m ключів по s групах, які містять $\eta_1,\eta_2,\ldots,\eta_s$ ключів, причому $\sum_{j=1}^{s} \eta_j = m$. Для того, щоб хеш-перетворення розподіляло ключі сторінками хеш-пам'яті, необхідно, щоб кількість хеш-адрес кожної з сторінок не перевищувала максимально допустимого числа $u=(\alpha+\delta)\cdot w$ записів на сторінці: $\forall j \in \{1,\ldots,s\}$: $\eta_j \le u$. При цьому в кожній зі сторінок хеш-пам'яті забезпечується наявність вільного об'єму пам'яті, достатнього для розміщення $w\cdot(1-\alpha-\delta)$ записів.

Якщо виходити з того, що хеш-перетворення $H_K(X)$ формує рівномірно розподілені хеш-адреси, то в одну сторінку, в середньому, потрапляє m/s хеш-адрес. В якості теоретичної моделі розподілу хеш-адрес коректно використовувати ймовірнісну модель Бернуллі. Відповідно до цієї моделі, попадання хеш-адрес m ключів у межі фіксованої сторінки пам'яті можна розглядати як m дослідів, у кожному з яких з імовірністю $1/s$ відбувається подія: потрапляння хеш-адреси в адресний простір цієї сторінки. Тоді згідно з властивостями даної моделі, можна вважати, що математичне очікування числа хеш-адрес, що потрапляють у фіксовану сторінку, дорівнює m / s з дисперсією $m \cdot \frac{1}{s} \cdot (1 - \frac{1}{s})$. Тоді, відповідно до теореми Муавра-Лапласа, кількість хеш-адрес, що потрапляють у рамки сторінки, підпорядкована розподілу Гауса з математичним очікуванням m/s і дисперсією $m \cdot \frac{1}{s} \cdot (1 - \frac{1}{s})$. Якщо врахувати, що число s сторінок хеш-пам'яті достатньо велике, то дисперсію кількості хеш-адрес, що потрапляють у фіксовану сторінку, можна наближено вважати рівною m / s.

Імовірність P_{os} переповнення сторінки, тобто ймовірність того, що число хеш-адрес, що потрапляють у фіксовану сторінку хеш-пам'яті, перевищить u, для нормального розподілу з урахуванням (1) визначається наступним виразом:

$$P_{os} = 0.5 - \Phi(\frac{u - m/s}{\sqrt{m/s}}) = 0.5 - \Phi(\frac{w\cdot(\alpha+\delta) - m/s}{\sqrt{m/s}}) = 0.5 - \Phi(\delta \cdot \sqrt{\frac{w}{\alpha}}) \quad (2)$$

Для постійного масиву ключів, тобто в ситуації, коли в сторінці немає необхідності мати запас вільної пам'яті для $w\cdot(1-\alpha-\delta)$ записів, $\delta = 1-\alpha$ і ймовірність P_{osw} переповнення сторінки, тобто ймовірність того, що число хеш-адрес, що потрапляють у фіксовану сторінку хеш-пам'яті, перевищить w, визначається наступним виразом:

$$P_{osw} = 0.5 - \Phi(\sqrt{w}\cdot\frac{1-\alpha}{\sqrt{\alpha}}) \quad (3)$$

З виразу (3) випливає, що ймовірність переповнення сторінки визначальною мірою залежить від значення коефіцієнту δ запасу вільної пам'яті в сторінці, від коефіцієнта α заповнення хеш-пам'яті, а також від об'єму сторінки і пам'яті, що потрібна для зберігання одного запису .

Для того, щоб виключити переповнення всіх s сторінок при заповненні фіксованою множиною Ω з m ключів, необхідно підібрати відповідне хеш-перетворення. Ймовірність P_0 того, що за одну пробу хеш-перетворення вдасться виключити переповнення всіх сторінок хеш-пам'яті визначається добутком ймовірностей того, що кожна з s сторінок не буде переповнена:

$$P_0 = (1-P_{OS})^S = \left[0.5 + \Phi(\delta\cdot\sqrt{\frac{w}{\alpha}})\right]^S \quad (4)$$

Середня кількість g проб, які необхідно виконати для підбору хеш-перетворення, при якому виключається переповнення сторінок хеш-пам'яті, визначається наступним виразом:

$$g = \sum_{j=1}^{\infty} j\cdot P_0\cdot(1-P_0)^{j-1} = \frac{1}{P_0} \quad (5)$$

Проведені експериментальні дослідження на основі статистичного моделювання довели адекватність викладеної математичної моделі хеш-адресації квазістатичного масиву ключів в хеш-пам'яті, розділеної на сторінки. Запропоновану модель можна використовувати для оптимізації параметрів хеш-пам'яті.

Висновки. Розроблено модель хеш-пошуку в квазістатичних масивах ідентифікаційних кодів користувачів, що враховує багаторівневу організацію пам'яті. На основі розробленої моделі запропонована організація хеш-пошуку в квазіпостійних масивах ключів. Доведено, що час доступу до інформації по ключу визначається зверненням не більше, ніж до однієї сторінки пам'яті нижнього рівня. Запропонована організація хеш-пам'яті може бути ефективно використана для підвищення швидкості пошуку в сучасних комп'ютерних системах для широкого кола практично важливих застосувань.

Список використаних джерел

1. Марковский А.П. Хеш-память с ограниченным временем поиска по ключу/ А.П. Марковский, Е.В. Порхун, А.В. Мнацаканов // Вісник НТУУ "КПІ". Інформатика, управління та обчислювальна техніка.- 2008,- № 49,- С.156-162.

2. Марковський О.П., Максимук В.Р. Організація мультиадресації в хеш-пам'яті // Збірник тез доповідей 9-ї наукової конференції Прикладна математика та комп'ютинг ПМК-2017. Київ 19-21 квітня 2017. – К.:Просвіта, 2017 – С.195-199.

ДОВІДКА ПРО АВТОРІВ

Марковський Олександр Петрович – кандидат технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

Oleksandr Markovskyi – Associate Professor, *PhD,* Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Аббасі Шагін – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Abbasi Shagin –student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Максимук Вікторія Романівна – студентка, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Viktoriia Maksymuk –student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

MarkovskyiO.P.,
AbbasiShagin,
MaksymukV.R.

EFFECTIVENESS MULTI-LEVEL HASH-MEMORY ORGANIZATION

Acute problem. The Hash function is considered to be the traditional technique for the distribution of keys in an unstructured memory, providing direct access. While the ideal condition would be that one and only one key corresponds to a certain address, in practice, during the use of common algorithms of Hashing, this condition is not fulfilled. There are occasions where different keys correspond to the same address. This situation is called a *collision*. Collisions can be deal with using a variety of methods. Common characteristic for all these methods is the reduction of the speed at which information can be searched in the memory, along with an increase in memory requirements.

Target setting. The ability to search quickly for information is considered a major problem in a wide area of applications, like searching information systems. Hashing as a method of searching for information has existed for many years, but in the above mentioned method the problem of collisions is an inhibitory factor in practical applications.

One of the most effective methods for resolving the problems of collisions, for constant number of keys is the function "Hash transformation, free of collisions", which is called Perfect Hashing Boolean Function.

From the above table we can conclude that in the field of searching, a Hash algorithm based method proves the relation between the Hash Function and the forming of the Hash address bits is a linear function of the key bits.

Actual scientific researches and issues analysis. The negative aspects of the Perfect Hash Function are considered to be the following:

a) The use only of constant (not varied) number of keys.

b) The fact that the process for searching the Perfect Hash function for a given number of keys demands useful time. Recently several of algorithms for searching Perfect Hash Functions have been proposed [9-21]. All these algorithms use, more or less, the force attack searching method to define the subset G (which includes all the Perfect Hash Functions), from the set of all the possible Hash transformations. The effectiveness of these algorithms is determined from the rate of settlement of accommodation in fulfilling the following conditions:

- The algorithm should obtain the Perfect Hash Function (or the Minimum Hash Function) for any number of keys. That means requires force attack searching in the determinate subset G, in which are included all the Perfect Hash Functions.
- The number of functions of the subset G, in which force attack searching is executed, should be the smallest one.
- The time required to calculate the Hash Function should be the least possible.
- We consider in advance that we must have the highest possibility that the Hash Function does not collide for a given set of keys.

Uninvestigated parts of general matters defining. The algorithm that is mentioned in study and which is considered to be one of the most effective, ensures the searching time of the Minimum Perfect Hash Function is proportional to m, where m=number of keys which must be distributed in memory.

The research objective. The fractional use of Hash Function for the forming of bits, in practice, does not cause any changes in the expected results. Having as initial state the aforementioned theoretical research we can form a practical method for the design of both a Minimum and not a Minimum Perfect Hash Function, while being able at any time to create the software in Assembly code.

The statement of basic materials. In paper an approach for increasing effectiveness of organization of multi-level hash-memory for rapid retrieval by key. Upper level of such hash-memory is retrieval buffer in cash memory unit and low level is in main memory. The peculiarities of researched hash-memory organization cocsist of near-permanent characters of key array. It means that intensively of retrieval operation by few order is height in compare to operations connected with key change. The probabilistic model of hash-searching in near-permanent keys arrays which are storages in two-level memory has been developed. On the base of this model the effectiveness organization of hash-access to information has been proposed. It has been shown that proposed hash-access organization requires ensure no more then one access to slow low-level memory.

Conclusions. In paper the new hash-searching organization has been proposed. The essence of this organization consist of using multi hash addressing for several key. It is allowed to remove the grouping keys in hash memory. The procedures of recording keys into the hash-memory and key searching been developed. The experimental evaluation of average of memory access for key recording and key searching has been obtained.

Key words: hash-memory, hash-retrieval, addressing collign.

УДК 004.056

Марковский А.П.,
Искаков Эмиль Русланович.,
Гарасимович Г.В.

КОМБИНАТОРНЫЙ АНАЛИЗБУЛЕВЫХ ФУНКЦИЙ, СПЕЦИАЛЬНЫХ КЛАССОВ ДЛЯ СИСТЕМ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ

В статье предложен новый метод построения булевых балансных функций, которые соответствуют критерию строго лавинного эффекта. Преимуществом предложенного метода является простота реализации и существенно большее по сравнению с известными методами количество функций этого класса, которые могут быть синтезированы. Подробно описана формализованная процедура конструирования балансных функций, обладающих строгим лавинным эффектом. Определена нелинейность синтезируемых функций. Приведен пример синтеза функции.

Ключевые слова: криптографические алгоритмы, булевы функции, балансные функции, функции, обладающие лавинным эффектом.

In this paper new method for designing of Boolean balanced function that satisfies the Strict Avalanche Criterion (SAC) are presented.The advantage of the suggested method is the simplicity of realization and the significant greatest number of the generated functions compared to the known methods. The formalized procedure for construction balanced SAC-functions is described in detail. The nonlinearity of synthesized functions have been determined. Examples of function design are given.

Key-words: Cryptographic algorithms, Boolean functions, balance functions, SAC functions.

Актуальность темы исследования. Развитие распределенных систем в значительной мере зависит от уровня технологии защиты в них информации. В настоящее время в основе большинства систем защиты информации лежат криптографические механизмы, базирующиеся на аналитически неразрешимых математических задачах теории чисел, эллиптических кривых и булевых функций. Использование последних играет особенно важную роль, поскольку вычисление булевых преобразований выполняется на 3-4 порядка быстрее по сравнению со сложными мультипликативными операциями модулярной арифметики, выполняемыми над числами, длина которых на порядок превышает разрядность процессоров.

Исходя из этого развитие теории булевых функций применительно к их криптографическим применениям, а также совершенствование на этой основе технологических аспектов построения на основе таких функций алгоритмов и протоколов защиты информации является важной и актуальной задачей для современного этапа развития информационных и компьютерных технологий.

Постановка проблемы. Булевы функции, используемые в системах защиты информации должны обладать рядом специфических свойств, важнейшим из которых является свойство строгого лавинного эффекта (SAC-StrictAvalancheCriterion), которое характеризуется максимальным значением дифференциальной энтропии, что

обеспечивает устойчивость к нарушению защиты дифференциальным криптоанализом [1,2]. Помимо этого, булевы функции должны обладать высоко нелинейностью с тем, чтобы успешно противостоять линейному криптоанализу [3].

Для практического использования булевых балансных функций, которые удовлетворяют критерию строгого лавинного эффекта, стоит задача разработки формализованных методов их синтеза. Это позволит существенно повысить эффективность использования булевых функций специальных классов для многих применений.

Анализ известных исследований и публикаций. Быстрый прогресс интегральной технологии позволяет создавать эффективные аппаратные реализации таких реконфигурируемых функций с использованием матриц программируемых элементов.

Булева функция $f(x_1,x_2,\ldots,x_n)$, определенная на множестве Z состоящем из 2^n возможных значений наборов X из n переменных называется балансной, если она с равной вероятностью принимает нулевые и единичные значения:

$$\sum_{X \in Z} f(X) = 2^{n-1}$$

Булева функция $f(x_1,x_2,\ldots,x_n)$, удовлетворяет критерию строгого лавинного эффекта, если при изменении значения любой из n переменных значение функции меняется с вероятностью 0.5:

$$\begin{gathered}\sum_{X \in Z} f(X) \oplus f(X \oplus \Delta_j) = 2^{n-1}, \forall j \in \{1,\ldots, n\}, \\ \Delta_j = (d_1,\ldots, d_j,\ldots, d_n), d_j = 1, d_i = 0, \forall i \in \{1,\ldots, n\}, i \neq j,\end{gathered} \quad (1)$$

где Δ_j- n-компонентный двоичный вектор, j-тая компонента которого равна единице, а остальные – нулю. Любая булева функция $f(x_1,x_2,\ldots,x_n)$, может быть представлена в виде канонического разложения Шеннона по j–той переменной x_j:

$$f(x_1,x_2,\ldots,x_n) = x_j \cdot \varphi_j(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n) \oplus \psi_j(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$$

где φ_j и ψ_j – булевы функции, не зависящие от x_j. Поскольку $f(X) \oplus f(X \oplus \Delta_j) = \varphi_j(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$, то булева функция $f(x_1,x_2,\ldots,x_n)$, удовлетворяет критерию строго лавинного эффекта, если при любом $j \in \{1,\ldots,n\}$ функции $\varphi_j(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$ балансны, то есть $\sum_{X \in Z} \varphi_j(X) = 2^{n-1}, \forall j \in \{1,\ldots,n\}$. Поскольку функция $\varphi_j(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$ не зависит от x_j, ее можно рассматривать как функцию от n-1 переменных, для которых существует 2^{n-1} возможных наборов X_j переменных $x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n$, образующих множество Z_j. Тогда условие соответствия функции $f(x_1,x_2,\ldots,x_n)$, критерию строгого лавинного эффекта может быть приведено к виду:

$$\sum_{X_j \in Z_j} \varphi_j(X_j) = 2^{n-2}, \forall j \in \{1,\ldots, n\}. \quad (2)$$

Функция $\varphi_j(x_1,\ldots,x_{j-1},x_{j+1},\ldots,x_n)$ отождествляется рядом исследователей [3] с дифференциалом булевой функции $f(x_1,x_2,\ldots,x_n)$ по переменной x_j:

$$\frac{\partial f(x_1,\ldots, x_j,\ldots, x_n)}{\partial x_j} = \varphi_j(x_1,\ldots, x_{j-1}, x_{j+1},\ldots, x_n). \quad (3)$$

Соответственно, булева функция $f(x_1,x_2,\ldots,x_n)$ удовлетворяет SAC, если ее дифференциалы по всем переменным балансны.

Принимая во внимание практическую важность проблемы автоматизированного синтеза балансных SAC–функций для современных средств защиты информации, за последние 15 лет предложен ряд подходов по решению этой проблемы [1-3].

С начала 90-х годов опубликовано большое число работ, посвященных синтезу булевых функций специальных классов, ориентированных для использования в системах кодирования и защиты информации. Методы, синтеза SAC-функций можно условно разделить на три группы:

- генетические;
- на основе ортогонального базиса;
- комбинаторные.

Генетические методы синтеза булевых функций специальных классов, основаны на том, что каким-то образом получается одна функция класса, а потом из нее путем специальных математических преобразований формируется множество функций, обладающих такими же свойствами. Для практической реализации этого подхода чаще всего используются спектральные преобразования [1,4]. Общим недостатком генетических методов синтеза булевых SAC-функций с является то, что проблема получения первичной булевой функции этого класса остается открытой.

Другим, часто используемым на практике подходом к получению SAC-функций является использование свойств ортогональных систем булевых функций и преобразований над нами. В рамках этого подхода разработано большинство существующих на сегодняшний день методов синтеза балансных SAC–функций [3,4]. Наиболее известный метод получения балансовых SAC-функций на основе ортогональных преобразований предложен в работе [5]. Сущность предложенного ими метода состоит в том, что $_n$ переменных разделяются на два непересекающегося множества, с s и t переменных ($n=s+t$), дальше формируются линейная функция $g(x_1,...,x_s)$ от s переменных и бинарная матрица Q , размерностью $s \times t$, причем число единичных компонентов произведения $\mathrm{Q} \cdot \gamma_1$ матрицы Q на любой $_s$-компонентный вектор γ_1 с одним ненулевым компонентом и произведения $\gamma_2 \cdot Q$ любого t-компонентного вектора γ_2 с одним ненулевым компонентом больше или равно единице. Вектор образован коэффициентами функции $g(x_1,...,x_s)$ может быть линейно-независимой от векторов, образованных столбцами матрицы Q . Балансная SAC-функция формируется согласно формуле:

$$f(x_1,...,x_n)=[x_1,...,x_s]\cdot Q\cdot[x_{s+1},...,x_n]^T \oplus g(x_1,...,x_s)$$

Недостатками этого метода является сложность получения матрицы Q при достаточно больших значениях *n*.

Выделение неисследованных частей общей проблемы.Общим недостатком существующих методов синтеза булевых функций на является принципиально малое число SAC-функций. Основным их недостатком является то, что они позволяют получать относительно небольшое подмножество SAC-функций.

К настоящему времени не существует аналитического выражения для исчисления количества булевых SAC-функций от n переменных.

Постановка задачи.Целью исследований является повышение эффективности синтеза булевых функций, обладающих высокой нелинейностью и свойством лавинного эффекта за счет увеличения количества таких функций.

Комбинаторный анализ булевых функций специальных классов. Для решения задачи синтеза SAC-функций и оценки их количества предлагается использовать комбинаторный подход. Его сущность состоит в установлении комбинаторных зависимостей между значениями SAC-функций на различных наборах и использования этих зависимостей для синтеза функций и определения их количества.

Показано, что для того, чтобы булева функция $f(x_1,x_2,\ldots,x_n)$ была балансной и удовлетворяла SAC по переменной x_i, необходимо, чтобы на половине (2^{n-2}) 2^{n-1} возможных значений остальных n-1 переменных, функция при изменении x_i меняла свое значение на противоположное, на одной четвертой (2^{n-3}) пар наборов принимала строго единичное значение и на оставшихся (2^{n-3}) наборах принимала нулевое значение. Число вариантов размещения значений функций в таблице истинности, удовлетворяющее этим условиям определяет число KSAC-функций:

$$K = C_{2^{n-1}}^{2^{n-2}} \cdot C_{2^{n-2}}^{2^{n-3}} \cdot 2^{n-2} \approx \frac{2^{3\cdot n+2.5}}{\pi}. \quad (4)$$

В частности. при n=4 количество SAC-функций согласно формуле (4) составляет 1680 (или 26%) из общего числа 6435 балансных булевых функций от 4-х переменных. Для n=8 количество SAC-функций согласно формуле (4) составляет уже $3\cdot10^{7}$ (или $5\cdot10^{-65}$ %) из общего числа $5.8\cdot10^{75}$ балансных булевых функций от 8-х переменных. Из приведенных в качестве примера данных, а частности, следует, что поиск имеющих практическое значение SAC-функций от большого числа переменных, представляет собой сложную задачу.

Комбинаторные свойства SAC-функций могут быть использованы и для задач синтеза. Так, доказано, что можно определить базовые фрагменты таблицы истинности и правила складывания их них таблицы истинности SAC-функций. Например, можно определить набор фрагментов E={1000,0100,0010,0001} и набор их инверсий: N={0111,1011,1101,1110} и показать, что в таблице истинности SAC-функции число фрагментов множеств E и N должно быть одинаково, а номера фрагментов в симметричных участках таблицы истинности должны быть различными. Используя эти зависимости можно построить достаточно просто комбинаторно построить таблицы истинности SAC-функций.

Предложенный подход не только может быть использован как для задач практического синтеза криптографически устойчивых булевых преобразований, но и имеет значение для задач теории булевых функций.

Выводы. Предложенный метод является развитием подхода к синтезу балансных булевых функций, удовлетворяющих критерию строго лавинного эффекта на основе исследованных комбинаторных свойств булевых функций, удовлетворяющих критерию строго лавинного эффекта. Главной особенностью разработанного метода является то, что он может использоваться как для синтеза фиксированных балансных SAC-функций, так и для построения управляемого кодом генераторов таких функций. При синтезе фиксированных SAC-функций,

предложенный метод отличается от известных простотой и высокой технологичностью, поскольку не использует представления синтезируемых функций в виде таблицы истинности. Разработанный метод практически реализует построение функций со специфическими свойствами от *n* переменных на основе подмножества функций указанного класса от *n*/2 переменных. Благодаря использованию комбинаторных свойств функций, удовлетворяющих лавинному эффекту обеспечивается возможность построения большего числа SAC-функций по сравнению с известными методами.

Разработанный метод ориентирован на использование в перспективных алгоритмах защиты информации на основе булевых преобразований: алгоритмах симметричного шифрования, потоковых алгоритмах шифрования и хеш-алгоритмах.

Список использованных источников :

1. Forre R. The strict avalanche criterion: spectral properties of Boolean functions and extend definition. / R. Forre // Advanced in Cryptology – Crypto'88 Proceeding, Lecture Notes in Computer Sciences, 403 – 1990-P.450-468.

2. Самофалов К.Г. Комбинаторный подход к получению булевых функций, обладающих строгим лавинным эффектом. К.Г. / Самофалов, А.П. Марковский // Электронное моделирование.- 2004,- Том. 26, - № 3, - с.27-40.

3. Tang D. Highly Nonlinear Booleam Function with optimal algebraic immunity and good behavior against fast algebraic attack / D. Tang, C. Carlrt, X. Tang. // IEEE Transactions on Information theory.-Vol.59.- No. 1.- 2013.- P.653-664.

4. Gao G/ Recent recent result of balanced symmetric Boolean functions / G. Gao, Y. Zhao. // IEEE Transactions on Information theory.-Vol. 62.- No.9.- 2016.- P.5199-5203.

ДОВІДКА ПРО АВТОРІВ

Марковський Олександр Петрович – кандидат технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

Oleksandr Markovskyi – Associate Professor, *PhD,* Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Марковский А. П.,
Искаков Эмиль Русланович.,
Гарасимович Г. В.

КОМБИНАТОРНЫЙ АНАЛИЗ БУЛЕВЫХ ФУНКЦИЙ, СПЕЦИАЛЬНЫХ КЛАССОВ ДЛЯ СИСТЕМ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ

Актуальність теми дослідження. Виходячи з розвитку теорії булевих функцій стосовно їх криптографічним застосуванням, а також вдосконалення на цій основі технологічних аспектів побудови на основі таких функцій алгоритмів і протоколів захисту інформації є важливим і актуальним завданням для сучасного етапу розвитку інформаційних і комп'ютерних технологій

Постановка проблеми.Для практичного використання булевих балансних функцій, які відповідають критеріям суворого лавинного ефекту, стоїть завдання розробки формалізованих методів їх синтезу. Це дозволить істотно підвищити ефективність використання булевих функцій спеціальних класів для багатьох застосувань.

Аналіз останніх досліджень і публікацій. Швидкий прогрес інтегральної технології дозволяє створювати ефективні апаратні реалізації таких реконфігурованих функцій з використанням матриць програмованих елементів. Беручи до уваги практичну важливість проблеми автоматизованого синтезу балансних SAC-функцій для сучасних засобів захисту інформації, за останні 15 років запропонований ряд підходів щодо вирішення цієї проблеми

Виділення недосліджених частин загальної проблеми. Загальним недоліком існуючих методів синтезу булевих функцій на є принципово мале число SAC-функцій. Основним їх недоліком є те, що вони дозволяють отримувати відносно невелика підмножина SAC-функцій. До теперішнього часу не існує аналітичного виразу для обчислення кількості булевих SAC-функцій від n змінних.

Постановка завдання. Метою досліджень є підвищення ефективності синтезу булевих функцій, що володіють високою нелінійністю і властивістю лавинного ефекту за рахунок збільшення кількості таких функцій.

Викладення основного матеріалу. Для вирішення завдання синтезу SAC-функцій і оцінки їх кількості пропонується використовувати комбінаторний підхід. З наведених як приклад даних, а зокрема, випливає, що пошук мають практичне значення SAC-функцій від великого числа змінних, являє собою складну задачу.

Висновки. Запропонований метод є розвитком підходу до синтезу балансних булевих функцій, що задовольняють критерію строго лавинного ефекту на основі досліджених комбінаторних властивостей булевих функцій, що задовольняють критерію строго лавинного ефекту. Головною особливістю розробленого методу є те, що він може використовуватися як для синтезу фіксованих балансних SAC-функцій, так і для побудови керованого кодом генераторів таких функцій.

Ключові слова. Криптографічні алгоритми, булеві функції, балансні функції, функції, що володіють лавинним ефектом.

UDK 004.75

Artem Volokyta,
Artem Kaplunov,
Heorhii Loutskii

METHOD OF EVALUATING THE EFFECTIVENESS OF DISTRIBUTED COMPUTING SYSTEM TOOLS MONITORING

В статті проаналізовано місце засобів моніторингу в роботі розподіленої обчислювальної системи, а також розглянуто метод оцінки ефективності засобів моніторингу на основі ступеня однозначності стану системи, що аналізується.

Ключові слова: моніторинг обчислювальних ресурсів, екстрені ситуації, ефективність системи моніторингу.

Рис.: 1.

In this article the importance of monitoring tools during work of distributed computer systems is analyzed, as well as a method of assessing the effectiveness of means of monitoring distributed computer systems based on the degree of single-valued of the state of the analyzed system.

Key words: monitoring of computing resources, emergency situations, effectiveness of the monitoring system.

Fig .: 1.

The relevance of those studies. In order to effectively use complex distributed systems, their reliability and adaptability must be maximized.

Monitoring systems are the main mechanism of control, thus their constant improvement is one of the priority components in distributed systems development.

Problem formulation. Distributed computer systems (DCS) monitoring provides both general information on the functioning of infrastructure, systems, services and users, as well as information about errors, unusual situations and "bottlenecks".

Monitoring is a process of systematic or continuous data collection and operation about parameters of a complex object or a process. This information can be used in order to optimize the decision-making process, to inform the relevant persons or as a feedback tool.

Analysis of recent scientific researches and issues. The most dynamically developing distributed systems include grid (global integration of computing and information resources in one computer infrastructure) and cloud computing (Internet services that provide user access to computer resources, platforms and software).

The key to efficient and troubleproof systems operation is a high-quality monitoring system, that notifies about failures in a timely manner and gives a detailed overview of individual elements performance, which are objects of monitoring.

Using the method of functional decomposition, the key parts can be identified, that are inherent in any monitoring system [1]:

1) Subsystem of data collection - permanent or regular data collection of a system's components. It is possible to perform a surface data processing to detect any emergencies.

2) Storage subsystem - accumulation and storage of data collected by the data collection system. This includes tools for interacting with databases, stored information archiving, etc.

3) Data analysis subsystem - solving tasks related to background monitoring, i.e. systematic long-term accumulation and analysis of data on the operation of objects (the collection of statistical data, the study of information by the accumulated system, the identification of regularities, etc.)

4) Notification subsystem - notification of responsible persons about emergency situations occurring in the system.

5) Output subsystem - displays information about the operation of the system and the results of checks in a user-friendly format.

6) Subsystem of correction - the system performs certain actions to eliminate any occurred emergency situation. Based on the interaction with the data analysis subsystem, it selects and implements the appropriate actions in accordance with the type of a problem.

These subsystems allow you to assess the performance of controlled objects and immediately respond to a variety of emergency situations.

Specification of uninvestigated parts of general matters. Monitoring system of a distributed computer system is software designed to control the efficient and trouble-free operation of a system, as well as ensuring the safety of equipment and alerting the operator when emergencies arise. Current monitoring systems are able to ensure the safety of equipment, however, reducing the response time of the system to emergencies still remains an issue for large systems [2].

Problem Statement. An important area of monitoring systems appliance is the monitoring of DCS performance. Modern monitoring systems are still not able to detect the effectiveness of DCS usage [3]. However, it is possible to estimate the efficiency of available resources usage by monitoring performance of specific nodes and tasks.

The statement of basic materials. Monitoring is a continuous or systematic process of collecting and processing information about status and parameters of complex process or object [4].

The most important monitoring task is emergency situations detecting. These can be power problems and other event that require immediate response to preserve hardware and to minimize damage. Notifications from the state monitoring sensors, that are processed by the monitoring system at the time of their arrival (passive monitoring), allow the system to rapidly react to emergency situation.

If sensors have no possibility to notify about emergency situation occurrence, then a regular and active monitoring should be organized.

Even large systems have a small number of sensors of infrastructure equipment, which allow to operate with a small data flow.

The next task is to monitor the efficiency of the compute nodes and their components. Timely disconnection of a node with a malfunction increases the chances of recovering its operability in the shortest possible time. Reliability has an assigned key role, because late detection of a malfunction inside a node can lead to loss of its operability for a long time. Modern equipment has built-in protection against certain types of malfunctions. Usually, there are several sensors on each compute node.

Another task is to monitor the suitability of the compute nodes for a task. Considering the fact that in a same situation the problem is reduced to the loss of working time, verification can be carried out not periodically, but before every task launch. There are less of reliability requirements since this kind of situations leads to a failure or an extended task execution time, but does not lead to physical damages of the equipment. At the same time, the data volume spreads to several tens of sensors per compute node.

The reliability requirements are set to produce performance monitoring. Performance monitoring is a tracking of various OS and hardware parameters in order to find what prevents executable program from achieving a better performance. The failure of this subsystem does not lead to any critical consequences, as the tasks keep executing, and the equipment does not suffer. However, there are inconveniences when debugging programs occur.

The generalized view of a monitoring system architecture is shown in Fig. 1:

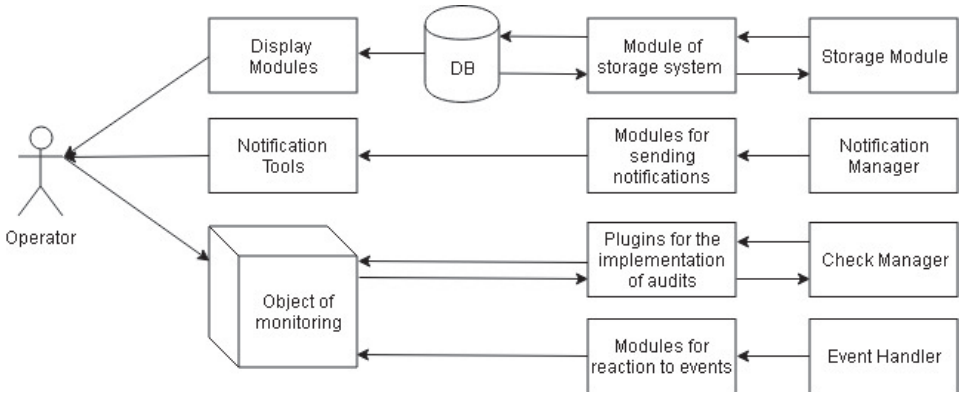


Fig. 1 Architecture of the monitoring system

The principle of interaction of the monitoring system with the analyzed system can be described in the form

$$S = \langle O, P, L, M, {}_p, {}_u \rangle$$

where O is the set of objects of the analyzed system; P - set of properties of objects of the analyzed system; L - a set of interconnected objects, reflecting the dependencies of a stable state of themselves or the properties of each other; M - a set of sensors or methods for analyzing a monitoring system; ${}_p$ is a function that represents the presence of a certain object property:

$${}_p = O \otimes P \to \{0,1\};$$

${}_u$ – function that represents the option of using a specific sensor or a method of analysis for monitoring the current property state of the object:

$${}_u = M \otimes P \to \{0,1\}.$$

The degree of single-valued state of an analyzed system can be represented as

$$E = \frac{\sum_{i=1}^{c} ep_i w_{p_i} n_{p_i}}{\sum_{j=1}^{d} w_{p_j} n_{p_j}}, \quad (1)$$

where $p_i$are the properties of analyzed objects by a monitoring system (for which there exists ${}_u(p_i, s) = 1$); p_j- all properties of an analyzed system's objects (for which there exists ${}_p(o, p_i) = 1$); w - the contribution of a property to an analyzed system functioning; n_p- the number of objects that have property $p$in an analyzed system.

ep_i– effectiveness evaluation of the correction subsystem element that performs control actions over a propertyp_i.

Therefore, L can not be greater than 1 and for each property in a monitoring system there is a method applicable for monitoring the state of this property.

Conclusions. By using the described method, it is possible to evaluate the effectiveness of a monitoring system, as well as to identify the system's analyzed parts that are not well covered by sensors or correction means of the monitoring system.

References

1. Кореньков В. В. АРХИТЕКТУРА И ПУТИ РЕАЛИЗАЦИИ СИСТЕМЫ ЛОКАЛЬНОГО МОНИТОРИНГА РЕСУРСНОГО ЦЕНТРА / В. В. Кореньков, П. В. Дмитриенко. // Системный анализ в науке и образовании. – 2011. – №3.

2. Стефанов К. С. СИСТЕМА МОНИТОРИНГА ПРОИЗВОДИТЕЛЬНОСТИ СУПЕРКОМПЬЮТЕРОВ / К. С. Стефанов. // Вестник ПНИПУ. Аэрокосмическая техника. – 2014. – №39. – С. 17–34.

3. МЕТОДЫ И СРЕДСТВА МЕТАМОНИТОРИНГА РАСПРЕДЕЛЕННЫХ ВЫЧИСЛИТЕЛЬНЫХ СРЕД / И. А.Сидоров, А. П. Новопашин, Г. А. Опарин, В. В. Скоров. // Вестник ЮУрГУ. Серия «Вычислительная математика и информатика» – 2014. – №2. – С. 30–42.

4. Кореньков В. В. Методология развития научного информационно – вычислительного комплекса в составе глобальной грид-инфраструктуры : дис. докт. техн. наук : 05.13.01 / Кореньков Владимир Васильевич – Дубна, 2012. – 260 с.

ДОВІДКА ПРО АВТОРІВ

Волокита Артем Миколайович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Volokyta Artem – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: artem.volokita@kpi.ua

КаплуновАртемВолодимирович – аспірант, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kaplunov Artem – PhD student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: art.kaplunov@gmail.commailto:artem.volokita@kpi.ua

Луцький Георгій Михайлович – професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Loutskii Heorhii – professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

Волокита А. М.,
Каплунов А. В.,
Луцький Г. М.

МЕТОД ОЦІНКИ ЕФЕКТИВНОСТІ ЗАСОБІВ МОНІТОРИНГУ РОЗПОДІЛЕНОЇ ОБЧИСЛЮВАЛЬНОЇ СИСТЕМИ

Актуальність теми дослідження. Проблема ефективності засобів системи моніторингу розподіленої обчислювальної системи стає більш актуальною в останні дні у зв'язку зі зростаючим попитом на використання великих grid- та cloud-систем. Дана робота присвячена оцінці ефективності використання засобів системи моніторингу розподіленої обчислювальної систем.

Постановка проблеми. Актуальні на даний момент системи моніторингу здатні забезпечити збереження обладнання, однак, зменшення часу відгуку системи на виникнення екстрених ситуацій залишається актуальним завданням для великих систем.

Аналіз останніх досліджень і публікацій. Протягом останніх років з’являється все більше статей присвячених тематиці моніторингу розподілених обчислювальних систем. Проте підходи до системи моніторингу все ще не здатні відстежувати ефективність використання розподілених обчислювальних систем в цілому.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена вивченню та аналізу розглянутого підходу для оцінки ефективності засобів моніторингу розподілених обчислювальних систем. Дослідження сфокусовано на можливості використання степені однозначності стану системи, що аналізується, для визначення ефективності моніторингу в цілому.

Постановка завдання. Завданням є визначення методу, за допомогою якого можливо оцінити ефективність засобів системи моніторингу розподіленої системи що розробляються, чи використовуються на даний момент.

Викладення основного матеріалу. Проведено аналіз завдань моніторингу розподілених обчислювальних систем. Описано підхід для оцінки ефективності моніторингу, на основі степені однозначності стану системи, що аналізується.

Висновки. Використовуючи описаний метод, можливо оцінити ефективність функціонування системи моніторингу, а також виявити аналізовані ділянки системи, недостатньо добре охоплені датчиками або засобами корекції системи моніторингу.

УДК 681.518.54

Оксана Кучмий, Виктор Корнейчук, Анна Дорошенко

НАДЕЖНОСТЬ ЭЛЕКТРОСНАБЖЕНИЯ КОМПЬЮТЕРНЫХ СИСТЕМ И МЕТОДЫ ЕЕ ПОВЫШЕНИЯ

RELIABILITY OF COMPUTER SYSTEMS ELECTRICAL SUPPLY AND METHODS OF ITS INCREASE

В статье приведены экспериментальные данные по интенсивности отказов сети электропитания ~220V. Предложена методика расчета надежности сети при использовании блоков бесперебойного питания, позволяющая выбрать указанные блоки с параметрами, обеспечивающими заданный уровень надежности.

Ключевые слова: надежность, интенсивность отказов, перерывы электроснабжения, время восстановления.

Рис.:1. Табл.: 1. Библ.: 6.

Experimental data regarding the fail intensity and recovery of a ~220W network is given. A methodology for evaluation the reliability of a network while using uninterruptible power supplies that allows for choosing the specified blocks with parameters that provide the required reliability level of a computer network is suggested.

Keywords:reliability, failure intensity, power outage, recovery time.

Fig.: 1. Tabl.: 1. Bibl.: 6.

Актуальность темы исследования. При построении высоконадёжных компьютерных систем (КС) возникает задача расчёта их надёжности, которая во многом зависит от надёжности сети электроснабжения (СЭ). Для указанных расчётов необходимо иметь достоверные данные о надёжности СЭ.

Постановка проблемы. Отсутствие достоверных данных о надёжности работы СЭ делает невозможным расчёт надёжности КС и их оптимизацию по критерию стоимость – надёжность.

Анализ последних исследований и публикаций. Приведённые в литературе данные о надёжности СЭ отличаются друг от друга на несколько порядков и из-за этого не могут быть использованы для расчёта надёжности КС и их оптимизации по критерию надёжности.

Выделение неисследованных частей общей проблемы. В данной статье приведены данные о интенсивности отказов (отключений) в подаче электроэнергии и интенсивности восстановления подачи электроэнергии в СЭ.

Постановка задания. Заданием есть получения достоверных числовых данных о надёжности СЭ, пригодных для расчётов надёжности КС, и оптимизации на их основе мероприятий по обеспечению надёжности по критерию стоимость – надёжность.

Изложение основного материала. Приведены данные о 140 отключениях в подаче электроэнергии, полученные путём наблюдения за работой СЭ в течении почти 10 лет. Это позволило определить интенсивность отказов СЭ $\lambda = 0{,}00169$ час$^{-1}$, наработку на отказ $T_0 = 593$ часа, время восстановления $T_в = 0{,}585$ час и интенсивность возобновлений $\mu = 1{,}71$ час$^{-1}$. Показано, как эти данные

могут быть использованы в расчёте надёжности КС и выборе блоков бесперебойного питания.

Ключевые слова: надежность, интенсивность отказов, перерывы электроснабжения, время восстановления.

Обеспечение надежности, согласно ГОСТ 27.205-97, является одной из основных по объему и сложности задач разработчика компьютерных систем (КС). Так как сколь угодно надежная аппаратура не способна выполнять свои функции без электроснабжения, то первым по важности мероприятием по повышению надежности КС является обеспечение бесперебойного электропитания КС. Оптимизация технико-экономических показателей КС при проектировании невозможна без учета надежности сети электропитания (СЭ). Для этого необходимы данные о надежности. Приведенные в литературе данные вызывают сомнения, так как отличаются друг от друга на 3 порядка. Так, например, в [1] указано, что интенсивность отказов λ и восстановлений μ СЭ равны $\lambda=270\cdot10^{-6}$ и $\mu=3$ час$^{-1}$, а в [2] приведены данные, из которых следует, что $0{,}33\geq\lambda\geq3\cdot10^{-3}$ час$^{-1}$. Для получения более достоверных данных с 0 часов 24.09.2008 года и по 3 часа 13.03.2018 г. (82971часов) нами проводились наблюдения за сбоями СЭ в городских условиях, которые продолжаются и по сей день. За указанное время зафиксировано 140 перерывов электроснабжения (Табл.1), где время суток округлено до целых часов. Отсюда [3] $\lambda=140/82971\approx0{,}001687$ час$^{-1}$., $T_0=\lambda^{-1}=593$ часа, интенсивность восстановлений СЭ $\mu=60/2107{,}63=0{,}0285$ мин$^{-1}=1{,}71$ час$^{-1}$; $T_в=\mu^{-1}=0{,}585$ час=35,1 мин.

Таблица 1

Перерывы электроснабжения в 2008-2018 годах

где N – номер перерыва (1,2,3…), T – суммарное время в часах от начала наблюдения (0 часов 24.09.2008 г.) до восстановления электроснабжения после N-ного перерыва, t – время в минутах отсутствия электроснабжения после N-ного перерыва (величина t указана только в тех случаях, когда ее удалось зафиксировать), *N=70,71,72 и 73 – четыре перерыва с интервалами 5 сек, 5 сек и 182 сек, при этом t=2 сек = 0,03 мин.

N:T-t	N:T-t	N:T-t	N:T-t	N:T-t
1:65	27:22680-1	53:45700-45	82:57925	112:70615
2:258	28:22681-80	54:46511-60	83:58248-0,02	113:72440
3:497	29:23575	55:48975-120	84:58376	114:73135
4:1633	30:23674	56:49201-44	85:58423	115:74606
5:2065	31:24117	57:49558-0,02	86:59834	116:75012-0,02
6:2296-13	32:24421	58:49597	87:61536-0,02	117:75183-0,02
7:3080	33:25728-0.02	59:49774	88:61814-19	118:75310-0,02
8:4487-15	34:31504	60:50527-140	89:62144	119:75543-480
9:4525	35:32105	61:50599	90:62170	120:76273-5
10:5027-20	36:32894	62:50611	91:62660	121:76489
11:5208-2	37:33121-0.17	63:50650	92:63348-70	122:76501
12:5573	38:33804-0.02	64:50848	93:63351	123:76770-120
13:7639	39:34887-3	65:52303	94:63397-27	124:77556
14:8421	40:34981-10	66:52313	95:63492-156	125:78541
15:8886	41:35321	67:52824-0,02	96:63946-25	126:79178
16:9194-15	42:36352	68:53046	97:64007-80	127:79530-1
17:11714	43:37645	69:54946-140	98:64918	128:79530-1
18:12803	44:38155	70:55200-0,03*	99:66554-0,1	129:79530-1

Оконьчание табл.

N:T-t	N:T-t	N:T-t	N:T-t	N:T-t
19:14300-0.1	45:38459	71:55200-0,03*	100:67268	130:80049-51
20:15240-25	46:38653	72:55200-0,03*	101:67776-0,02	131:80630-14
21:15281	47:40673	73:55200-0,03*	102:67824-0,02	132:81036
22:16296-0.02	48:40705-1	74:55331-200	103:68849	133:81082
23:17707	49:42995-1	75:55550	104:68968	134:81208
24:18593-0.02	50:44185	76:55883	105:69725	135:81786
25:21497	51:44376-20	77:56221	106:69761	136:82068
26:21592	52:44798	78:56842-10	107:69840-1	137:82074-7

Для устранения влияния перерывов в электроснабжении в КС используют блоки бесперебойного питания (ББП) со встроенным аккумулятором, одним из основных параметров которых является время T_a автономной (т.е. без электроснабжения) работы ББП [5]. Оценить влияние T_a на надежность ББП можно путем рассмотрения графа его работы (Рис.1), где цифрами 0, 1 и 2 обозначены состояния: 0 – КС питается от сети ~220V,1 – сеть ~220V отказала, КС питается от ББП,2 – сеть ~220V не восстановилась, а ББП исчерпал свой ресурс,λ и μ – интенсивности отказов и восстановления сети ~220V.

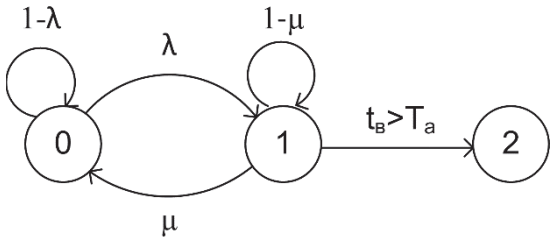


Рис. 1. Граф работы КС с БПП

Среднее время пребывания КС в состоянии 0 (до перехода в состояние 1) будет равно λ^{-1}. Следовательно, за время tКСt·λ раз будет переходить в состояние 1. Если время $t_в$ восстановления ~220V будет меньше чем время T_a, то КС перейдет в состояние 0 и отказа не будет. Вероятность этого события будет равна $1\text{-}e^{-\mu T_a}$.Для того, чтобы отказов КС за время t не было, необходимо, чтобы КС все t·λ переходов совершила в состояние 0. Отсюда следует, что вероятность безотказной работы КС за время t будет равна

$$P(t) = (1 - e^{-\mu T_a})^{t\lambda} = e^{(\ln(1-e^{-\mu T_a}))^{t\lambda}} \quad (1)$$

Так как ББП имеет, как и любые другие блоки КС, свою интенсивность λ_u отказов (величина порядка 10^{-5}час$^{-1}$), то результирующая интенсивность отказов ББП будет равна

$$\lambda_u^{'} = \lambda_u - \lambda \ln (1 - e^{-\mu T_a}) \quad (2)$$

Для увеличения T_a применяют последовательное включение ББП, при котором первый ББП питает второй ББП в течение времени T_a, после чего первый ББП отключается, а второй начинает автономную работу. В некоторых ББП для увеличения T_a в n раз предусмотрена возможность смены аккумуляторов в «горячем» (т.е. рабочем) режиме, естественно, что для этого необходимо иметь n-1 запасных заряженных аккумуляторов [4,5].

В этом случае $\lambda_u^{'} = \lambda_u - \lambda \ln (1 - e^{-\mu n T_a})$.

Как и многие другие мероприятия по повышению надежности энергетике [3,4] величина n имеет оптимальное значение, при которой суммарные затраты на обеспечение надежности и потери от отказов будут минимальными.

Если все же перечисленные мероприятия не позволяют добиться заданного уровня надежности сети питания, то применяют ББП, включающий дизельный двигатель и генератор электроэнергии, позволяющий неограниченное время работать без внешней сети питания ~220V[5,6].

ЛИТЕРАТУРА

1. Китаев В. Е. Расчет источников электропитания устройств связи.- М.: Радио и связь, 1993.- 232 с.

2. Смирнов Ю.К. Секреты эксплуатации жестких дисков ПК.- СПб.: БХВ-Петербург, 2006.- 384 с.

3. Тарасенко В. П., Маламан А.Ю., Черниченко Ю.П., Корнійчук В.І. Надійність комп'ютерних систем.- К.: «Корнійчук», 2007.- 256 с.

4. Трубицын В. И. Надежность электростанций: Уч. для вузов.- М.: Энергоатомиздат, 1997.- 240 с.

5. Методические вопросы исследования надежности больших систем энергетики: сб. науч. тр./Отв. ред.: Н.И.Воропай.- Иркутск: ИСЭМ СО РАН. Вып. 57.- К.: Знания Украины, 2007.- 396 с.

6. Воробьев А. Ю. Электроснабжение компьютерных и телекоммуникационных систем.- М.: Эко-Тренд, 2002.-280с.

ДОВІДКА ПРО АВТОРІВ

Кучмій Оксана Олександрівна – асистент, кафедра системного програмування и спеціалізованих комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kuchmiy Oksana – assistant, Department of System Programming and Specialized Computer Systems, National Technical Universityof Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: ok_kuch@i.ua.

Корнійчук Віктор Іванович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Korneychuk Victor –associate professor, Departmentof Computer Engineering, National Technical Universityof Ukraine “IgorSikorskyKyivPolytechnicInstitute”.

E-mail: vikorn1940@gmail.com.

Дорошенко Анна Юріївна – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Doroshenko Anna – student,DepartmentofComputerEngineering,NationalTechnical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: annadoroshenko03@gmail.com

Kuchmiy Oksana,
Korneychuk Victor,
Doroshenko Anna

RELIABILITY OF COMPUTER SYSTEMS ELECTRICAL SUPPLY AND METHODS OF ITS INCREASE

Relevance of the research topic. When constructing highly reliable computer systems (CS) there emerges the problem of evaluation of their reliability, which mostly depends on the reliability of the power source (PS), namely the network. Trustworthy specification of a PS are required for these evaluations.

Problem introduction. The absence of trustworthy specifications of the reliability of a PS makes it impossible to evaluate the reliability of CS and their optimizations with respect to the cost-reliability criteria.

The analysis of recent research and publications. The data about the reliability of PS provided in articles and publications differs significantly by the order of multiple magnitudes and therefor cannot be considered as trustworthy.

Determining the innovative part of the article. In this article the data concerning the fail intensity and the intensity of recovery in supplying the electricity with PS is given.

Goal introduction. The purpose of this article is to derive trustworthy specifications of PS and their optimization with respect to the cost-reliability criteria.

The essence. The data related to 140 network fails that was gathered during the analysis of a PS operation for 10 years is presented. This allowed to derive the intensity of failure as $\lambda = 0{,}00169$ hour^{-1}, time until a fail occurs as $T_0 = 593$ hours, the recovery time as $T_r = 0{,}585$ hours and the recovery intensity as $\mu = 1{,}71$ hour^{-1}. It is shows how this data can be used to evaluate the reliability of a CS and when choosing uninterruptible power sources.

Keywords:reliability, failure intensity, power outage, recovery time.

Секція 2.RT (Інтернет речей, системи реального часу). Section 2. RT (Internet of Things, Real-Time Systems).

УДК 004.383

Анатолій Сергієнко,
Кадір Сафван Хусейн,
Анастасія Сергієнко

ФІЛЬТРИ ЗІ СКІНЧЕННОЮ ХАРАКТЕРИСТИКОЮ З МІНІМІЗОВАНИМИ АПАРАТНИМИ ВИТРАТАМИ

MINIMIZEDHARDWAREFIRFILTERDESIGN

Розглядається розробка паралельних нерекурсивних фільтрів, які реалізуються в програмованих логічних інтегральних схемах. Новий метод полягає в тому, що блоки множення на коефіцієнти, які мають невелику амплітуду, заміняються на блоки постійної пам'яті, які зберігають кратні значення цих коефіцієнтів. За рахунок цього зменшуються апаратні витрати на реалізацію фільтра та збільшується його пропускна спроможність.

Ключові слова: ПЛІС, нерекурсивний фільтр, конвеєр.

Рис.: 3. Табл.:1. Бібл.: 4.

The development of the parallel finite impulse response filters for the FPGA implementation is considered. A new method consists in substituting the multipliers to the small coefficients to the constant coefficient multipliers, which store the multiplied values of these coefficients. Due to this, the filter hardware volume is minimized and its throughput is increased.

Key words: FPGA, FIR filter, pipeline.

Fig.: 3. Tabl.:1. Bibl.: 4.

Вступ. Програмовані логічні інтегральні схеми (ПЛІС) широко використовуються для високошвидкісної обробки цифрових сигналів. Архітектура ПЛІС адаптована до реалізації фільтрів зі скінченною імпульсною характеристикою (СІХ). Для цього, наприклад, ПЛІС фірми Xilinx містять блоки DSP48, кожен з яких призначений для розрахунку одного ступеня СІХ-фільтра в конвеєрному режимі. Але порядок такого фільтра обмежений об'ємом ПЛІС та кількістю блоків DSP48 в одному стовпці елементів мікросхеми. Як результат, порядок СІХ-фільтра, який генерується утилітою Xilinx Coregen для ПЛІС Spartan-6, обмежена діапазоном від 8 до 48 [1]. Крім того, якщо ПЛІС використовується лише для фільтрації, то в ній неефективного використовується решта програмованих ресурсів, таких як конфігуровані логічні блоки та їх логічні таблиці (ЛТ).

Існує багато методів реалізації СІХ-фільтрів, які не мають блоків множення. Більшість із них використовують схеми множення на константу (СМК), які мають мінімізований об'єм апаратного забезпечення. Вони широко використовуються у ПЛІС протягом десятиліть. При цьому такі фільтри зовсім не застосовують апаратні блоки множення, які входят у склад блоків DSP48 [2,3,4].

У роботі запропоновано новий підхід до розробки СІХ-фільтру, який використовує як апаратні блоки множення, так і СМК. Він забезпечує як збільшення порядку фільтра, так і мінімізовані апаратні витрати, а також підтримує високу пропускну здатність фільтра.

Структура СІХ-фільтра. СІХ-фільтр, який призначений для реалізації у ПЛІС фірми Xilinx FPGA, має відому систолічну структуру [1]. Граф синхронних потоків даних (ГСПД) систолічної структури фільтра *k*-гопорядку показана на рис.1, а його умовне зображення — на рис.1, б. Тут x_i, y_i є вхідними та вихідними даними, кола, трикутники та товсті відрізки представляють собою додавання, множення на коефіцієнт і затримку на один такт, відповідно. Цей граф відображається у відповідну паралельну конвеєрну структуру з максимальною тактовою частотою за допомогою відображення один до одного.

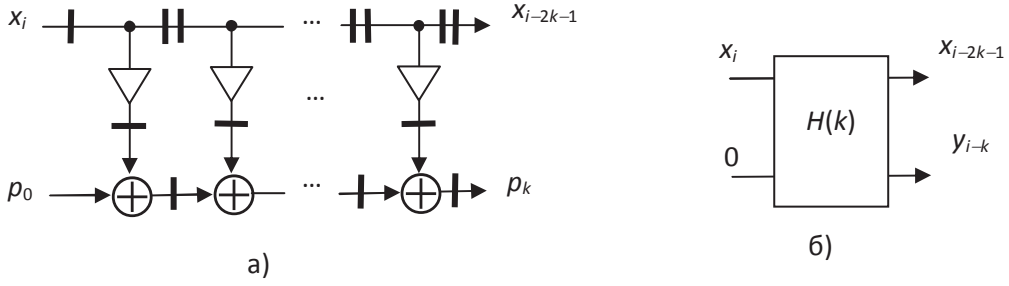


Рис. 1. ГСПД систолічного СІХ-фільтра (а) та його умовне зображення (б)

Набір коефіцієнтів СІХ-фільтра *n*-гопорядку формує його імпульсну характеристику $H(n)$. Пропонується реалізувати СІХ-фільтр з трьох блоків, які обчислюють згортку з лівою $H_1(n_l)$, середнєю $H_2(n_m)$ та правою $H_3(n_r)$ підмножинами коефіцієнтів, $n = n_l + n_m + n_r$. Відповідний модернізований ГСПД показано на рис.2. У більшості випадків розрядність коефіцієнтів $H_1(n_l)$ та $H_3(n_r)$ набагато менше, ніж коефіцієнтів $H_2(n_m)$. Тому бажано реалізувати частину фільтра, позначену на рис.2 як $H_2(n_m)$, використовуючи блоки DSP48, а решту фільтра —з використанням СМК.

Третій блок фільтра, який стоїть останнім у графі на рис.2, повинен мати підвищену розрядність проміжних результатів, щоб зберегти низький рівень помилок від їх усікання. Це може суттєво знизити ефективність фільтра, оскільки багаторозрядний суматор на основі ЛТ є значно повільнішим, ніж відповідний суматор у блоці DSP48.

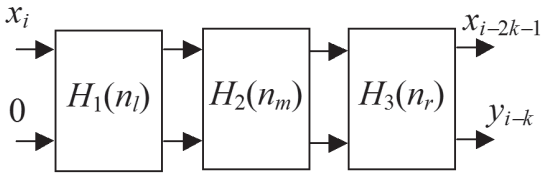


Рис. 2. Модернізований ГСПД фільтра

Щоб звести до мінімуму цей фактор, запропоновано удосконалений ГСПД, який показано на рис.3. Відповідний удосконалений фільтр містить дві частини $H_1(n_l)$ та $H_3(n_r)$ з однаковою розрядністю, яка набагато менша, ніж розрядність операндів блоку DSP48. Результат фільтрації формується в окремому суматорі

зі збільшеною розрядністю. Як результат, розроблений СІХ-фільтр містить лише n_m блоків DSP48, яких може бути набагато менше, ніж порядок фільтра n.

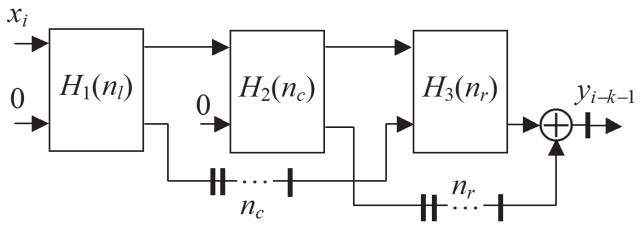


Рис. 3. Удосконалений ГСПД фільтра

Експериментальні результати. Для оцінки ефективності НІХ-фільтрів з запропонованою структурою, було випробувано три реалізації фільтрів нижніх частот порядку n = 35, коефіцієнти, вхідні та вихідні дані яких мають розрядність 16. При цьому СМК були побудовані з використанням представлення коефіцієнтів у канонічній двійковій системі числення [4], причому дерево суматорів СМК має не більше п'яти вхідних складових. За цими умовами, перший та третій блоки фільтрів (див. рис. 2, 3) мають як мінімізовані апаратні витрати, так і високу пропускну здатність.

ГСПД, представлені на рис. 1-3 були описані мовою VHDL. Потім проекти фільтрів були синтезовані для ПЛІС Xilinx Spartan-6. Результати синтезу для деяких фільтрів низьких частот показані в таблиці 1. Щоб визначити інтегральну характеристику апаратних витрат Q, вважалося, що для реалізації 18-розрядного блоку множення, що входить у склад блоку DSP48, потрібно 200 ЛТ.

Таблиця 1

Прараметри СІХ-фільтрів

Структура фільтра	Апаратні витрати				Максимальна частота f_C, МГц	f_C/Q
	ЛТ	Тригерів	DSP48	Q		
Систолічна, рис.1	0	0	33	33	390	11.8
Модернізована, рис.2	772	1538	17	20,9	146	7.0
Удосконалена, рис.3	914	1639	17	21,6	267	12.4

Аналіз таблиці показує, що структура фільтра на базі лише блоків DSP48 забезпечує максимальну тактову частоту дискретизації f_C за рахунок збільшених еквівалентних апаратних витрат Q. Модернізована структура має меншу величину f_C, але забезпечує мінімальний обсяг апаратного забезпечення. І удосконалена структура має менші апаратні витрати ніж систолічна, а її відношення пропускної здатності до апаратурних витрат f_C/Q досягає максимуму.

Висновки. Запропоновано модернізацію та удосконалення систолічної структури цифрового фільтра зі скінченною імпульсною характеристикою, яка реалізована у ПЛІС. Модернізація полягає в тому, що в структурі використовуються як універсальні апаратні блоки множення, так і блоки монження на коефіцієнт, які реалізовані на основі суматорів. Нова структура фільрів забезпечує як оптимальне відношення пропускної здатності до апаратних витрат,

так і збільшення максимального порядку реалізованих фільтрів. Одержані фільтри описані мовою VHDL і мають приблизно удвічі менше число блоків множення, ніж фільтри, які згенеровані програмою Xilinx Coregen.

Список використаних джерел

1. Spartan-6 FPGA DSP48A1 Slice User Guide. UG389 (v1.2) / Xilinx Inc. / May 29, 2014. —46 p.

2. Meyer-Baese. U. Digital Signal Processing with Field Programmable Gate Arrays. Springer, 4-th Ed. 2014. — 930 p.

3. SergyienkoA. FIRfiltersoftcoregenerator / A. Sergyienko,V. Vasylienko, O. Maslennikow // PraceIVKonferencjiKrajowej „Reprogramowalneukladycyfrowe", RUC'2001. –Szczecin, Poland. 2001. — P. 167-172.

4. Kumm M. Multiple Constant Multiplication Optimizations for Field Programmable Gate Arrays. Springer, 2016. — 206 p.

ДОВІДКА ПРО АВТОРІВ

Сергієнко Анатолій Михайлович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Sergiyenko Anatoliy – associate professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: aser@comsys.kpi.ua

Кадір Сафван Хусейн – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Mustafa Rekar Quasim – student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: safwan.h.qadir@gmail.com

Анастасія Анатоліївна Сергієнко– студентка, кафедра системного програмування таспеціалізованих комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Anastasia Anatoliivna Serhienko – student, Department of System Programming and Application Specific Computer Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: ananserr@gmail.com

Anatoliy Sergiyenko,
Quadir Safwan Hussein,
Anastasia Serhienko

FINITE IMPULSE RESPONSE FILTERS WITH A MINIMIZED HARDWARE

Relevance of research topic. Field programmable gate arrays (FPGAs) are widely used for the high-speed digital signal processing, in particular for the signal processing by the finite impulse response(FIR) filters. In order to increase the efficiency of the use of the FPGA configurable resources and reduce their energy consumption, it is necessary to introduce more perfect structures of the FIR filters.

Problem consists in, one hand, large configurable resources of modern FPGAs, other hand, ineffective resource utilization in the filters which are configured in FPGAs using the usual methods.

Analysis of recent research and publications. Implementation of the FIR filters in FPGA is usually based on the use of the multiplication blocks. At the same time, the rest of the programmable resources is used irrationally.

Selection of unexplored parts of the general problem. This article is devoted to the study and implementation of the IIR filters in FPGAs. Such filters use both hardware multiplication blocks, and application specific multiplication blocks based on the configurable FPGA resources.

Setting objectives. The goal is to improve the FIR filter structure and its multiplication blocks in order to increase the effectiveness of the FPGA resource utilization.

Presentation of the main material. The modernization and improvement of the systolic structure of the digital FIR filter, which is implemented in the FPGA, is proposed. The modernization consists in the fact that universal hardware multiplication blocks, and constant coefficient multipliers are coexist in the structure. And the last multipliers are implemented on the basis of adders. The new filter structure provides both the optimal ratio of bandwidth to hardware costs and an increase in the maximum order of implemented filters. The resulting filters are described in VHDL and have approximately double the number of multiplication blocks than the filters generated by the Xilinx Coregen tool.

Conclusions. The proposed structure of the FIR filter provides the optimal ratio of bandwidth to hardware costs, as well as increasing the maximum order of the filters implemented in FPGA. The resulting filters are described in VHDL and have approximately double the number of multiplication blocks than the filters generated by the Xilinx Coregen program.

Key words: FPGA, FIR filter, pipeline.

УДК 004.383

Анатолій Сергієнко,
Хасан Мухамед Джамал,
Павло Сергієнко

АЛГОРИТМ І СТРУКТУРА МОДУЛЯ ДЛЯ ОБЧИСЛЕННЯ КВАДРАТНОГО КОРЕНЯ У ПЛІС

ALGORITHM AND STRUCTURE OF THE SQUARE ROOT CALCULATOR IMPLEMENTED IN FPGA

Розглядається розробка апаратних пристроїв для обчислення функції квадратного кореня за ітераційним алгоритмом. Запропонований алгоритм дає змогу прискорити обчислення функції квадратного кореня та зменшити апаратні витрати за рахунок обчислення кількох ітерацій табличним методом. Запропонований алгоритм розрахований на реалізацію у програмованих логічних інтегральних схемах.

Ключові слова: ПЛІС, квадратний корінь, CORDIC, конвеєр.

Рис.: 3. Бібл.: 6.

The development of the hardware units for the square root (SQRT) function calculations is considered, which is based on the CORDIC-like iterative algorithm. The proposed algorithm helps both to speed-up the SQRT function calculations and to minimize the hardware volume due to substituting some iterations by the look-up tables. The algorithm is intended for the SQRT function implementation in FPGA.

Key words: FPGA, square root, CORDIC, pipeline.

Fig.: 3. Bibl.: 6.

Вступ.Функція квадратного кореня $\sqrt{x}$ - важлива елементарна функція в наукових розрахунках, обробці цифрових сигналів та обробці зображень [1]. Наприклад, вона використовується у нейронних мережах [2]. В даний час багато задач вирішуються у програмованих логічних інтегральних схемах (ПЛІС), де також необхідно розраховувати функцію $\sqrt{x}$.

Існують різні віртуальні модулі для обчислення функції $\sqrt{x}$, які пропонуються виробниками ПЛІС та сторонніми компаніями [3]. Але ці модулі були розроблені десятиліття тому, і вони, як правило, не враховують особливості нових поколінь ПЛІС. Тому такі модулі потребують модернізації. У роботі [4] запропоновано вдосконалений алгоритм обчислення функції $\sqrt{x}$, який орієнтований на реалізацію у ПЛІС. У даній роботі пропонується ще один алгоритм, який ефективно реалізується у ПЛІС.

Алгоритми "цифра за цифрою". Виконання алгоритму розрахунку елементарної функції типу "цифра за цифрою" зводиться до повторення однотипних ітерацій, результатами яких є чергові точні цифри результату. Відомий алгоритм CORDIC, який призначений для розрахунків $\sqrt{x}$, полягає в наступному. Він обчислює функцію atanh(x/y). Але побічним результатом є функція $K\sqrt{x^2+y^2}$, а за рахунок заміни $x = A + 0.25$, $y = A - 0.25$, отримують $K\sqrt{A}$ [3,5]. Недоліками цього алгоритму є додаткове множення результату на коефіцієнт $1/K \approx 1.207$ і повторення деяких ітерацій для збіжності алгоритму.

Більш конструктивним алгоритмом є алгоритм “цифра за цифрою” обчислення функції $\sqrt{x}$ [4], який базується на наступних співвідношеннях. Для кожного числа x$\in$ [0.25; 1.0] знаходять коефіцієнти $a_i \in$ [0; 1], такі що

$$\prod_{i=1}^{\infty}(1 + a_i 2^{-i})^2 = 1.0. \quad (1)$$

Звідси

$$1/\sqrt{x} \approx \prod_{i=1}^{m}(1 + a_i 2^{-i}) \text{ або } \sqrt{x} \approx x\prod_{i=1}^{m}(1 + a_i 2^{-i}). \quad (2)$$

Отже, обчислення функції $\sqrt{x}$ полягає у виконанні конвергентного процесу, згідно з яким вираз (1) наближається до одиниці, в той час як вираз (2) наближається до шуканого значення. Алгоритм цього процесу виражається наступним чином:

```
x[0] = x; y[0] = x;
for(i = 0, i < n, i++) {
        t = x[i] + 2^(-i)*x[i];
    q = t + 2^(-i)*t;
    if (q < 1) {
       x[i+1] = q;
       y[i+1] = y[i] + 2^(-i)*y[i];}// a[i]=1
    else {
       x[i+1] = x[i];
       y[i+1] = y[i];}// a[i]=0
}
```

Результатом є $y[n] = \sqrt{x}$.

Модернізований алгоритм. Найбільшу затримку розглянутого алгоритму дає подвійне додавання зсунутих даних з результатами *t* та *q*. Ці обчислення можуть бути замінені одним етапом:

$$q = x_i + 2^{-m}x_i + 2^{-m}(x_i + 2^{-m}x_i) = x_i + 2^{-m+1}x_i + 2^{-2m}x_i. \quad (3)$$

Оскільки у сучасних ПЛІС трьохвходовий суматор реалізується як один ступінь, який будується на основі шестивходових логічних таблиць (ЛТ), то такі обчислення можуть бути виконані за один такт без додаткових затримок і витрат на обладнання.Аналіз алгоритму показує, що коли *i* досягає межі $n/2$, то старші розряди числа x_iстають одиничними а *i*старших розрядів y_iє точними старшими розрядами результату. Отже, решту отриманих бітів можна обчислити після аналізу та розрахунку різниці $1 - x_i$.

Нехай $\varepsilon_1 = 1 - x_{n/2}$, $\varepsilon_x = \sqrt{x} - y_{n/2}$. Ці величини згідно з (1) і (2) дорівнюють

$$\varepsilon_1 = 1 - x\prod_{i=1}^{n/2}(1 + a_i 2^{-i})^2 ; \varepsilon_x = \sqrt{x} - x\prod_{i=1}^{n/2}(1 + a_i 2^{-i}). \quad (4)$$

Представимо $z = \sqrt{x}\prod_{i=1}^{n/2}(1 + a_i 2^{-i})$, тоді з (4)

$$\varepsilon_1 = 1 - z^2 = (1 + z)(1 - z); \qquad \varepsilon_x = \sqrt{x}\,(1 - z). \tag{5}$$

Оскільки $z \approx 1$, то $\varepsilon_1 \approx 2(1 - z)$; і $\varepsilon_x \approx \sqrt{x}\,\varepsilon_1/2 \approx y_{n/2}(1 - x_{n/2})/2$. Тоді результат дорівнює $y_n = y_{n/2} + y_{n/2}(1 - x_{n/2})/2$ та модернізований алгоритм є наступний:

```
x[0] = x; y[0] = x;
for (i = 0; i < n/2; i++){
        q = x_i + 2^-i+1 *x_i + 2^-2i *x_i;
        if (q < 1.0){
          x_i+1 = u;
          y_i+1 = y_i + 2^-i-1 *y_i;
        }else {
          x_i+1 = x_i ;
          y_i+1 = y_i ;
        }
}
y = y_i+1 + y_i+1 *(1.0 - x_i+1)/2;
```

Експериментальні результати. Отриманий алгоритм був описаний мовою VHDL як віртуальний модуль. Цей модуль було зконфігуровано для ПЛІС Xilinx Spartan-6 для різної розрядності вхідних і вихідних даних. На рисунках 1 та 2 наведено залежність апаратних витрат в кількості ЛТ, а також максимальної тактовової частоти від розрядності *n* вхідних даних і результатів для комбінаційної і конвеєрної схем цього модуля, відповідно. Слід відзначити, що модулі з розрядністю 32 додатково мають один блок множення DSP48, а решта — чотири таких блоки.

Для порівняння, на рис. 1, 2 показані характеристики віртуальних модулів, які пропонуються компанією Xilinx Inc. Загалом, запропонований модуль має більші апаратні витрати і меншу тактову частоту. Це пояснюється тим, що модуль, який генерується засобом XilinxCoregen, описаний на рівні ЛТ і тригерів і тому адаптований до архітектури ПЛІС конкретного типу. А запропонований модуль, хоча і не потребує тривалого генерування, повинен бути зкомпільованим синтезатором, який виконує ефективну, але не оптимальну оптимізацію.

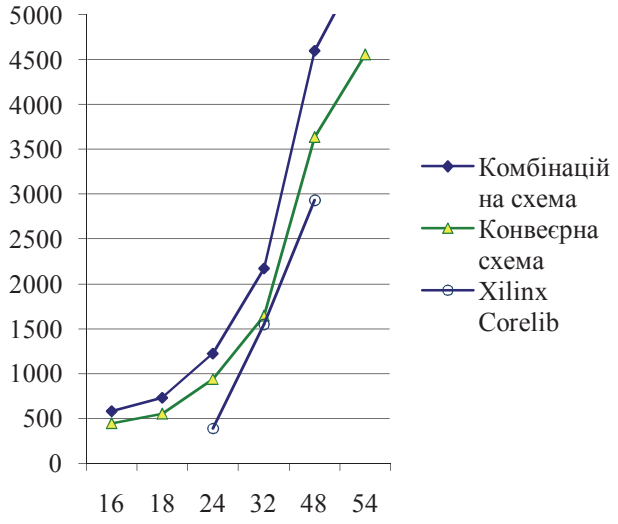


Рис. 1. Апаратні витрати блоку $\sqrt{x}$, ЛТ, в залежності від *n*

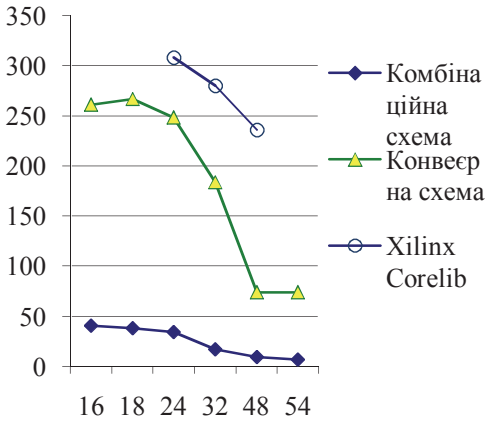


Рис. 2. Максимальна тактова частота, МГц, блоку $\sqrt{x}$ в залежності від *n*

Переваги запропонованого модуля полягають в тому, що він є безкоштовним і може бути налаштований на довільну розрядність вхідних та вихідних даних, а також для будь-якого типу ПЛІС. Крім того, запропонований модуль має нижчу латентну затримку, що є важливим, наприклад, при виконанні на його основі операцій з плаваючою комою.

Наприклад, для розрядності 24 біт, латентна затримка становить лише 15 тактів проти 25 тактів у конкурентного ядра. Це означає, що при виконанні розрахунків з плаваючою комою запропонований модуль забезпечує більшу продуктивність.

Висновки. Запропоновано модифікований алгоритм “цифра за цифрою” для обчислення функції квадратної кореня. Алгоритм відрізняється мінімізованою кількістю ітерацій, яка приблизно удвічі менша за кількість розрядів результату. Алгоритм описаний мовою VHDL і призначений для реалізації у ПЛІС будь-якої серії. Найбільш ефективна його реалізація при обчисленні функції $\sqrt{x}$ з плаваючою комою.

Список використаних джерел

1. Woods R. FPGA-based Implementation of Signal Processing Systems / J. McAllister, G. Lightbody, Y. Yi / J. Wiley and Sons, Ltd., Pub. 2008. 364 p.

2. FPGA Implementations of Neural Networks". A. R. Omondi, and J. C. Rajapakse, Eds. Springer. 2006. 360 p.

3. Yoshikawaa K. Development of Fixed-point Square Root Operation for High-level Synthesis / N. Iwanagaa, A. Yamawaki // Proc. 2nd Int. Conf. on Industrial Application Engineering. 2014. P. 16 — 20.

4. Сергієнко А. М. Реалізація функції квадратного кореня у ПЛІС / П. А. Сергієнко // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка: [зб. наук. пр.] Київ. 2014. Т.60, С. 40 — 45.

5. Бікташева С. Р. CORDIC-метод обчислення квадратного кореня / С. Р. Бікташева, Л.В. Мороз, М. Ю. Стахів // Вісн. Нац. Ун-ту "Львівська політехніка". Сер.: Електроніка : [зб. наук. пр.] Львів : Вид-во Нац. ун-ту "Львів. політехніка", 2006. С. 152—155.

6. Chen T.C. Automatic computations of exponentials, logarithms, ratios and square roots. // IBM J. Res. and Develop. 1972. №4. P. 380 — 388.

ДОВІДКА ПРО АВТОРІВ

Сергієнко Анатолій Михайлович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Sergiyenko Anatoliy – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: aser@comsys.kpi.ua

Хасан Мухамед Джамал – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Hasan Muhamad Jamal – student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".
E-mail: muhamad.software@gmail.com

Павло Анатолійович Сергієнко– студент, кафедра системного програмування таспеціалізованих комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».
Pavlo Anatolijovich Serhienko – student, Department of System Programming and Application Specific Computer Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".
E-mail: paulsrgnk002@gmail.com

Anatoly Sergienko,
Hassan Muhamed Jamal,
Pavlo Sergienko

ALGORITHM AND STRUCTURE OF THE MODULE FOR THE SQUARE ROOT CALCULATION IN FPGA

Relevance of research topic. Field programmable gate array (FPGA) is a modern element base that is designed for high performance implementation of the application specific algorithms with the fixed-point numbers. Very often, such algorithms encounter the calculation of elementary functions. There is a shortage of device designs for calculating elementary functions in the FPGA and they need to be refined.

The problem consists in the inbalance of the exisiting SQRT IP core parameters and the affordances of the FPGA system designers, as well as the modern FPGA features.

Analysis of recent research and publications. Over the past few years, there are more articles devoted to high-performance modules for calculating elementary functions, but ready-made modules are inaccessible for the widespread use. Most of the modules calculate the square root using an inefficient CORDIC algorithm.

Selection of unexplored parts of the general problem. This article is devoted to the study and implementation of a rare algorithm for calculating a square root of the iterative type in the FPGA.

Objectives. The goal is to improve the iterative algorithm for calculating the square root, intended for configuring in the FPGA.

Presentation of the main material. An improved iterative algorithm for calculating the square root function is proposed. The improvement is to reduce the difficulty of calculating the algorithm steps and to replace the last steps with the table calculations. This increases the speed of the calculation of the function more than twice.

Conclusions. The proposed modified algorithm is distinguished by a minimized number of iterations. The algorithm is described in VHDL language and is intended for implementation in FPGA of any series. Its most effective implementation is when calculating the function of a square root with a floating point.

Key words: FPGA, square root, pipeline, CORDIC.

УДК 004.942

Анатолій Сергієнко,
Мустафа Рекар Касім

АПАРАТНЕ МОДЕЛЮВАННЯ ПОШИРЕННЯ УЛЬТРАЗВУКОВИХ ХВИЛЬ У ТВЕРДОМУ ТІЛІ

MODELING THE ULTRASOUND WAVE PROPAGATION IN THE SOLID BODY BY THE HARDWARE

Розглянуто удосконалений хвилевий алгоритм моделювання поширення ультразвуку, який полягає у представленні середовища у вигляді системи хвилевих фільтрів та відрізняється тим, що завдяки реалізації багатоканальних фільтрів з програмованою затримкою, зменшується похибка моделювання дисперсійного поширення звуку. Алгоритм при реалізації у ПЛІС дає змогу виконувати моделювання у реальному часі.

Ключові слова: ПЛІС, дисперсія звуку, хвилевий фільтр.
Рис.: 3.Бібл.: 8.

The improved ultrasound propagation wave modeling algorithm is proposed, which consists in representing the medium in the form of a system of the wave digital filters. The algorithm differs in that due to the implementation of the multichannel filters with the programmable delays, the simulation error of the sound propagation dispersion is decreased. The algorithm implementing in FPGA provides the modeling in real time.

Key words: FPGA, sound dispersion, wave digital filter.
Fig.: 3. Bibl.: 8.

Вступ.Акустичні процеси в твердих тілах зазвичай моделюються в комп'ютерах за допомогою методу різницевих рівнянь, який вимагає великих обчислювальних ресурсів [1]. Для мінімізації цих ресурсів часто використовують метод цифрових хвилеводів, або методDWG (digital waveguide), який забезпечує високошвидкісні обчислення [2]. Багато успішних прикладів впровадження методуDWG показано в моделюванні та звуковому синтезі струнних, духових та віртуальних музичних інструментів, а також в вокодерах [2, 3, 4]. Але ця модель не враховує розсіювання (дисперсію) хвиль у поширенні звуку.

У даній роботі пропонується модифікація методу DWG, яка забезпечує моделювання звукової дисперсії в твердій смужці.

Основи моделі DWG.DWG – це цифрова модель розповсюдження хвиль в ідеальному хвилеводі. Вона базується на принципах, описаних у роботі [5], але адаптованих до моделювання звукових хвиль [2,4]. При цьому у i-й точці хвилеводу розглядаються прямі $f_i = Rv_f$і зворотні хвилі $b_i = –Rv_b$, де f_iта b_i – це тиск прямих і зворотних (віддзеркалених) хвиль, відповідно R – хвилевий імпеданс, v_f, v_b – швидкості частинок речовини. Для жорсткого циліндра з площею перетину A цей імпеданс дорівнює $R = \rho c_o A$, де ρ – густина матералу, c_o– швидкість поздовжних хвиль. Реальне значення тиску дорівнює сумі тисків прямої та зворотної хвиль$u = f_i + b_i$.

У цифровій моделі DWG сигнали квантуються з частотою дискретизації F_s, яка, щонайменше, в два рази перевищує максимальну частоту хвиль, які досліджуються. Тоді i-й сегмент хвилеводу довжиною L виглядає як дві лінії затримки на $n = L/(c_o F_s)$ тактів сигналу дискретизації.

Кілька однорідних хвилеводів з'єднануються між собою в одній точці за допомогою вузла адаптера, який також називають вузлом розсіювання. Функціонування адаптера повинне підлягати закону Кірхгофа. Для випадку з'єднання двох хвилеводів це означає виконання рівнянь:

$$v_{b1} = rv_{f1} + (1 - r)\, v_{f2}; \qquad (1)$$
$$v_{b2} = (1 + r)\, v_{f1} - rv_{f2};$$

д $r = (R_2 - R_1)/(R_2 + R_1)$ – коефіцієнт віддзеркалювання. Так само, коли підключено три хвилеводи, то у вузлі адаптера розраховуються наступні формули:

$$v_{b1} = (2g_1 - 1)\, v_{f1} + 2g_2\, v_{f2} + 2g_3\, v_{f3};$$
$$v_{b2} = 2g_1\, v_{f1} + (2g_2 - 1) v_{f2} + 2g_3\, v_{f3}; \qquad (2)$$
$$v_{b3} = 2g_1\, v_{f1} + 2g_2\, v_{f2} + (2g_3 - 1)\, v_{f3};$$

де v_{f1}, v_{f2}, v_{f3} – хвилі, що входять у вузол, v_{b1}, v_{b2}, v_{b3}, – хвилі, що виходять з вузла, g_1, g_2, g_3 – питомі імпеданси хвилеводів, такі що $g_1 + g_2 + g_3 = 1$, та $g_i = R_i/(R_1 + R_2 + R_3)$, $i = 1, 2, 3$.

Удосконалена модель DWG. Згідно з теорією поширення акустичних хвиль у твердих тілах, які мають обмежені розміри, наприклад, у стрижнях, розглядають основні – повздовжні хвилі з фазовою швидкістю c_p та вторинні – поперечні хвилі зі швидкістю c_s, причому $c_s < c_p$ [6]. При взаємодії цих хвиль вони можуть перетворюватись одна в іншу. Це можна промоделювати при з'єднанні хвилеводів з такими хвилями у вузлі адаптера.

Крім того, повздовжні хвилі мають дисперсію, тобто, їх швидкість c_pзалежить від довжини хвилі λ. Наприклад, швидкість повздовжних хвиль у циліндрі оцінюється формулою [6]:

$$c_p = c_o(1 - \nu^2 p^2 a^2/\lambda^2) = f(\lambda), \qquad (3)$$

де ν – коефіцієнт Пуассона, що дорівнює 0.29 для сталі, a – радіус циліндра. Тому модель DWG повинна бути скоректована відповідно до цієї формули.

Модифіковану модель DWG деякого стрижня або циліндра можна представити структурою, яка показана на рис.1. Ця модель складається з лівого (LA) та правого (RA) адаптерів, хвилеводу P поздовжніх хвиль та хвилеводу S поперечних хвиль. У кожному з трьохпортових адаптерів обчислюються вирази (2). Їхні порти, які означають торці стрижня, підключені до мереж, які реалізують віддзеркалення хвиль з коефіцієнтом придушення s_L та s_R згідно з рівняннями (1). Сигнал збудження v_{in} подається у модель через суматор.

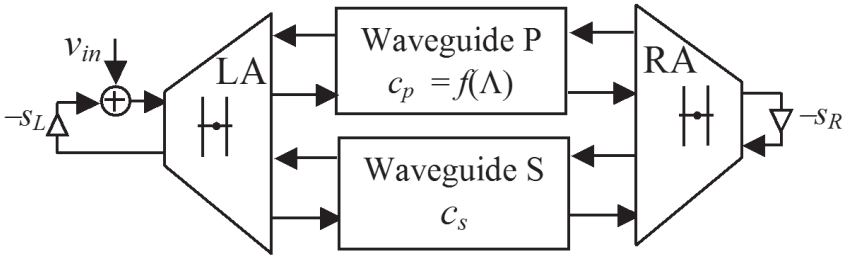


Рис. 1. Структура моделі стрижня

Хвилевід S виконує затримку на задану кількість тактів частоти дискретизації та певне ослаблення хвилевого сигналу. Хвилевід P складається з кількох каналів, кожен з яких має окремі діапазон частот пропускання, затримку і згасання в залежності від відповідної довжини хвилі λ згідно з (3).

Отже, модифікована модель DWG здатна приймати до уваги дисперсність поширення звуку у твердому тілі, тобто, залежність фазової швидкості звуку від його довжини хвилі або частоти.

Експериментальні результати. Була розроблена програмна модель стрижня зі структурою як на рис. 1. В ній до восьми каналів хвилеводу P реалізуються на основі смугових фільтрів на базі фазового хвилевого фільтра, який описано у [7]. Кожен канал хвилеводу має динамічно настроюваний смуговий фільтр, який реалізовано так, як показано в [8]. Сумарна амплітудно-частотна характеристика хвилеводу S показана на рис.2. Тут частота *f* вимірюється в частках частоти дискретизації *Fs*. На рис.2 також показані частотні характеристики двох сусідніх каналів хвилеводу.

Модель збуджувалась вузьким ультразвуковим імпульсом, а інтенсивниість звуку *v* вимірювалась на виході вузла суматора (див. рис.1). Часова діаграма, яка представляє хвильовий сигнал, показана на рис.3. Ця діаграма показує, що хвилі дійсно зазанають дисперсії під час розповсюдження через модель стрижня.

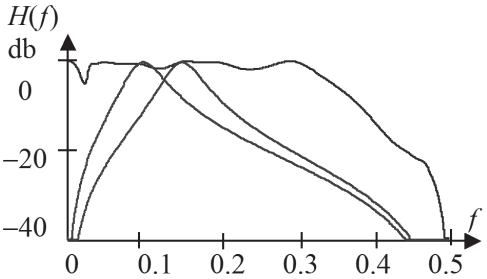


Рис. 2. Сумарна амплітудо-частотна характеристика хвилеводу P

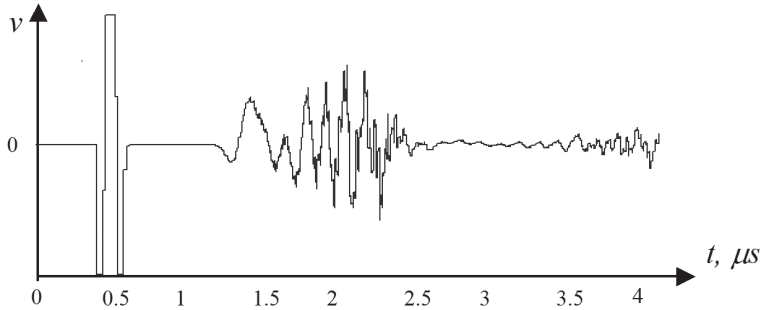


Рис. 3. Часова діаграма поширення ультразвукового імпульсу у стрижні

Модель стрижня описана мовою VHDL, і тому вона може бути реалізована як у симуляторі VHDL, так і в ПЛІС. Модель, описана стилем для синтезу і при конфігуруванні у ПЛІС XilinxSpartan-6, містить 2680 логічних елементів, 60 блоків множення DSP48, та 16 блоків пам'яті BlockRAM. При цьому забезпечується частоту дискретизації сигналу $F_s <$ 100 МГц. Така частота забезпечує моделювання в режимі реального часу.

Висновки. Запропоновано модифікований метод DWG для моделювання твердого тіла, який враховує дисперсність поширення хвиль. Метод заснований на моделі цифрового хвилеводу, затримка якого залежить від довжини хвилі. Приклад моделі стрижня, описаної на VHDL, показує ефективність цієї моделі.

Список використаних джерел

5. Самарский А.А. Математическое моделирование. Идеи. Методы. Примеры / Михайлов А.П. / М.: Физматлит, 2001. -320 с.

6. Smith III J. O. Physical Modeling Using Digital Waveguides // Computer Music Journal, V.16. 1993. №4.P.74-91.

7. Karjalainen M. Digital Waveguides versus Finite Difference Structures: Equivalence and Mixed Modeling / C. Erkut // EURASIP Journal on Applied Signal Processing 2004. №7. P. 978–989.

8. Маркел Дж. Д. Линейное предсказание речи. – М.: Связь. 1980. 308 с.

9. Fettweis A. Wave digital filters: Theory and practice // Proc. IEEE, V74, 1986. №2. P. 270 – 327.

10. Kolsky H. Stress Waves in Solids. Dover Publications Inc. 2012.

11. Regalia P. A., The Digital All-Pass Filter: A Versatile Signal Processing Building Block / S. K. Mitra, P. P. Vaidyanathan //Proc. IEEE. V.76. 1988. №1. P. 19-37.

12. Сергиенко А. М, Перестраиваемые цифровые фильтры на ПЛИС / Т. М. Лесик // Электронное моделирование, Т.32. 2010, №6. С. 47-56.

ДОВІДКА ПРО АВТОРІВ

Сергієнко Анатолій Михайлович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Sergiyenko Anatoliy – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: aser@comsys.kpi.ua

Мустафа Рекар Касім – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Mustafa Rekar Quasim – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: rekoqasm91@gmail.com

Anatoly Sergiyenko,
Mustafa Rekar Quasim

HARDWARE MODELING OF ULTRASONIC WAVES EXPANSION IN SOLIDS

Relevance of research topic. Numerical modeling of the physical phenomenon of the wave propagation is the basis of many studies and implementations at the fields of acoustics, ultrasound technology, and radio electronics. The hardware implementation of such a simulation allows us to significantly accelerate and improve the efficiency of devices that use the wave processes. The use of the field programmable logic arrays (FPGAs) to simulate the wave processes allows us to increase speed, reduce power consumption and increase the reliability of tools that use such simulation.

The problem. The known methods and tools do not provide the real-time ultrasound wave expansion modeling.

Analysis of recent research and publications. Over the past few years, there are more articles devoted to hardware sound propagation techniques, which tend to use systems based on graphic accelerators that perform parallel floating point calculations. Their drawbacks are high cost, high power consumption, and low reliability.

Selection of unexplored parts of the main problem. This article is devoted to the study and analysis of the approach based on the digital waveguides implemented in FPGA.

Objectives. The task is to create a method for designing the high-performance processors for modeling the acoustic phenomena in the application specific processor, configured in FPGA.

Presentation of the main material. The advanced ultrasound propagation wave modeling algorithm, which consists in representing the medium in the form of a system of wave filters, is proposed. It differs in that the error of modeling of dispersion of the sound propagation is decreased due to the implementation of multichannel filters with the programmable delay.

Conclusions. The modified method of digital waveguides for solid state modeling, which takes into account the dispersion of wave propagation, is proposed. An example of a solid bar model described on VHDL shows the performance of it.

Key words: FPGA, sound dispersion, wave filter.

UDC 631.432:626.86

Демчук Дмитро, Валерій Сімоненко

РОЗРОБКА СИСТЕМИ АВТОМАТИЧНОГО ПРОГНОЗУВАННЯ ЗРОШЕНЬ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР З ВИКОРИСТАННЯМ МОДУЛЬНИХ ПРИСТРОЇВ

DEVELOPMENT OF A SYSTEM FOR AUTOMATIC FORECASTING IRRIGATION OF AGRICULTURAL CROPS USING MODULAR DEVICES

У статті розглядається питання автоматизування, а саме прогнозування зрошення сільськогосподарських угідь. Для прогнозування використовуються пристрій-збирач даних (пристрій спостереження) з місця поливу, сервер обробки отриманої інформації.

Ключові слова: вологоперенесення, прогнозування зрошення, пристрій спостереження, модульність, сервер

Рис.: 5, Бібл.:8

This paper concerns the task of smart using resources for big watering systems. Types of other existing devices used in watering systems are investigated. Evaluation criteria are defined. Developed the devices prototype responding the specified conditions. The text topic is assigned using a set of keywords.

Key words: moisture transfer, irrigation forecasting, modularity, observing device, server

Fig.: 5, Bibl.:8

Target setting. Due to long-term intensive agricultural use of irrigated land, increased the diversity of agro-ecological conditions, changes in soil formation, ecological and land reclamation conditions, the economic problems of recent years have led to unsustainable scarcity of resources in irrigated agriculture and deterioration of the technical condition of irrigation systems. Effective and stable irrigation of a particular agricultural crop has become an actual topic in the recent years.

Actual scientific researches and issues analysis. In connection to the invention of new more chipper devices and approaches in the modular electronics, the topic of forecasting irrigation using simple devices has become more studied in recent years.

Uninvestigated parts of general matters defining. Despite a large number of works and devices devoted to the prediction of moisture reserves of the soil, the problem of creating system using a universal modular device collector and transmitter and server with forecasting was not considered.

The research objective. The purpose of this paper is to develop alert system for forecasting irrigation using cheap universal modular devices and server data handler.As a solution, the article will focus on creating a model consists of universal device that periodically collect and send data to central server, which processes received data and informs users about irrigation forecast.

The statement of basic materials. For efficient and stable irrigation of a particular agricultural crop, it is necessary to know the properties of soil and culture, to

identify and predict certain changes in the soil on which this crop is growing. Among the main properties of the soil is the ability to produce and contain the water needed for plants. Study of the quantitative patterns of moisture transfer in the soil involved by Voronin, Muromets and others. The table 1 shows the algorithm for forecasting irrigation.

General model structure. The structure of the model is a simple automatic device (one or more devices could be) and server. Device has modular structure, so it consists of three parts – data collector, data transmitter that periodically send collected data and autonomous work module. Server has a data receiver, analyzer block, user interface part and database. Fig. 1 illustrates the principle of interaction between the device and the server. Fig. 2 illustrates device`s modularity.

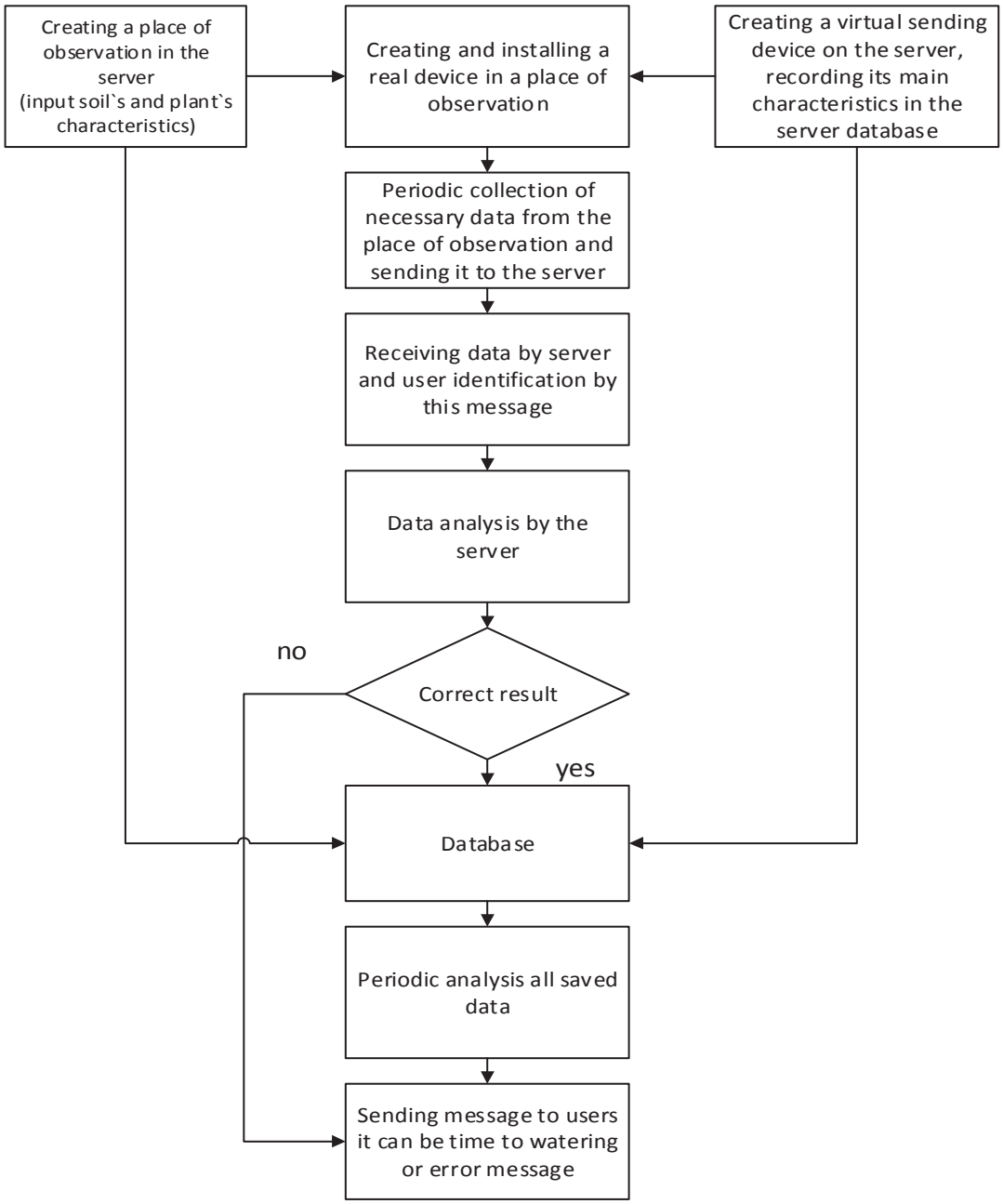


Fig. 1. Structure of a system for forecasting irrigation of agricultural crops using ready simple device

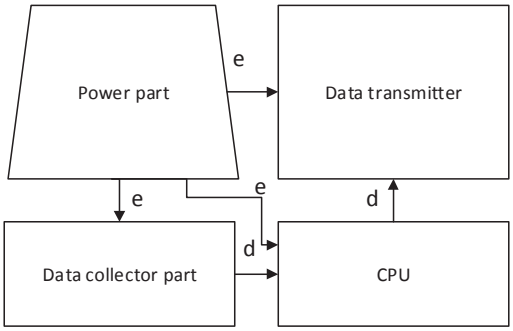


Fig. 2 Device`s modularity (e –energy, d - data)

Experiments. The device was created using Arduino as CPU, SIM800L as data transmitter and temperature and pressure sensors. Device named as tenso-meter. The figures depict the temperature graphs of the sensor located in the vacuum tube, which measures vacuum pressure.

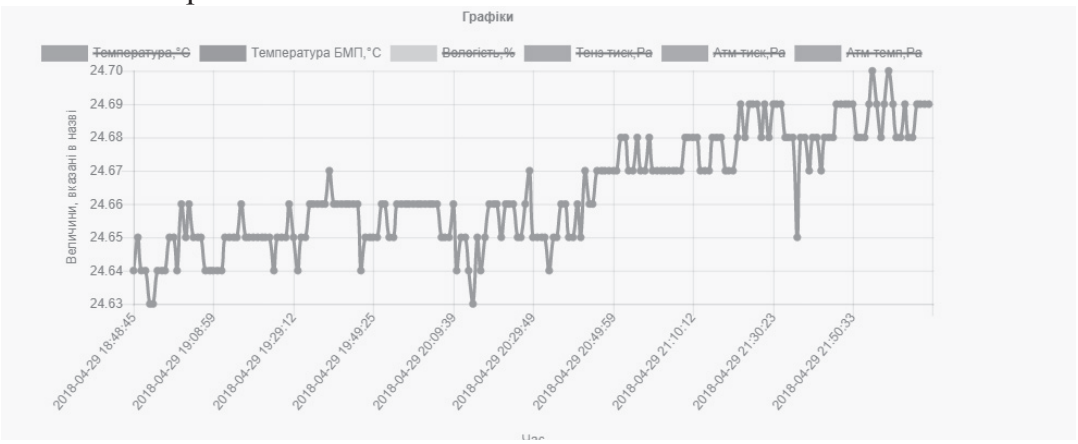


Fig. 3. Graph of the temperature sensor

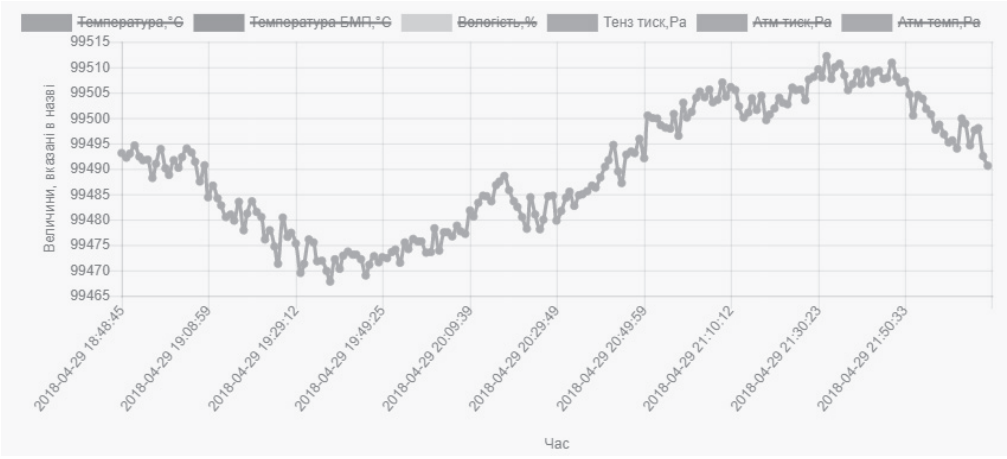


Fig.4. Graph of pressure sensor in the vacuum tube

id	Назва	Прив'язка	Ред.	Вид.	Історія	Графіки	ВКЛКлапан	Полив
1	BMP180	Testing(Місто Київ)	=	-	+	+	Вкл	12:30 29.04.2018 10мм

Fig.5. Result of analysis by the server

Having analyzed the data on the server using generalized the Stoic and other methods calculating water consumption for irrigated agricultural fields, which requires mid-hour temperature, humidity, pressure difference between the vacuum tube and the atmosphere, common history all of it, we can get directions on the prediction of forecast soil moisture content. Setting into place properties in the server the physical and chemical properties of the soil and the biological properties of the agricultural crop, it is possible to predict the estimated time and the amount of water needed. From Figure 4 you can see that irrigation requires 2018-04-29 at 12:30 in the amount of 10mm in the observation place.

Conclusions. The paper has demonstrated simple working system for forecasting irrigation. A simple modular-based observing device and server has been developed. It can be seen that the system produces qualitative results. The forecast is calculated up to five days. The system has no drawbacks to similar systems, such as the price of the device. Any disadvantages of the system are primarily due to the shortcomings of the device. It also would be interesting to test other types of device or methods of forecasting.

References

1. Воронин А. Д. Основы физики почв: учеб. пособие / А. Д. Воронин – М. : Изд-во Моск. ун-та, 1986. – 244 с.

2. Муромцев Н. А. Мелиоративная гидрофизика почв / Н. А. Муромцев – Л. : Гидрометеоиздат, 1991. – 265 с.

3. Рекомендації з оперативного контролю та управління режимом зрошення сільськогосподарських культур із застосуванням тензіометричного методу / М. І. Ромащенко, В. М. Корюненко, М. М. Муромцев – К. : ІВПіМ, 2012. – 71 с.

4. Офіційна документація по SM200-UM-1.1 [Електронний ресурс] – Режим доступу:https://www.delta-t.co.uk/wp-content/uploads/2016/11/SM200-Soil-Moisture-Sensor-UM.pdf(дата звернення 15.03.2018)

5. CropSense, офіційний інтернет-портал [Електронний ресурс] – Режим доступу: https://www.cropsense.uni-bonn.de/(дата звернення 15.03.2018) – CROP.SENSe.net

6. Metos, офіційний інтернет-портал[Електронний ресурс]Режим доступу: http://metos.at/home/ru/imetos-eco-d3/(дата звернення 15.03.2018) –METOS

7. Tsivinsky, G. V., Pendak, N. V., Idayatov, V. A.: Instruction on operative calculation of irrigation regimes and forecast of irrigation of agricultural crops due to lack of moisture stores. 2nd edn. Ukrainian Ecological League, Kherson (2010). (In Ukrainian)

ДОВІДКА ПРО АВТОРІВ

Сімоненко Валерій Павлович – професор, доктор технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Simonenko Valerii–Professor, Doctor of Technical Sciences, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: svp@comsys.ntu-kpi.kiev.ua

Демчук Дмитро Олександрович – студент-бакалавр, доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Demchuk Dmytro - student bachelor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail:ddima199703@gmail.com

Демчук Дмитро,
Валерій Сімоненко

РОЗРОБКА СИСТЕМИ АВТОМАТИЧНОГО ПРОГНОЗУВАННЯ ЗРОШЕНЬ СІЛЬСЬКОГОСПОДАРСЬКИХ КУЛЬТУР З ВИКОРИСТАННЯМ МОДУЛЬНИХ ПРИСТРОЇВ

Актуальність теми дослідження. Проблема нестабільного дефіцитного ресурсозабезпечення у зрошувальному землеробстві та погіршення технічного стану зрошувальних системстає більш актуальною у зв'язку з економічними проблемами останніх років. Дана робота присвячена проблеміавтоматичного прогнозування поливів на заданому місці із вказаними характеристиками ґрунту і сільськогосподарської культури.

Постановка проблеми. Відсутність автоматичних систем з доступним пристроєм спостереження. Проблема вибору пристрою, який мав універсальний процесор і передавач, проте змінні модулі збирання інформації з навколишнього середовища.

Аналіз останніх досліджень і публікацій. Мала кількість діючих аналогічних систем, які при цьому досить часто використовують закриті методи обчислень, а також використовують дорогі пристрої спостереження, які надаються власниками даних систем. Дані пристрої не є модульними, тому не є повністю універсальними.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена розробці простої відкритої системи для прогнозування поливів на заданому місці з використанням модульного пристрою спостереження та сервера приймача-аналізатора.

Постановка завдання. Завданням є розробка системи оповіщення для прогнозування зрошення з використанням дешевих універсальних модульних пристроїв та обробника серверних даних.

Викладення основного матеріалу. Для ефективного та стабільного зрошення певної сільськогосподарської культури необхідно знати властивості ґрунту та даної культури, визначити та прогнозувати певні зміни у ґрунті. Серед основних властивостей ґрунту є здатність виробляти і містити воду, необхідну для рослин. Використовуючи узагальнену формулу Штойка розроблено систему, яка прогнозує кількість і час зрошення на період до 5 днів.

Висновки. У роботі продемонстрована проста робоча система прогнозування зрошення. Був розроблений простий пристрій спостереження. Видно, що система показує якісні результати. Прогноз розрахований до п'яти днів. Система недоліків аналогічних існуючим системам, таких як ціна пристрою, закритий алгоритм аналізу. Будь-які інші недоліки системи, насамперед, пов'язані з недоліками пристрою. Надалі актуально протестувати типи модульних частин пристроїв або інші методи прогнозування.

Ключові слова: вологоперенесення, прогнозування зрошення, пристрій спостереження, модульність, сервер.

УДК 681.511: 629.127 + 62-526

Даніленко Наталія,
Сімоненко Андрій

МОДЕЛЬ ПІДВОДНОЇ ЖИТТЄЗДАТНОЇ СИСТЕМИ НА ОСНОВІ МІКРОКОНТРОЛЕРА ATMEGA328 ПЛАТФОРМИ ARDUINO UNO

MODEL OF UNDERWATER VIABLE SYSTEM BASED ON ATMEGA328 MICROCONTROLLER ARDUINO UNO PLATFORM

У статті наведено принципруху, що може бути застосований для моделі підводної життєздатної системи. Показано спосіб поєднання електричних елементів для створення рухомого підводного апарату на основі платформи Arduino. Зроблено аналіз програмної реалізації моделі, розглянуто функцію «автопілот» для даної системи. Наведено результати тестування.

Ключові слова: підводна життєздатна система, коливання, Arduino Uno, ефект Магнуса, АПА.

Рис.: 5. Бібл.: 14.

The article presents the principle of motion, which can be applied to the model of the underwater viable system. The method of combining electrical elements to create a moving device (submarine) based on the Arduino platform is shown. The analysis of software implementation of the model is made, the "autopilot" function for this system is considered. The results of testing are presented.

Key words: underwater viable system, fluctuations, Arduino Uno, Magnus effect, AUV.

Fig.: 5. Bibl.: 14.

Актуальність теми дослідження. Створення моделі життєздатної системи[1], такої, як малогабаритний неприв'язний автономний підводний апарат (АПА)[2], - досить складний та часозатратний процес, адже необхідно забезпечити адаптивність такої системи та врахувати безліч факторів впливу водного середовища. Для пришвидшення моделювання та реалізації ідей, особливо для конструювання механічних моделей з використанням відповідного програмного забезпечення, все частіше використовується апаратна обчислювальна платформа Arduino. Її низька вартість і крос-платформеність приваблює розробників. Так, за допомогою Arduino можна створювати автономні апарати.

Постановка проблеми. Зазвичай при моделюванні неприв'язних АПА використовуються гвинтові рушії[3], що значною мірою впливає на швидкість руху апарата. ККД гвинтового рушія рідко перевищує позначки 0,5-0,7, при цьому верхня межа вважається дуже високою і досягається лише на малооборотних гвинтових рушіях[3,4]. При швидкому русі (неживого) предмета через воду завихрення утворюються лише позаду предмета, як і у випадку з використанням гвинтових рушіїв. За законом Бернуллі, тиск у завихреннях падає, що створює гальмівний ефект. Постає проблема значного зниження опору води. Також водорості, сітки та інші предмети намотуються на гвинтовий рушій,

гальмуючи рух апарата та призводячи до його несправності. Постає питання заміни гвинтових рушіїв на рушії біонічного типу (розглядаємо спосіб руху ската з великими грудними плавниками та тонким і слабким хвостом) для забезпечення більшої плавності рухів, швидкості руху апарата, його маневреності.

Аналіз останніх досліджень і публікацій. Протягом останніх двох десятиліть з'являється все більше статей, присвячених моделюванню підводних життєздатних систем, зокрема завдяки появі нових технічних засобів моделювання та доступності апаратно-обчислювальної платформи Arduino і подібних до неї платформ. Проте підходи до реалізації підводних автономних апаратів ще не достатньо вивчені.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена вивченню та аналізу запропонованої підводної моделі життєздатної системи, зокрема принципу руху, що може бути застосований для даної моделі. Дослідження зосереджене на вивченні способу руху підводних мешканців (скат) та можливості реалізації даного способу при моделюванні системи, а також на створенні відповідного програмного забезпечення (функція «автопілот»).

Метою даної роботи є дослідження способу руху підводного об'єкта (ската) і відтворення даного способу за рахунок створення моделі життєздатної системи на основі мікроконтролера ATmega328 платформи Arduino UNO (малогабаритний неприв'язний АПА).

Постановка завдання. Завданням є створити модель автономного підводного апарата як життєздатну систему, принцип руху якої базується на складних коливаннях, а також забезпечити адаптивність такої системи та врахувати фактори впливу водного середовища.

Викладення основного матеріалу. Спосіб руху ската за рахунок складних коливань. Скати – це пластиножаберні хрящові риби, які вважаються найближчими родичами акул. Зазвичай вони мають форму плоского диска[5]. Коли скат переміщується у воді, його грудні плавники здійснюють складні коливання[6], подібні до змахів крил пташок, а хвіст виконує функцію руля, дякуючи чому скат може виконувати плавні рухи, ковзаючи в товщі води, та розвивати дуже високу швидкість, яка на сьогоднішній день лише частково доступна сучасним підводним апаратам. Скати «літають» у воді, середовищі, густина якого приблизно у 800 разів більша густини повітря. Тому техніка «польоту» цих риб, хоч і нагадує політ пташок, все ж є менш складною. Скатам необхідно не лише «парити» у воді (рис. 1), але й іноді плисти проти водного потоку чи течії, а наявність відразу двох сил: виштовхувальної сили Архімеда та сили тяжіння, що повністю (або частково) компенсують одна одну, - змушує ската вивертати грудні плавники так, щоб кожен елемент плавника здійснював рухи, які схожі на рухи дельфіна. У дельфінів підйомна сила, за допомогою якої вони рухаються, залежить від форми тіла, випуклої зверху та згладженої знизу, що перетворює тіло дельфіна в аналог крила літака.

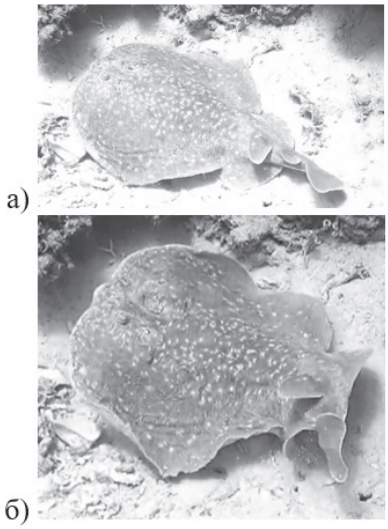


Рис. 1. Вид збоку на ската, що піднімає плавники (а),
і на ската, що опускає плавники (б)

Розглянемо траекторію руху плавника ската (кореневої та кінцевої частини), з урахуванням викиду кінця плавника вперед. Траекторія руху плавника ската збігається з траекторією руху крила чайки (рис. 2).

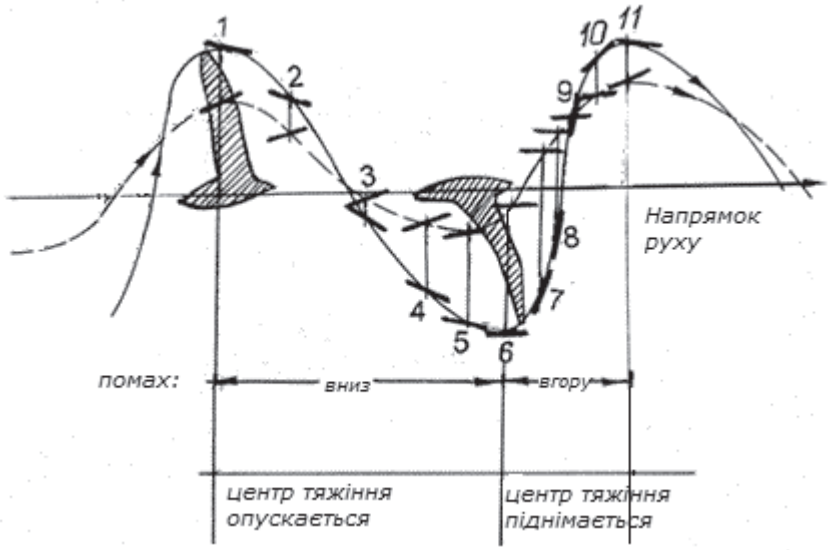


Рис. 2. Траекторія руху крила чайки під час спостереження збоку

Аналізуючи динаміку підйомних сил, що виникають при переміщенні, за увесь цикл, коли скат рухає плавниками та змінює форму від варіанту 1 (рис.3, зверху) до варіанту 6 (рис.3, знизу), бачимо, що сили плавно змінюються як у напрямку, так і за модулем, але напрямок результуючої сили буде приблизно спрямований туди, куди буде рухатись скат. При поворотах ската даний механізм дає збій, тому скат регулює рух на поворотах за допомогою хвоста. Таким чином, це не скат пливе, відштовхуючи воду назад, а вода за рахунок перепадів тиску з протилежних боків плавників ската (вертикальна площина) переміщує його практично без затрат енергії з боку ската. Ковзання у воді

досягається через той же ефект: вода за рахунок перепадів тиску з протилежних боків тіла ската (горизонтальна площина) плавно штовхає йоговперед.

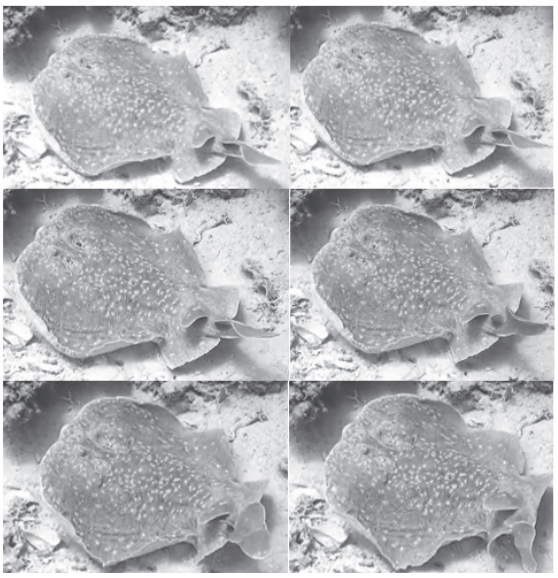


Рис. 3. Рухи ската, розбиті на фази при плаванні

Відомо, що в основі підйомної сили закладений ефект Магнуса від взаємодії приєднаного до тіла вихра з набігаючим водним потоком [7]. Саме ефект Магнуса є тим механізмом, що дозволяє скатам переміщуватися з високою швидкістю і малими затратами енергії.

При застосуванні даного механізму переміщення [8, 9] для малогабаритного АПА потрібно динамічно змінити «геометрію» апарату. Щоб це зробити, необхідно, щоб апарат був обладнаний механічними «плавниками» та «хвостом», за допомогою яких з'явиться можливість протистояти опору води і за рахунок цього розвинути значну швидкість.

Для створення «плавників» і «хвоста» можна використати властивість ската відтворювати складні коливання [8,10, 11] з більшою чи меншою амплітудою, однаково вигинаючи грудні плавники, але з різною силою.

Існують точні двигуни для управління параметрами руху – сервоприводи. Саме їх можна використати в якості з'єднувальних повертальних механізмів для частин механічного «хвоста» та «плавників». До того ж вони легко програмуються у середовищі Arduino IDE (середовище має спеціально створені бібліотеки під різні типи сервоприводів).

Переваги використання двигунно-рушійного комплексу, побудованого на складних коливаннях:

а) збільшення швидкості плавання АПА;

б) більша маневреність моделі;

в) «хвостовий» та «плавникові» рушії безпечні для людини та навколишнього середовища (водорості не заплутуються навколо «хвоста» та «плавників»);

г) ККД становить 0,6-0,85.

Платформа Arduino як засіб реалізації автономного апарату. Автономний апарат досить легко реалізувати за допомогою платформи Arduino. Arduino – це апаратна обчислювальна платформа, яка легко програмується, заснована на простому

Open Hardware проекті, що базується на вмонтованих мікроконтролерах в єдину друковану плату із вбудованою підтримкою вводу/виводу та стандартизованою мовою програмування. Мова програмування Arduino ґрунтується на середовищі підключення та в основному на мові C / C ++. Arduino є хорошою альтернативою у випадку недоступності більш «просунутих» мікроконтролерів, завдяки її загальнодоступності, відносно низькій ціні, гнучкості та крос-платформеності. Типова плата Arduino містить мікроконтролер (МК), цифрові та аналогові лінії вводу/виводу та інтерфейс UART або USB для підключення до комп'ютера.

Однією з модифікацій плати Arduino є клон Arduino Uno Rev-3. За допомогою цієї платформи можна змоделювати автономний апарат, запрограмувати та протестувати його.

Arduino Uno Rev-3 – це пристрій на основі МК ATmega328. До його складу входять усі необхідні для зручної роботи з МК компоненти: 14 цифрових входів/виходів (з них 6 можуть використовуватися в якості ШІМ-виходів), 6 аналогових входів, кварцевий резонатор на 16 МГц, роз'єм USB, роз'єм живлення, роз'єм для внутрішньосхемного програмування (ICSP) і кнопка Reset.

Плата має flesh-пам'ять 32 КБ (ATmega328), з яких 0.5 КБ використовуються завантажувачем.

Параметри плати Arduino Uno Rev-3 зовсім не відрізняються від параметрів оригінальної плати Arduino Uno, проте ціна в кілька разів нижча.

Використання сучасної елементної бази дозволяє не лише вдосконалити старі методи проектування та конструювання, а й створити нові. Також застосовуючи елементну базу з покращеними технічними та експлуатаційними характеристиками, зокрема платформу Arduino Uno, міні-сервоприводи SG90, інфрачервоні датчики руху HC-SR501 тощо, можна суттєво скоротити витрати на створення продукту.

Варто зазначити, що використання такої елементної бази призводить до значного зменшення розмірів та зниження маси продукту, що при його розробці відіграє значну роль.

Використання даної елементної бази забезпечує автоматичну обробку інформації.

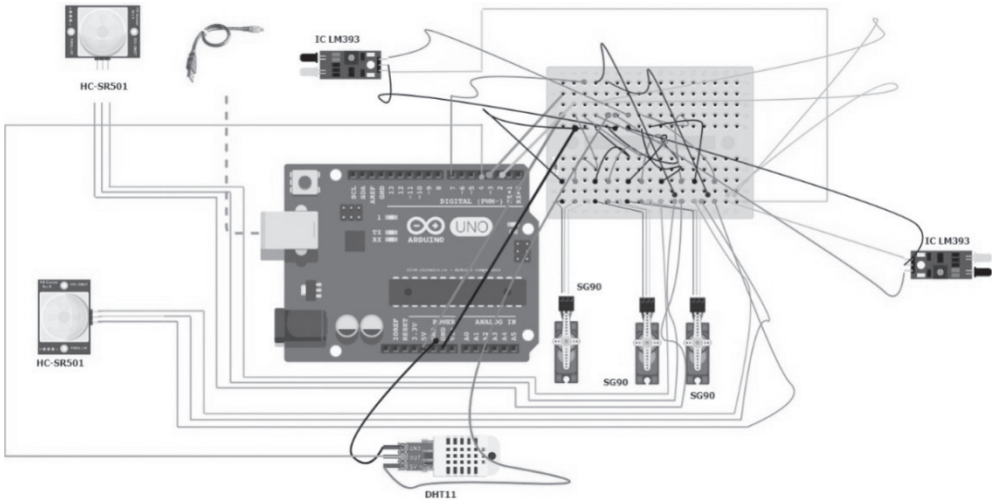


Рис. 4. Схема поєднання елементів (плата ArduinoUno, безпаєчна макетна плата, сервоприводи SG-90, датчики приближення ICLM393, датчик температури DHT11, датчики руху HC-SR501)

При моделюванні продукту добре використовувати принцип подвоєння[12] таких елементів, як інфрачервоні датчики приближення і руху, що дозволяє підвищити надійність роботи моделі. Використання ретельно підібраної елементної бази дає можливість значно зменшити масу і габарити готового продукту, що забезпечує надійність системи.

На рис. 4 наведено схему з'єднання елементів моделі. Для з'єднання було використано перемички типу мама-мама, папа-мама, безпаєчну макетну плату.

Аналіз програмної реалізації системи.Життєздатна система – це автономна система, що може сама адаптуватись до змін навколишнього середовища, автоматично змінюючи свою структуру або алгоритм функціонування з метою досягнення оптимального стану. Задля створення такої системи необхідно повністю продумати логіку її автономного функціонування, передбачаючи наперед виняткові ситуації, та створити функцію «автопілот».

Для програмування мікроконтролера ATmega328 ArduinoUno можна використовувати ArduinoIDE, Programino, B4R, CodeBlocksforArduino та інші середовища.

Середовище ArduinoIDE містить увесь необхідний мінімум для розробки програм: написання коду, перевірка коду, компіляція, завантаження скетча в Arduino, монітор послідовного порту тощо.

Programino – небезкоштовне середовище розробки. Однією з головних вимов цього середовища є встановлення ArduinoIDE. Дане середовище має дуже корисний інструмент AnalogPlotter, він дозволяє візуалізувати те, що відбувається в COM-порт від Arduino. AnalogPlotter найчастіше використовують для відображення показників аналогових датчиків: температури, вологості, освітленості тощо.

Середовище розробки B4Rунікальне тим, що використовує мову Basic, а не C.

Існують і інші середовища розробки для Arduino, наприклад Codeblocks for Arduino. Але простота і легкість у використанні середовища ArduinoIDE заслуговує на особливу увагу. Середовище постійно оновлює бібліотеки, за потреби завантажує приклади використання тієї чи іншої бібліотеки, а також дає можливість створювати бібліотеки, драйвери для своїх пристроїв, аматорських плат тощо.

Програма для даної моделі була написана в ArduinoIDE.

Щоб відтворити рухи плавників та хвоста справжнього ската, необхідно забезпечити постійний рух механічних плавників та хвоста [13]. Щоб досягнути такого ефекту, модель повинна складатися з 4-х частин: тулуб, два «плавники» і «хвостова» частина (рис. 5), що поєднані між собою сервоприводами. Дана система оснащена датчиками приближення та руху, що дає роботизованому «скату» змогу обпливати перешкоди та фіксувати рух підводних об'єктів, що випромінюють тепло. Також модель містить датчик температури та вологості і визначає температуру води, а також рівень вологості всередині пристрою[14].



Рис. 5. Спрощений вигляд системи «скат»

Таким чином на програмному рівні потрібно забезпечити безперервну роботу трьох сервоприводів. Для цього можна використати бібліотеку Servo.h, яка вбудована в ArduinoIDE. Для кожного з трьох сервоприводів обов'язково

необхідно виділити окремий ШІМ-вихід на платі ArduinoUnoRev-3. У програмі слід зазначити, до якого порту «прив'язаний» сервопривід. Це можна зробити за допомогою функції Servo.attach(number_of_pin).

Інфрачервоні датчики приближення та руху (в даному випадку IC LM393 та HC-SR501) потрібно налаштувати апаратно та програмно.

Датчик руху HC-SR501 може працювати у двох режимах: H (сигнал на виході зникне лише після закінчення терміну затримки з моменту останньої фіксації руху в зоні пошуку) та L (датчик, зафіксувавши рух, подає на вихід сигнал високого рівня). Крім того на даній платі встановлено два змінних опори, що регулюють зону спрацювання (Sensitivity Adjust) та затримку вимкнення (Time Delay Adjust). Програмно датчик легко налаштовується, не вимагає жодних додаткових бібліотек. У середовищі ArduinoIDE потрібно вказати цифровий (не ШІМ) вхід та для коректної роботи датчика визначити час калібрування.

Датчик приближення IC LM393 теж налаштовується програмно і апаратно. На апаратному рівні можна відрегулювати відстань приближення від 2 до 30 см. Програмуючи датчик, можна визначити ефективну відстань спрацьовування. Для цього за допомогою функції Serial.begin() налагоджують сеанс зв'язку з комп'ютером.

Датчик температури DHT11 вимагає підключення додаткової бібліотеки DHT.h. Даний датчик працює лише при підключенні до цифрового (не ШІМ) виходу. Після ініціалізації датчика можна використати дві функції для зчитування температури та вологості: dht.readHumidity(), dht.readTemperature().

Результати тестування. Під час тестування системи було отримано задовільний результат. «Скат-робот» з успіхом відтворював рухи справжнього ската, обпливав перешкоди, реагував на об'єкти, що випромінюють тепло, відповідно до алгоритму заданого в програмі.

Висновки.Було проведено дослідження способу руху підводного об'єкта (ската), а також відтворено даний спосіб за рахунок створення моделі життєздатної системи – автономного підводного апарату – на основі мікроконтролера ATmega328 платформи ArduinoUno. Даний спосіб руху підводних об'єктів можна використовувати для відтворення «ковзаючого» руху або «польоту» підводних апаратів, а також для забезпечення повної чи часткової автономності підводних апаратів. Створення таких апаратів на основі платформи Arduino досить просте, як показує спроба реалізації. Це значною мірою розширює діапазон їх використання.

Список літератури

1. Модель жизнеспособной системы [Електронний ресурс] – Режим доступу до ресурсу: https://ru.wikipedia.org/wiki/Модель_жизнеспособной_системы. (Дата звернення 18.04.2018)

2. Агеев М. Д., Киселев Л. В., Матвиенко Ю. В. Автономные подводные роботы: системы и технологии / отв. ред. Киселев Л. В. — М.: Наука, 2005. — 398 с. — ISBN 5-02-033526-6.

3. Движители кораблей и судов [Електронний ресурс]. – 2010. – Режим доступу до ресурсу: http://korabley.net/news/dvizhiteli_korablej_i_sudov/2010-04-06-527. (Дата звернення 26.04.2018)

4. Steven W. Underwater robotics: science, design & fabrication / W. Steven, B. Harry, J. Vickie. – Hong Kong: MATE Center, 2010. – 758 c.

5. Скаты [Електронний ресурс] // Энциклопедия животных – Режим доступу до ресурсу: http://www.animalsglobe.ru/skatyi/. (Дата звернення 17.04.2018)

6. Симаков Ю. Г. Живые приборы / Ю. Г. Симаков. – Москва: Знание, 1986. – 176 с.

7. Движения и перемещения у рыб [Електронний ресурс] – Режим доступу до ресурсу: http://www.zoofirma.ru/knigi/zoologija-pozvonochnyh/3403-dvizhenija-i-peremeschenija-u-ryb.html. (Дата звернення 18.04.2018)

8. Приспособление рыб к абиотическим факторам среды [Електронний ресурс] – Режим доступу до ресурсу: http://portaleco.ru/ekologija-ryb/sposoby-peredvizhenija-ryb.html. (Дата звернення 18.04.2018)

9. Петров А. В. Энергетические методы увеличения подъемной силы крыла / Альберт Васильевич Петров. – Москва: ФИЗМАТЛИТ, 2011. – 404 с.

10.Understanding Helicopter Automatic Flight Control Systems (AFCS) [Електронний ресурс] // Helicopter Maintenance. – 2014. – Режим доступу до ресурсу: http://www.helicoptermaintenancemagazine.com/article/understanding-helicopter-automatic-flight-control-systems-afcs. (Дата звернення 19.04.2018)

11. Подводные аппараты с движителем колебательного типа [Електронний ресурс] // Vortex Oscillation Technology Ltd. – 2018. – Режим доступу до ресурсу: http://www.vortexosc.com/modules.php?name=Content&pa=showpage&pid=74. (Дата звернення 19.04.2018)

12. Власов В.Н. Механизм плавания рыб и дельфинов, а также полета птиц как вариант движения с отталкиванием от среды. [Електронний ресурс] / Власов В.Н.. – 2012. – Режим доступу до ресурсу: http://vitanar.narod.ru/Delphin/Delphin.htm. (Дата звернення 22.04.2018)

13. Мазейкин Е.М. Конструирование и моделирование технических устройств [Електронний ресурс] / Мазейкин Е.М., Шмелёв В.Е. – Режим доступу до ресурсу: www.tula.net/tgpu/resources/construct/index.html. (Дата звернення 18.04.2018)

14. Сахновский Б.М. Модели судов новых типов [Електронний ресурс] / Сахновский Б.М. // Судостроение. – 1987. – Режим доступу до ресурсу: http://www.shipmodeling.ru/books/NewTypeShips/newtypeships.pdf. (Дата звернення 24.04.2018)

ДОВІДКА ПРО АВТОРІВ

Даніленко Наталія Олегівна – студентка, 4 курс, факультет інформатики та обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Danilenko Nataliia - student of the 4th year Faculty of Informatics and Computer Technology of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: n.danilenko@hotmail.com

Сімоненко Андрій Валерійович - старший викладач факультету інформатики та обчислювальної техніки Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Simonenko Andrii - senior lecturerFaculty of Informatics and Computer Technology of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: svp@comsys.ntu-kpi.kiev.ua

Danilenko Nataliia,
SimonenkoAndrii

MODEL OF UNDERWATER VIABLE SYSTEM BASED ON ATMEGA328 MICROCONTROLLER ARDUINO UNO PLATFORM

Relevance of research topic. Creating a viable system model such as a small, non-intrusive, autonomous submarine (AUV) is a complex and time-consuming process since it is necessary to ensure the adaptability of such a system and take into account the many factors affecting the water environment. To accelerate the design and implementation of ideas, especially for the design of mechanical models with the use of appropriate software, Arduino hardware computing platform is increasingly being used. Its low cost and cross-platformism attract developers. So, with Arduino, you can create standalone devices.

Formulation of the problem. Ensure greater smoothness of movements, speed, and maneuverability of a small-size underwater vehicle. Autopilot development. Ensuring system adaptability.

Analysis of recent research and publications. Over the past two decades, there are more articles devoted to the modeling of underwater viable systems, in particular, due to the emergence of new technical tools for modeling and accessibility of the Arduino hardware platform and similar platforms. However, approaches to the implementation of submarine autonomous devices have not yet been sufficiently studied.

Uninvestigated parts of general matters defining. This article is devoted to the study and analysis of the proposed underwater model of a viable system, in particular, the principle of motion, which can be applied to this model. The research focuses on the study of the mode of movement of underwater inhabitants (skate) and the feasibility of implementing this method in modeling the system, as well as on the creation of appropriate software (in particular, ‘autopilot’ function).

Target setting. The task is to create an small, unattached autonomous underwater vehicle model as a viable system, the principle of which is based on complex fluctuations, and to ensure the adaptability of such a system and to take into account the factors of the influence of the water environment.

The statement of basic materials. The study of the principle of movement of underwater inhabitants (skates) was carried out. Approaches are described for using this method of motion when creating an autonomous underwater vehicle model. The scheme of the combination of electric elements for the creation of a model of "skate" type is proposed. The principle of programming the Arduino platform for this model in various software environments is described. The test results of the model were satisfactory and foreseeable.

Conclusions. As the implementation attempt showed, the creation of a submarine model as a viable system based on the Arduino platform is relatively simple. The results of hardware and software testing of the developed model are presented and their analysis is made.The article presents a method of motion for small, unattached autonomous underwater device, which is realized due to the implementation of complex oscillations. Deals method for combining electrical elements to create a moving device based on the Arduino Uno platform. The analysis of program implementation of the system is made, results of testing are given.

Keywords: underwater viable system, fluctuations, Arduino Uno, Magnus effect, AUV.

УДК 004.4

Владислав Панійван

ПОБУДОВА МОДЕЛІ СИСТЕМИ МАСОВОГО ОБСЛУГОВУВАННЯ ДЛЯ ВИДАННЯ РЕКОМЕНДАЦІЙ ЩОДО ОБРАННЯ ЧЕРГ

BUILDING THE MODEL OF QUEUEING SYSTEM FOR PRODUCING RECOMMENDATIONS OF QUEUE CHOICE

У даній статті описано проект створення системи контролю черг для багатоканальних систем масового обслуговування (СМО) з необмеженим очікуванням. В рамках даної статті буде описано побудову моделі СМО для емуляції роботи гірськолижного курорту та створення системи видачі рекомендацій на основі характеристик та стану створеної моделі.

Ключові слова: система масового обслуговування, черга, видача рекомендацій, модель системи.

This article describes the design of a queue monitoring system for multichannel queuing system (QS) with unlimited expectations. Within the framework of this article will be described the construction of the QS model for emulating the work of the ski resort and the creation of a system for issuing recommendations based on the characteristics and state of the created model.

Key words: queuing system, queue, recommendation system, system model.

Актуальність теми дослідження. Проблема керування чергами гостро стоїть у всіх системах масового обслуговування (СМО). Метою даної статті є опис проекту створення системи контролю людського потоку для багатоканальних СМО. Подібна система може знайти своє використання у транспортних вокзалах, торгівельних центрах і магазинах, курортах і місцях масового відпочинку, тощо, за умови можливості оперативного моніторингу людського потоку та рівнозначності черг.

Постановка проблеми. На сьогоднішній день у більшості СМО відсутня система рекомендацій, або ця система не є достатньо ефективною для вирішення проблеми черг. Створення моделі СМО для емулювання її роботи в реальному часі є першим кроком для створення системи рекомендацій для керування чергами.

Аналіз останніх досліджень і публікацій. Варіанти вирішення проблем керування черг значно відрізняються в залежності від характеру системи масового обслуговування, тому не існує єдиного вирішення цього питання. Замість цього протягом років у багатьох галузях масового обслуговування були сформовані неоднозначні практики щодо покращення ситуації з чергами. Більшість публікацій і досліджень на дану тему розглядають вузьку частину питання та мають теоретично-інформаційний характер.

Виділення недосліджених частин загальної проблеми. Існуючі системи моделювання СМО здебільшого розроблені з ціллю моделювання потоків виробництва і обробки товарів на підприємствах та, у деяких випадках, для моделювання потоків людей, але в останніх відсутній ряд критично важливих

параметрів моделі, таких як місце знаходження клієнта відносно цілей та врахування часу переміщення клієнтів, тому їх неможливо використовувати для створення системи рекомендацій.

Постановка задачі. У даній роботі описується побудова моделі СМО для емуляції роботи гірськолижного курорту та створення системи видачі рекомендацій на основі характеристик та стану створеної моделі. Основною ціллю видачі рекомендації є мінімізація часу очікування між катаннями.

Викладення основного матеріалу. *Побудова моделі.*Систему гірськолижних трас можна представити у вигляді графу. Вершинами графу будуть точки підйому та спуску. Вершини мають якість наявності людей (черг). Ребрами графу є траси підйому та спуску, а їх вагами – час, який потрібний для спуску або підйому. Окрім черг, вершини графу мають ще додаткову характеристику у вигляді структури даних, що відображає прогнозовану кількість людей, які прибудуть до цієї вершини, а також час через який це здійсниться. Нижче наведено тестовий приклад спроектованої системи та результати її роботи. На рисунку 1 представлена схема системи з трьох вершин (зображені у вигляді кіл та пронумеровані від 0 до 2 всередині). Зв'язки між вершинами позначені стрілками, довжини спусків та підйомів вказані поруч.

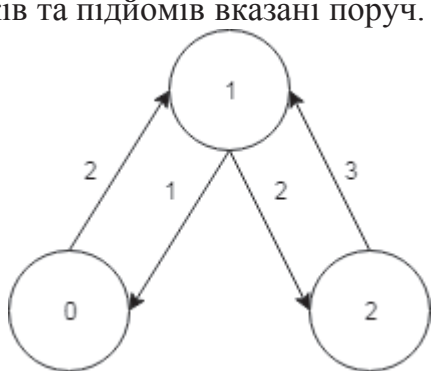


Рис. 1. Схема тестової системи.

У таблиці 1 описано статичні характеристики вершин графу.

Таблиця 1

Статичні дані вершин графу

№ вершиини	Тип вершини	Пропускна здатність чол/хв
0	Підйомник	10
1	Траса	-
2	Підйомник	20

У таблиці 2 описані змінні характеристики графу такі, як: поточна довжина черг та очікувана кількість людей, що прибудуть на підйомник, у деякий час. Для наочності дані наведено для перших декількох хвилин роботи моделі з кроком в одну хвилину.

Таблиця 2

Змінні характеристики вершин протягом роботи системи

Хвилин з початку	№ вершини	Поточна черга	Кількість очікуваних людей
0	0	200	-
	1	0	-
	2	300	-
1	0	190	-
	1	0	10 людей через 1 хвилину 20 людей через 2 хвилину
	2	280	-
2	0	180	-
	1	10	30 людей через 1 хвилину 20 людей через 2 хвилини
	2	260	-
3	0	175	-
	1	30	30 людей через 1 хвилину 20 людей через 2 хвилини
	2	240	5 людей через 1 хвилину
4	0	180	-
	1	30	30 людей через 1 хвилину 20 людей через 2 хвилини
	2	225	15 людей через 1 хвилину

Як можна бачити із таблиці, модель працює коректно, бо кількість людей у системі залишається незмінною. В рамках даного прикладу для симуляції поведінки лижників був використаний закон рівномірного розподілу.

*Видача рекомендацій.*Для лижника, що стоїть на вершині, найкращим шляхом можна вважати той, що мінімізує час очікування після спуску, тобто призводить до найменших затримок між катаннями.

Час очікування можна виразити наступною формулою (1):

$$t = \frac{q}{p} + v, \quad (1)$$

де t – час очікування, q – довжина черги на момент прибуття, p – пропускна здатність підйомника, v – час підйому підйомника. При цьому довжину черги можна обрахувати за формулою (2):

$$q_n = q_0 - p * n + \sum_{i=1}^{n} q'_i, \quad (2)$$

де q_n – довжина черги через n хвилин, q_0 – довжина черги на момент видачі рекомендації, n –тривалість спуску, q'_i – довжина черги яка додасться протягом хвилини i, p – пропускна здатність підйомника.

Приклад роботи системи видачі рекомендацій. Для моделі, що описана у таблицях 1 - 2 та рис. 1, за результатом роботи протягом 4 хвилин система набула наступного вигляду (табл. 3):

Таблиця 3

Стан системи після чотирьох хвилин роботи

№ вершини	Поточна черга	Прибуття людей очікується
0	180	-
1	30	30 людей через 1 хвилину 20 людей через 2 хвилини
2	225	15 людей через 1 хвилину

Для лижника, що знаходиться у вершині 1, постає задача вибору шляху спуску: до вершини 0 чи до вершини 2. У даному випадку, за результатами розрахунків, найкоротший час очікування буде досягнутий при виборі траси до вершини 2 (13.375 хвилин у порівнянні з 19.75 хвилинами), незважаючи на те, що черга у вершині 2 набагато довша. На вибір спуску впливає багато параметрів: час спуску, поточна довжина черги, прогнозована кількість людей за результатами емуляції на момент прибуття, пропускна здатність підйомника та час підйому.

Складання маршруту катання. Наступним кроком у розробці системи може стати покращення системи видачі рекомендацій, а саме: видача рекомендації щодо маршруту катання, який складається з послідовності вершин, що має відвідати лижник (замість видачі рекомендації щодо обрання одного спуску). У цьому випадку, замість мінімізації часу очікування є розумним обрати шлях максимізації корисного часу користувача, тобто відношення катання/очікування:

$$k = \sum_{i=1}^{n} \frac{a_i}{b_i}, \quad (3)$$

де k – коефіцієнт корисності катання, ai – час спуску до вершини i, bi – час очікування на вершині i, n – кількість вершин маршруту.

На практиці подібна рекомендація має бути видана не більше ніж на 40 хвилин вперед, в залежності від точності роботи моделі. Таким чином, у середньому кількість вершин n у формулі 3 не буде перевищувати 3. У цьому випадку для вирішення задачі обрання маршруту можна скористатися алгоритмом повного перебору усіх варіантів катання. Кількість підйомників, до яких лижник може спуститися знаходячись на вершині трас, зазвичай лежить у діапазоні від двох до п'яти. Таким чином, найбільша кількість варіантів маршруту, що має бути розглянута для обрання оптимального, буде дорівнювати 5^3,що не є обчислювально складним. Враховуючи специфіку задачі, подібні обчислення можуть бути виконані достатньо швидко для досягнення прийнятного часу відгуку для кожного окремого клієнта системи.

Точність моделі. Для покращення роботи системи, модель СМО має бути доповнена додатковими параметрами в залежності від її характеру. Так, для гірськолижних курортів, одним із ключових параметрів є час спуску лижника та складність трас. У попередніх прикладах час спуску був фіксованої величиною та виражав середнє значення для більшості людей. Наступним кроком для покращення точності моделі може бути врахування навичок окремого користувача для розрахунку часу спуску при видачі рекомендації. Слід зауважити, що для розрахунку руху людської маси, використання середніх значень є оптимальним підходом, але при видачі рекомендації конкретному клієнту його швидкість спуску є критично важливим, так як від часу спуску буде залежати ситуація на підйомниках, з якою користувач зіткнеться.

Одним із рішень даної задачі є побудова регресивної моделі залежності швидкості лижника від його навичок та складності траси. Останні два параметри можуть бути виражені натуральними числами як якісна оцінка по деякій фіксованій шкалі. Дані швидкості для побудови моделі у цьому випадку мають бути взяті з результатів тестів. Таким чином, формула часу спуску лижника буде виглядати наступним чином:

$$w' = v(c, d) * l, \quad (4)$$

де *w'* – час спуску для окремого користувача, *v(c, d)* – функція лінійної регресії для знаходження швидкості лижника в залежності від його навичок та складності траси, *c* – якісна оцінка навичок лижника, *d* – оцінка складності траси, *l* – довжина спуску.

Висновки. У даній статті розглянуто побудову моделі СМО з метою емуляції її роботи для видачі рекомендацій користувачу щодо обрання черги. На практиці дані в подібній системі мають бути доповнені даними з зовнішніх систем спостереження для створення більш чіткої моделі. Також модель може бути ускладнена додатковими параметрами, такими як: ввід-вивід людей із системи, зміна часу спуску залежно від навичок лижника та інші. Використання подібної системи не обмежено гірськолижними курортами. Система може бути застосована для будь-яких СМО.

Список використаних джерел

1. Брамсон М. Стабільна мережа черг з нестабільною моделлю потоку. Вашингтон, 1999.

2. Сандаранпандіан В. Теорія масового обслуговування. Лондон, 2009.

3. Мануел Лагуна. Пірсон Освіта Індія. Моделювання бізнес процесів, симуляція та дизайн. Нью Делі, 2011. С. 178.

4. Харкол-Балтер М. Моделювання продуктивності та дизайн комп'ютерних систем. Політика складання розкладу черг на основі їх розміру. Лондон, 2012. С. 499.

ДОВІДКА ПРО АВТОРІВ

Панійван Владислав Юрійович – бакалавр, кафедра автоматизованих систем обробки інформації і управління, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Paniivan Vladyslav – bachelor, Department of Computer-Aided Management and Data Processing Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

VLADISLAV PANIIVAN

BUILDING THE MODEL OF QUEUEING SYSTEM FOR PRODUCING RECOMMENDATIONS OF QUEUE CHOICE

Relevance of research topic. The problem of queue management is acute in all queueing systems. The purpose of this article is to describe the project of creating a human-control system for multichannel queueing systems. Examples of queueing systems that can be used by such a system include transport stations, shopping malls and shops, resorts and places of mass recreation, and any other queueing systems.

Formulation of the problem. For today, the problem of queuing management in most queueing systems remains unresolved. The purpose of this article is to describe the project of creating a human-control system for multi-channel queueing systems with unlimited expectations. An example of such systems can be amusement parks or ski resorts.

Analysis of recent research and publications. The options for managing queue problems vary considerably depending on the nature of the queueing system, so there is no single solution to this issue, but instead, over the years, many areas of queueing systems have been shaped by ambiguous practices to improve the queue situation. Most publications and researches on this topic deal with a narrow part of the issue and have a theoretical and informal character.

Selection of unexplored parts of the general problem. Existing QS simulation systems are mainly designed to model the flow of production and processing of goods at enterprises and in some cases to simulate human flows, but the latter lack a number of critical parameters of the model, such as the location of the customer in relation to the goals and taking into account the time of moving customers, so they can not be used to create a system of recommendations.

Setting the task. Within the framework of this article, the construction of the queueing model for emulating the work of the ski resort and the creation of a system for issuing recommendations based on the characteristics and state of the created model will be described. The main goal of the project is to minimize the waiting time between rides.

Presentation of the main material. The article describes the main stages of creating a system model, its structure and work process. There were also described a model for issuing recommendations based on the created queueing model and options for improving the model.

Conclusion. In this article, the construction of the queueing model is considered in order to emulate its work for issuing recommendations to the user, regarding the choice of the queue. In practice data in a similar system should be supplemented with data from external monitoring systems to create a more precise model. Also, the model may be complicated by additional parameters, such as input/output of people from the system, changing the time of descent depending on the skills of the skier,etc. Using such a system is not restricted to ski resorts and can be applied to any queueing system.

UDC 004.8

Володимир Шимкович, Юрій Бердник

РЕАЛІЗАЦІЯ ТА ДОСЛІДЖЕННЯ НЕЙРОННИХ МЕРЕЖ ПРЯМОГО РОЗПОВСЮДЖЕННЯ НА ПЛІС

IMPLEMENTATION AND RESEARCH OF NEURAL NETWORKS OF DIRECT PROPAGATION ON FPGA

В статті наведені результати реалізації нейронних мереж прямого поширення на ПЛІС за розробленим раніше методом реалізації нелінійних функцій активації та алгоритмами реалізації штучних нейронів. Отримані результати дослідження, а саме використаній ресурс ПЛІС, похибка та швидкодія, представлені в таблиці.

Ключові слова: нейронні мережі, функції активації, ПЛІС, VHDL.

Рис.: 1.Табл.: 1. Бібл.: 16.

This paper presents the results of FPGA implementation of directpropagation neuronal networks for the previously developed method of realization of nonlinear activation functions and artificial neurons implementation algorithms. The obtained results of the research, which are presented in the table, are FPGA source, error and speed.

Key words: neural network, activation function, FPGA, VHDL.

Fig.: 1. Tabl.:1. Bibl.: 16.

Neural network control systems constitute a new high-tech branch of control theory and relate to the class of nonlinear dynamic systems. High speed of parallel processing of information coupled with the ability to teach neural networks makes this technology very attractive for creating effective implementation of automated control systems for complex dynamic objects [1, 2]. Neural networks can be used to construct devices for regulation and correction of control systems of reference, adaptive, nominal and inverse-dynamic object model for observation and evaluation of the object parameters, magnitude of the perturbation, search or computation of the optimal influence change program, object identification, prediction of the object state or another operation [3 - 5]. For realization of all neural network capabilities, the neural network controllers of control systems need to be built on chips such as the FPGA [6-9].

FPGA chips allow implementing complex parallel algorithms. This technology involves the implementation of a plurality of blocks on one crystal with intensive information exchange between them. The implementation of data parallel processing on general-purpose processors or digital signal processors requires considerable effort to build multimicroprocessor systems. But a developer does not encounter any rigid architectural constraints in FPGA and can effectively implement parallel algorithms.

Existing approaches to the digital realization of nonlinear functions use different methods of approximation, such as the tabular method, Taylor series expansions, lump-linear approximation, etc. Taylor series expansion requires many multiplications, and therefore unacceptable for FPGA implementation, because the multiplication block occupies a large amount of resources. The tabular method

involves the creation of a global variable, a table of the target function possible values, unpredictable and uncontrolled access to the tables of all neurons in the network, which in turn creates a large time delay. And the creation of a separate local table for each neural network is inappropriate in terms of the use of the FPGA resource. Therefore, the most optimal method for sigmoidal type activation functions implementing is lump-linear approximation of this function.

Issues related to the implementation of one neuron on an FPGA are considered in [6]. The implementation of 3-input neuron with different activation functions on FPGA Xilinx XCV400hq240 has been investigated. The comparison of performance and occupied neuron resources for different implementation variants is given.

In [7] a prototype of a cascade fragment of a neural network with straight sequential bonds was developed. The prototype is implemented on the basis of FPGA Virtex-E XCV400E-pq240. The fragment is intended for parallel input of 8 component input vector, parallel multiplication by the contents of internal blocks of memory of factors, parallel addition of 8 products and parallel activation of the received amounts in each neuron (8th senior bits of sum are given to the address port of memory in which the discrete values of the selected activation function are written). The following result is obtained. The maximum clock speed of the circuit is 90 MHz. Calculation time of the source vector by input is 70 ns.

In work [8] the analysis of the exponential and sigmoid artificial neurons activation functions by means of FPGA (Xilinx Virtex-5 XC5VLX110T) is analyzed. To calculate the values of the sigmoidal function, a table of function values at the reference points is used and a linear interpolation between these points. As a result, the performance and error of the calculations and FPGA resource are presented.

We propose the following approach to the realization of functions of activation of a sigmoid type with the necessary accuracy of its approximation. In the beginning, the activation function is studied for symmetry with respect to the axes. If the condition

$$1-f(x)=f(-x), \tag{1}$$

isfulfilledthen you can consider the function $f(x)$ only for positive arguments, and for negative ones, determine by formula (1), which accelerates the calculation of the function values and reduce the occupied resource of the FPGA.

Further, proceeding from the given accuracy of approximation, we form a piecewise-linear activation function with an appropriate number of intervals, and each of the intervals $(-\infty;x_1),(x_1;x_2);...(x_n;+\infty)$ represent by a separate formula. A general description of a piecewise-linear activation function is written as:

$$f(x)=\begin{cases} k_0x+b_0,\ x<x_1 \\ k_1x+b_1,\ x_1<x<x_2 \\ \cdots \\ k_nx+b_n,\ x_n<x \end{cases}, \tag{2}$$

Next, for each linear function, we calculate the values of the coefficients *k*and *b*and memorize them. Subsequently, for the calculation of the value of the function *f*(*x*)by the x argument from memory, the corresponding values of the coefficients *k*and

*b*are chosen and the value of $f(x)$is calculated according to the corresponding formula. And for the negative values of the argument, the value of the function is calculated by formula 3.

One of the main features of the neural network is the parallel processing of signals. Multilayer neural networks represent a homogeneous computing environment. According to the neuroinformatics terminology, these are universal parallel computing structures, intended for solving various classes of tasks. With the hardware implementation of artificial neural networks on the FPGA, each layer of the network works in parallel with another, which allows using the principle of the conveyor during calculations. Neurons in each layer also work in parallel on the principle of multiprocessor data processing. That is, each artificial neural network is a separate processor and the processing of information in each neuron passes simultaneously.

The results of simulation of neural networks such as used FPGA resource, speed and error are presented in Table 1.

Table 1

Results of neural networks simulation in FPGA

Neural network	Number of synapses	FPGA resource (LUT)	Output ANN value		Error	Speed, ns
			MatLab	FPGA		
1-1-1	3	484	0.56389366	0.546875	0.017019	120.24
1-2-1	5	585	0.62573413	0.609375	0.016359	126.22
1-3-2	10	723	0.68372392	0.671875	0.011849	127.77
1-5-1	11	523	0.78328235	0.765625	0.017657	128.75
2-2-1	8	488	0.63863534	0.625	0.013635	150.48
2-3-1	11	790	0.70144096	0.6875	0.013941	130.58
2-4-2	18	899	0.75747616	0.75	0.007476	133.60
4-4-4	36	1104	0.79223947	0.765625	0.026614	133.53
3-5-1	23	725	0.82549114	0.796875	0.028616	134.05
2-4-1	14	933	0.75747616	0.75	0.007476	133.54
1-3-1	7	723	0.68372392	0.671875	0.011849	127.71
2-1-1	5	602	0.57070375	0.5625	0.008204	126.22
1-4-1	9	829	0.73651211	0.734375	0.002137	127.63
1-6-1	13	1065	0.82373748	0.796875	0.026862	132.26
3-1-1	7	723	0.57708067	0.5625	0.014581	128.71
3-2-1	11	819	0.65058263	0.625	0.025583	134.14
2-8-2	34	1369	0.90702034	0.890625	0.016395	163.24

In the table, the structure of the studied neural networks is represented by three numbers, where the first is the number of neurons in the input layer, the second is the number of neurons in the hidden layer, and the third is the number of neurons in the output layer. The obtained values in LUTs may vary somewhat when the synaptic weights of the neurons change. As a result, an increase in the efficiency of the neural networks implementation on the FPGA is obtained. The increase in the efficiency based on indicators such as the speed and error of the calculations, and the usage of the FPGA resource.

Each neuron is represented as a separate block, as shown in Figure 1, which consists of several parallel processes, and the neural network is a multiprocessor system. The programming language allows you to explicitly specify the signals that start the process. To start the process of calculating, the input signal of this neuron is used.

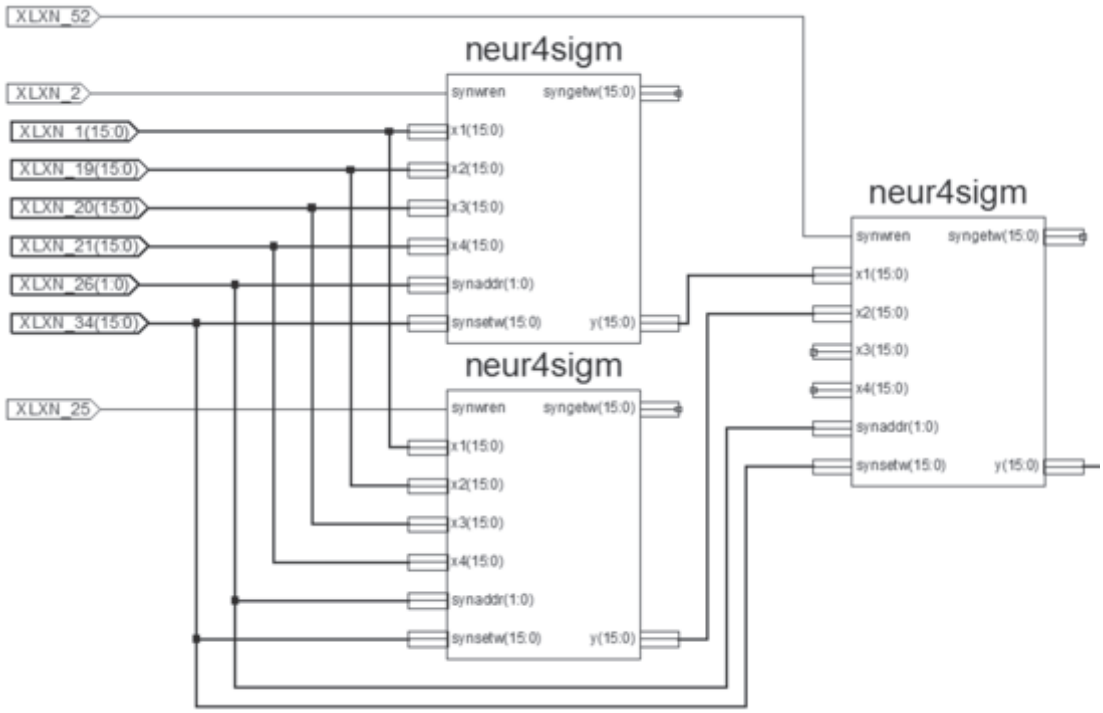


Fig. 1. A fragment of the neural network, implemented on FPGA

To solve the problem of selecting a neural network of minimal structure from a set of possible, [12] describes a technology and software package for the study and evaluation of neural network models (direct and inverse) of multidimensional control objects for their further implementation on the FPGA. Their essence consists in modeling all possible variants of "internal" structures within the "external" structure and determining the structure that provides the lowest mean square error, or the neural network of minimal structure whose mean square error satisfies the condition of the problem.

In order to debug and adapt the neural network models of dynamic objects in real time, it is proposed to use the genetic algorithm and implement it on the FPGA as a parallel search procedure [13].

Realized neural networks can be the basis for further synthesis of neural network components of dynamic control systems based on FPGA. The method of synthesis of hardware components of control systems based on ANN on the example of the implementation of the generalized neural network model of the control object is described in [14-16].

Conclusions. The results of the research of neural networks implemented on FPGAs, according to the developed method, showed their high performance and accuracy with optimal use of computing resource. These neural networks can be the basis for further synthesis of neural network components of dynamic control systems based on FPGA.

References

1. G. Dreyfus. (2005) Neural Networks: Methodology and Applications. SpringerVerlag Berlin Heidelberg. P. 498.

2. A.L. Edelen, S.G. Biedron, B.E. Chase, D. Edstrom, S.V. Milton, P. Stabile. (2016). Neural Networks for Modeling and Control of Particle Accelerators. IEEE Trans.Nucl.Sci. 63 no.2, pp. 878-897

3. M. Lawrynczuk. (2013). Computationally Efficient Model Predictive Control Algorithms: A Neural Network Approach. Springer. Studies in Systems, Decision and Control, vol. 3.

4. Oliver Nelles. (2013). Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models. Springer Science & Business Media, 2013. P. 786.

5. Sigeru Omatu. (2017). Classification of mixed odors using a layered neural network. International Journal of Computing, Volume 16, Issue 1. pp. 41-48.

6. A. Muthuramalingam, S. Himavathi, E. Srinivasan.(2008). Neural network implementation using FPGA: Issues and application. Int. J. Inf. Technol., vol. 4, no. 2, pp. 86-92.

7. Himavathi S., Anitha D. Muthuramalingam A. (2007).Feedforward Neural Network Implementation in FPGA Using Layer Multiplexing for Effective Resource Utilization. IEEE Transactions on Neural Networks. V. 18. № 3. P. 880–888.

8. Ortega-Zamorano F, Jerez J, Juarez G, Perez J, Franco L. (2014). High precision FPGA implementation of neural network activation functions. In: 2014 IEEE symposium on intelligent embedded systems (IES), pp 55–60

9. P. I. Kravets, T. I. Lukina, V. A. Zherebko, V. N. Shimkovich. (2011). Methods of Hardware and Software Realization of Adaptive Neural Network PID Controller on FPGA–Chip. Journal of Automation and Information Sciences. Begell House, New York, USA.Volume 43, Issue: 4, pp. 70-77. DOI: 10.1615/JAutomatInfScien.v43.i4.80

10. Kravets P.I., Shymkovych V.M., Zubenko G.A. (2012).Technology of Hardware and Software Implementation of Artificial Neurons and Artificial Neural Networks by Means of FPGA. Visnyk NTUU ‘KPI’ Informatics, operation and computer systems, , №55, pp. 174-180.

11. Kravets P.I., Shimkovich V.N., Ferens D.A. (2013). Method and algoritms of implementation on PLIS the activation function for artifical neuron chains.Elektronnoe modelirovanie, Vol. 37, no. 4, pp. 63-74.

12. Kravets P.I., Lukina T.I., Shymkovych V.M., Tkach I.I. (2012) Development and research the technology of evaluation neural network models MIMO-objects of control. Visnyk NTUU ‘KPI’ Informatics, operation and computer systems, №57, pp. 144-149.

13. Kravets P.I, Shimkovich V.N. (2013). Method of optimization of weight coefficients of neuron networks by means of genetic algorithm under implementation on programmed logical integral circuits.Elektronnoe modelirovanie, Vol. 35, no. 3, pp. 65-75.

14. Kravets P. I., Shymkovych V.M.,Omelchenko P. (2013) Neuronetwork components of the systems of control of dynamic objects and their hardware-software implementation on FPGA. Visnyk NTUU ‘KPI’ Informatics, operation and computer systems, 2013, №59, pp. 78-85.

15. Kravets P.I., Shymkovych V.M., Fedorchuk V.V., Goy A.A. (2015) Neural controller stability of moving object with the hardware and software realization on FPGA,Visnyk NTUU ‘KPI’Informatics, operation and computer systems, №63, pp. 4–11.

16. P. I. Kravets, V. M. Shymkovych, and V. Samotyy. (2017) Method and technology of synthesis of neural network models of object control with their hardware implementation on FPGA, 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Bucharest, Romania, pp. 947-951. doi: 10.1109/IDAACS.2017.8095226

ДОВІДКА ПРО АВТОРІВ

Шимкович Володимир Миколайович – асистент, кафедра автоматики та управління в технічних системах, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

VolodymyrShymkovych – assistant, Department of Automation and Control in Technical Systems, National Technical University of Ukraine ‘Igor Sikorsky Kyiv Polytechnic Institute’.

E-mail: shymkovych.volodymyr@gmail.com

Бердник Юрій – студент, кафедра автоматики та управління в технічних системах, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Yurii Berdnyk – student, Department of Automation and Control in Technical Systems, National Technical University of Ukraine ‘Igor Sikorsky Kyiv Polytechnic Institute’.

E-mail: berdniky@gmail.com

Володимир Шимкович,
Юрій Бердник

РЕАЛІЗАЦІЯ ТА ДОСЛІДЖЕННЯ НЕЙРОННИХ МЕРЕЖ ПРЯМОГО РОЗПОВСЮДЖЕННЯ НА ПЛІС

Актуальність теми дослідження. Проблема підвищення ефективності адаптивних систем управління на базі нейронних мереж стає більш актуальною в останні дні у зв'язку зі зростаючими вимогами до якості керування складними швидкодіючими об’єктами що працюють в умовах параметричної та інформаційної невизначеності. Дана робота присвячена проблемі підвищення ефективності роботи нейронних мереж при їх апаратній реалізації.

Постановка проблеми. Реалізація та дослідження ефективності роботи нейронних мереж прямого розповсюдження на ПЛІС.

Аналіз останніх досліджень і публікацій. Протягом останніх років з’являється все більше статей присвячених побудові систем управління на основі нейронних мереж та їх апаратній реалізації.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена реалізації та дослідженню нейронних мереж, реалізованих за раніше розробленим методом і алгоритмами. Дослідженнясфокусовано на показниках ефективності роботи нейронних мереж прямого розповсюдження на ПЛІС.

Постановка завдання. Завданням є дослідження ефективності реалізації ШНМ на ПЛІС за розробленим раніше методом реалізації нелінійних функцій активації та алгоритмами реалізації штучних нейронів.

Викладення основного матеріалу. Побудовано моделі нейронних мереж прямого поширення на ПЛІС за розробленим раніше методом реалізації нелінійних функцій активації та алгоритмами реалізації штучних нейронів у середовищі XilinxISEDesignSuite 14.3. Та досліджено їх. В результаті побудовано порівняльну таблицю, де наведено основні характеристики апаратної реалізації ШНМ: швидкодія, похибка обчислень, а також використаний ресурс ПЛІС.

Висновки. Результати досліджень реалізованих на ПЛІС нейронних мереж, за розробленим методом, показали їх високу швидкодію та точність при оптимальному використанні обчислювального ресурсу. Дані нейронні мережі можуть бути базовими для подальшого синтезу нейромережевих компонентів систем управління динамічними об'єктами на основі ПЛІС.

Ключові слова: нейронні мережі, функції активації, ПЛІС, VHDL.

УДК 004.75

Сергій Бугай, Юрій Гордієнко

ОПТИМІЗАЦІЯ АЛГОРИТМУ ВИЯВЛЕННЯ ПАДІННЯ НА ОСНОВІ АКСЕЛЕРОМЕТРА ТА ГІРОСКОПА

OPTIMIZATION OF THE ALGORITHM OF FALL IDENTIFICATIONBASED ON ACCELEROMETER AND GYROSCOPE

Падіння є поширеною і істотною причиною травми у дорослих людей похилого віку, що часто призводить до інвалідності та смерті. Мета цього дослідження полягає у розробці, оптимізації та оцінці ефективності алгоритму виявлення падіння на основі сенсорної системи смарт-годинника Hexiwear, що складається з 3-х осевого акселерометра та гіроскопа. Частина даних була використана для розробки алгоритму виявлення падіння, включаючи дослідження критичних значень датчиків, а решту набору даних було використано для оцінки алгоритму чутливості та специфічності.

Ключові слова: оптимізація, критичні порогові значення, прискорення, кутова швидкість.

Рис:3, табл:3

Fall is a common and significant cause of trauma in older adults, which often leads to disability and death. The purpose of this study is to develop, optimize and evaluate the effectiveness of the fall detection algorithm based on the sensory system of the smart-clock Hexiwear, consisting of a 3-axis accelerometer and a gyroscope. Part of the data was used to develop a fall detection algorithm, including the study of critical sensor values, and the rest of the data set was used to evaluate the sensitivity and specificity algorithm.

Keywords:optimization, critical thresholds, acceleration, angular velocity.

Актуальність теми дослідження. Падіння є найбільш істотними причинами травм в літньому віці. Внаслідок падінь виникає багато порушень, які можуть призвести до смерті через ускладнення, такі як інфекція або пневмонія. Автоматичне виявлення падіння є активною областю досліджень тому, що після падіння люди похилого віку зазвичай не мають змоги самотужки викликати допомогу.

Постановка проблеми. Реалізація ефективного алгоритму виявлення падінь дозволить підвищити ефективність систем для моніторингу стану здоров’я людей похилого віку.

Аналіз існуючих досліджень. Багато досліджень, які використовують акселерометри, в основному використовують зміни величини прискорення для визначення падіння. Коли прискорення перевищує критичну величину, падіння виявляється [1][2]. Однак, орієнтування лише на велике прискорення, може спричинити багато помилкових спрацьовувань, оскільки інші дії, такі як сидіння, біг та стрибки, також можуть призвести до значного зростання прискорення. Інші алгоритми виявлення падіння залежать від виявлення орієнтації тіла після падіння. Проте ці стратегії можуть бути менш ефективними, коли положення людини не є горизонтальним. В даній статті розглянуто використання комбінації датчиків

акселерометра та гіроскопа для надійного виявлення падіння. Акселерометр надає цінну інформацію щодо інерційних змін в організмі через імпульс, гіроскоп надає унікальну інформацію про швидкість обертання тіла під час падіння. Падіння забезпечує як високе прискорення, так і високі кутові швидкості, комбінація яких не спостерігається під час звичайної щоденної діяльності.

Виділення недосліджених частин загальної проблеми. Використання комбінації показників датчиків акселерометра та гіроскопа дозволить підвищити ефективність роботи алгоритму падіння

Постановка завдання. Основною темою даної статті є розробка та оптимізація алгоритму надійного виявлення падіння, одночасно оптимізуючи чутливість та специфічність.

Викладення основного матеріалу. Для зняття показників датчиків використовується смарт-годинник Hexiwear. Алгоритм виявлення базується на простому методі виявлення похибки, що використовує порогові значення акселерометра, знайдені в літературі. Порогове значення гіроскопа було визначено експериментально, використовуючи попередні зібрані дані. Дані збираються та перевіряються в спеціально розробленому мобільному додатку для Hexiwear. Параметри, що використовуються в аналізі, наступні: сума вектору прискорення Acc, що містить як статистичні, так і динамічні показники показники прискорення, обчислюються з зібраних даних використовуючи формулу

$$Acc = \sqrt{(A_x)^2 + \left(A_y\right)^2 + (A_z)^2}, \quad (1)$$

де параметри A_x, A_y, A_z – прискорення в 3х напрямках. Схожим чином, кутова швидкість обчислюється з вибірки даних наступним чином:

$$\omega = \sqrt{(\omega_x)^2 + \left(\omega_y\right)^2 + (\omega_z)^2}, \quad (2)$$

де ω_x, ω_y, ω_z – кутові швидкості в 3х напрямах. Коли об'єкт падає, прискорення швидко змінюється, а кутова швидкість створює різні сигнали вздовж напрямку падіння. Потім критичні пороги в прискоренні та кутовій швидкості використовуються для визначення події падіння. Ці критичні пороги є 2х видів: Lowerfallthreshold (LFT) – локальний мінімум, та Upperfallthreshold (UFT) – локальний максимум. В даній статті для виявлення падіння використовуються LFT_{Acc}та UFT_{Acc}в комбінації з UFT_{gyro}. Алгоритм виявлення падіння зображено на рис.1.

Ефективність запропонованого алгоритму оцінюється шляхом вимірювання чутливості та специфічності, яка зазвичай визначається як

$$Sensivity = \frac{No.TP}{No.TP+No.FN} \ (3), \ Specifity = \frac{No.TN}{No.TN+No.FP} \quad (4)$$

де No.TP (truepositive) – падіння відбувається, пристрій фіксує його, No.FP (falsepositive) – пристрій фіксує падіння, але воно не відбувається, No.TN (truenegative) – нормальний рух, падіння не фіксується, No. FN (falsenegative) – падіння відбувається, пристрій не фіксує його. Труднощі полягають у визначенні оптимальних порогових значень для LFT_{Acc}, UFT_{Acc} та UFT_{gyro}. Неправильно встановлений пороговий вибір може призвести до збільшення випадків помилкових спрацьовувань, виявлених під час щоденних дій, що призводить до зниження специфічності. Крім того, занадто жорсткий відбір порогу може призвести до несправності виявлення падіння, що в цілому призводить до збільшення помилко-

вих негативів та зниження чутливості системи. Багато алгоритмів експериментально визначають порогові значення [3] Щоб оптимізувати пороги, розраховується чутливість та специфічність для всіх можливих значень порогових значень за експериментальними даними. Чутливість зворотньо пов'язана з специфікою в тому сенсі, що чутливість зростає, коли специфічність зменшується в різних порогових межах. Характерна крива характеристик приймача відображає компроміс між чутливістю (осі x) та специфікою (вісь y). Оптимальним порогом є точка, яка може забезпечити мінімальну відстань між точкою (0, 1) і собою [4]. Відстань визначається на рис.4

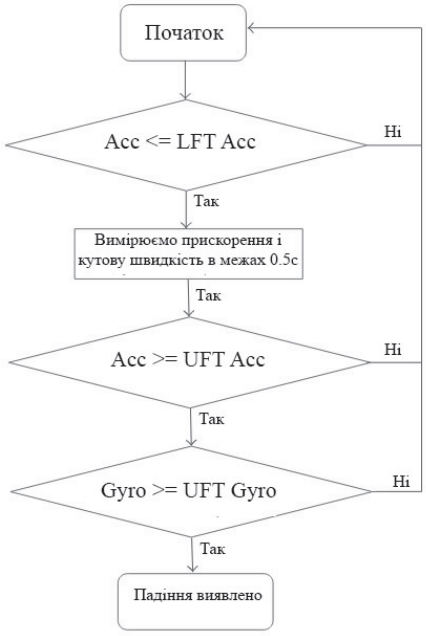


Рис. 1. Алгоритм виявлення падіння

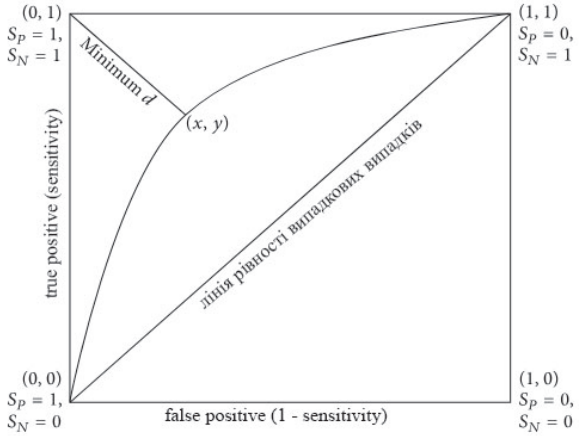


Рис. 2. Крива ПРХ

де, відповідно до чутливості та специфічності:

$$d_T = \sqrt{\left(1 - S_N{}^2\right) + S_p^2} \quad (5)$$

Результати. На рис.3 зображені показники Acc та ω для різних активностей. Також показані порогові значення дляLFT_{Acc}, UFT_{Acc} та UFT_{gyro} щоб визначити падіння. Виявлені події, які задовольняють кожному з компонентів алгоритму, показані внизу. Падіння визначається, коли в межах вікна падіння визначено всі 3 порогові значення.

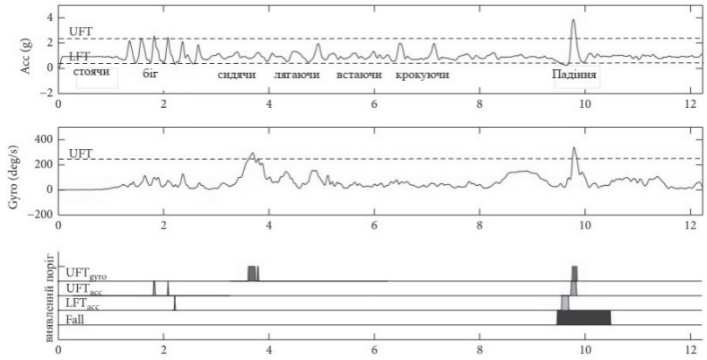


Рис. 3. Показники гіроскопа та акселерометра для різних активностей

З отриманих експериментальних даних, підмножина аналізується для ідентифікації LFT_{Acc}, UFT_{Acc} та UFT_{gyro}. Після цього інші дані використовуються для перевірки для визначення чутливості та специфічності алгоритму. В таблиці 1 представлено резюме критичних порогових значень для LFT_{Acc}, UFT_{Acc} та UFT_{gyro}, включаючи середню ± стандартну похибку, порогове значення в верхній та нижній 90% довірчих межах та 99% верхньої та нижньої межі достовірності.

Таблиця 1

Критичні порогові значення показників акселерометра та гіроскопа

	LFT_{Acc} (g)	**UFT_{Acc} (g)**	**UFT_{gyro} (°/s)**
Середня ± стандартна помилка	0.18 ± 0.06	3.27 ± 0.38	382.1 ± 98.9
Ліміт нижче 90%	0.10	2.77	254.52
Ліміт вище 90%	0.27	3.76	509.61
Ліміт нижче 99%	0.03	2.37	151.69
Ліміт вище 99%	0.33	4.16	612.43

Для підвищення точності системи верхня межа становить від 90% до 99% (0,27 g-0,33 g) для LFT_{Acc}, а нижня межа між 90% та 99% (2,37 g-2,77 g) для UFT_{Acc} та (151,7-254,5 ° / с) для UFT_{gyro} використовуються для алгоритму виявлення падіння. Для оптимізації порогових значень, що використовуються для остаточного алгоритму, запропонований алгоритм запускається на різних рівнях для діапазонів, заданих значенням,, і для комбінованого алгоритму розраховуються значення чутливості та специфічності. Оскільки ціль дослідження полягає у оптимізації алгоритму з цілеспрямованою точністю понад 90% як для чутливості, так і для специфічності, спочатку були обрані три пороги: 0,26 g, 2,77 g та 254,5 °/сек, відповідно аналізам з експериментальних даних. Потім кожен поріг змінюється для перелічених вище діапазонів та розраховується чутливість та специфічність. З кривих ПРХ вибирається

оптимальний поріг, який мінімізує відстань d. Таблиця 2 наводить короткий виклад результатів для чутливості та специфічності падіння без (наприклад, тільки акселерометра) та з включенням показників гіроскопа. У Оскільки пороги встановлені до бажаних рівнів довіри, стратегія оптимізації не реалізується.

Таблиця 2

Результати чутливості та специфічності падіння

	Акселерометр без гіроскопа		**Акселерометр з гіроскопом**	
Рівень довіри	90%	99%	90%	99%
Чутливість	83.56%	98.75%	81.37%	96.55%
Специфіка	96.12%	75.93%	99.38%	89.50%

Результати показують, що підвищення рівня довіри з 90% до 99% підвищує чутливість з 83,56% до 98,75%, як очікувалося, при цьому різко знижується специфічність (від 96,12% до 75,93%) для алгоритму лише з акселерометром. Ці результати схожі з результатами інших [9, 10]. Додавання інформації гіроскопа покращує специфіку шляхом здатності диференціювати падіння від інших активностей. Це досягається з трохи зменшеною чутливістю в порівнянні з методами, де тільки акселерометр. Коли критичні пороги встановлюються та оптимізуються за допомогою ПЧХ-аналізів, можна досягти підвищеної специфічності, одночасно досягти високого рівня чутливості для алгоритмів, що використовують тільки акселерометр та комбіновану стратегію акселерометра та гіроскопа (табл.3). Для комбінованого датчика акселерометра та гіроскопа+оптимізаційна стратегія, чутливість та специфічність для виявлення падіння склали 96,3% та 96,2%, відповідно.

Таблиця 3

Результати чутливості та специфічності з урахуванням стратегії оптимізації

	Акселерометр без гіроскопа	Акселерометр з гіроскопом
Чутливість	97.36%	96.3%
Специфіка	82.72%	96.2%

Висновки. У цьому дослідженні, після схеми оптимізації з використанням кривих ROC та ітеративних аналізів чутливості та специфічності було визначено критичні пороги для, LFT_{Acc}, UFT_{Acc} та UFT_{gyro}, що одночасно призводять до високої чутливості та специфічності виявлення падіння. Ці оптимізовані значення становили 0,30g-0,35g, 2,4g та 240 °/сек, відповідно. Результати демонструють ефективність комбінованого алгоритму акселерометра та гіроскопа, який оптимізований для чутливості та специфічності в порівнянні з традиційними лише акселерометричними стратегіями виявлення падіння. Оптимізована комбінована схема сенсора забезпечує як високу чутливість (96,3%), так і високу специфічність (96,3%), одночасно, набагато вищі, ніж попередні результати, про які йшлося в літературі. Збільшена чутливість досягається шляхом включення інформації про гіроскоп, що забезпечує цінні додаткові відомості про обертання тіла, з метою диференціювання події падіння від інших подій, таких як біг та стрибки. Стратегія оптимізації призводить до

різкого підвищення специфічності, а лише трохи жертвує чутливістю до виявлення падіння [5]. Даний алгоритм був реалізований та протестований при розробці системи для догляду за старими людьми Hexihexi. Код проекту HexiHexi опубліковано на порталі Github для проектів з відкритим вихідним кодом [6]

Список літератури

1. U. Lindemann, A. Hock, M. Stuber, W. Keck, and C. Becker, “Evaluation of a fall detector based on accelerometers: a pilot study,” Medical and Biological Engineering and Computing, vol. 43, no. 5, pp. 548–551, 2005.

2. M. Kangas, A. Konttila, P. Lindgren, I. Winblad, and T. Jämsä, “Comparison of low-complexity fall detection algorithms for body attached accelerometers,” Gait & Posture, vol. 28, no. 2, pp. 285–291, 2008.

3. C. Dinh, D. Tantinger, and M. Truck, “Automatic emergency detection using commercial accelerometers and knowledge,” Computers in Cardiology, vol. 36, pp. 485–488, 2009.

4. R. Kumar and A. Indrayan, “Receiver operating characteristic (ROC) curve for medical researchers,” Indian Pediatrics, vol. 48, no. 4, pp. 277–287, 2011.

5. Quoc T.Huynh,Uyen D. Nguyen, Lucia B. Irazabal, Nazanin Ghassemian,Binh Q. Tran “Optimization of an Accelerometer and Gyroscope-Based Fall Detection Algorithm”, Journal of Sensors, vol.2015, pp. 1-8, 2015.

6. HexiHexionGitHub [Електронний ресурс] 20.02.2018– Режим доступу:URL: https://github.com/RadMushroom/HexiHexi

ДОВІДКА ПРО АВТОРІВ

Гордієнко Юрій Григорович - професор, доктор технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Gordienko Yuri Grygorovych - professor, Doctor of Technical Sciences, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: yuri.gordienko@gmail.com

Бугай Сергій Васильович - студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Buhai Serhii Vasyliovych - student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: sergeybugay94@gmail.com

Serhii Buhai, Yuri Gordienko

OPTIMIZATION OF THE ALGORITHM OF FALL IDENTIFICATIONBASED ON ACCELEROMETER AND GYROSCOPE

Relevance of research topic. Falls are the most significant causes of trauma in the elderly. As a result of falls, there are many disorders that can lead to death due to complications such as infection or pneumonia. Automatic fall detection is an active area of research.

Formulation of the problem. Absence of optimal algorithm for drop detection.

Analysis of existing research. Many studies using accelerometers mainly use changes in the magnitude of acceleration to determine the incidence. When the acceleration exceeds the critical value, the fall manifests itself. However, focusing only on great acceleration can cause a lot of false positives, as other actions, such as sitting, running and jumping, can also lead to a significant increase in acceleration.

Selection of unexplored parts of the general problem. This article is devoted to the development of a fall detection algorithm using a combination of accelerometer and gyroscope.

Task objectives. The challenge is to create an optimal algorithm that, when used, will show high recognition results of the fall.

Presentation of the main material. An analysis of the algorithm's work for various indicators of threshold values was conducted. Approaches are described for optimization of its work. The results showed high recognition accuracy.

Conclusions. Different methods of optimization of the algorithm are analyzed. The results of experiments are presented and the analysis of the following steps is presented.

Keywords:optimization, critical thresholds, acceleration, angular velocity.

УДК 004.75

Сергій Бугай, Юрій Гордієнко

СПОСІБ ІНТЕГРАЦІЇ РОЗПОДІЛЕНИХ ОБЧИСЛЕНЬ НА БАЗІ НОСИМОЇ ЕЛЕКТРОНІКИ

METHOD OF INTEGRATION OF DISTRIBUTED COMPUTING BASED ON WEARABLE ELECTRONICS

На сьогоднішній день існує достатньо різноманітних пропозицій продуктів для догляду за старими людьми, проте більшість з них не може задовольнити всіх потреб даного незахищеного прошарку населення або ж коштують занадто дорого. У статті розглянуто існуючі рішення в області догляду за людьми похилого віку, а також розроблювану бюджетну систему на базі смарт-годинника Hexiwear, мобільного додатку та веб-серверу.

Ключові слова: догляд за похилими людьми, розподілені обчислення, носима електроніка, виявлення падінь.

Рис.:5

Nowadays, there are quite a variety of product offerings for the care of the elderly, but most of them can not meet all the needs of this vulnerable population, or they are too expensive. The article examines existing solutions in the field of care for the elderly, and also as the developed budget system based on the Hexiwear smart watch, mobile application and web server.

Keywords: elderly care, distributed computing, wearable electronics, drop detection

Актуальність теми дослідження. Технології змінили життя кожного з нас і продовжують робити це кожного дня. Люди похилого віку є однією з найбільших незахищених верств населення. Часто вони не можуть повноцінно слідкувати за своїм повсякденним життям та здоров'ям, проте їм можуть допомогти технологічні інновації. В Україні кількість людей похилого віку також зростає, наприклад, на Дніпропетровщині рік тому їх було 12 тисяч, а сьогодні – вже 15 тис. Сучасні технології мають хорошу можливість покращити життя таких людей, дозволяючи їм зберігати свою незалежність і жити повноцінним життям довше. Не менш важливо, завдяки технологіям, полегшиться життя піклувальників. Проте це складна задача, оскільки не всі керівники технологічних продуктів, чий функціонал підходить для задоволення певних потреб старих людей, пристосовують свої продукти до використання ними. Також люди похилого віку іноді можуть мати складнощі з використанням таких продуктів. Певною проблемою є також той факт, що не всі існуючі продукти продаються офіційно в більшості країн світу, а отже для певного кола користувачів не буда доступна підтримка від розробника, що певним чином ускладнює або навіть унеможливлює використання таких продуктів. Є безліч способів для того, щоб зробити життя людей старшого покоління простішим.

Постановка проблеми. Технологія розподілення обчислень з кожним днем стає все більш популярною і все частіше використовується в різних областях. Ця технологія має великий потенціал і тому вона активно розвивається. Тематика

даної статті, що представляє спосіб інтеграції інфрастуктури розподілених обчислень із компонентами носимої електроніки для організації моніторингу фізичного стану людей похилого віку, представляє науковий і практичний інтерес.

Аналіз останніх досліджень і публікацій. На сьогоднішній день існує достатньо різноманітних пропозицій продуктів для догляду за старими людьми, від фітнес-браслетів до систем «розумний будинок». Але більшість з них є занадто дорогими, потенційні користувачі не можуть собі їх дозволити, або ж вони є вузькопрофільними, тобто не здатними забезпечити виконання всіх потреб похилої людини. Система **GrandCare** дозволяє піклувальникам дивитись показники датчиків життєдіяльності, отримувати сповіщення при незвичайних чи небезпечних випадках (критичні показники датчиків, незвична активність тощо), а також виходити на зв'язок з рідними в зручний спосіб. Проте вона має високу ціну та відсутність офіційної підтримки в нашій країні. Система **BeClose** є варіацією системи "розумний будинок", адаптованою під потреби людей похилого віку. За допомогою безпровідних датчиків є можливість відстежувати повсякденне життя літніх людей. Але дана система не відслідковує життєві показники, а також має досить високу ціну. Фітнес-браслети **FitBit** та **JawboneUP** призначені для самостійного відслідковування користувачем таких показників життєдіяльності, як пульс, тиск, сон. Вони мають вузьконаправлений функціонал і підходять лише для профілактичних цілей. Система **MetriaWearableSensor** кріпиться до тіла за допомогою спеціального клею, нешкідливого для шкіри, збирає такі дані, як частота серцевих скорочень, артеріальний тиск і кількість сну, і передає їх по безпровідному з'єднанню, щоб далі інтерпретувати їх за допомогою складних комп'ютерних алгоритмів. Проте вона також підходить лише для профілактичного самостійного використання, оскільки не має можливості повідомити піклувальника про критичні показники датчиків або викликати допомогу. **Tellikin** – простий у використанні домашній комп'ютер, призначений для швидкої та зручної комунікації між людиною похилого віку та її родиною, якщо вони, наприклад, проживають на відстані. Проте даний продукт є занадто вузько направленим і не підходить для використання широким колом користувачів. **Independa** – комплексна система по догляду за старими людьми, що охоплює безпровідні датчики в будинку та датчики на тілі. Але висока ціна робить її недоступною для широкої аудиторії. **TabSafe** - система керування ліками, яка нагадує, розсилає, сповіщає та публікує інформацію про здоров'я, яка доступна з будь-якого пристрою, здатного використовувати Інтернет. Ця система є ефективною для тих людей, які потребують постійного прийому препаратів, але вона не здатна перевіряти стан користувача і попередити про небезпеку для здоров'я. Системи **GreatCallResponder** та **MobileHelp** – персональні системи аварійного відгуку, або також «тривожні кнопки». Вони дають змогу користувачам викликати екстрену допомогу в будь-якому місці. Проте їх функціонал вузько направлений, в нього не входить перевірка життєвих показників користувача.

Виділення недосліджених частин загальної проблеми. Існує безліч застосувань систем відстеження охорони здоров'я, але всі вони мають певні недоліки, такі як висока ціна або обмежений функціонал. Відсутні системи, які б поєднували в собі як доступність для широкого кола користувачів, так і прийнятний функціонал. Дану проблему можна вирішити за допомогою використання open-source носимої електроніки з наявністю пульсометра, акселерометра та гіроскопа.

Постановка завдання. Всі представлені продукти мають певні недоліки, а також дещо обмежений функціонал (наприклад, відсутність можливості сповіщення опікунів при виникненні критичних ситуацій). Тому доцільним є дослідження і реалізація способу інтеграції розподілених обчислень на базі носимої електроніки для догляду за людьми похилого віку, ефективнішого за аналоги, за рахунок використання оптимальнішого алгоритму виявлення падінь.

Викладання основного матеріалу. Розроблювана система складається з наступних компонентів: смарт-годинник Hexiwear, що зчитує показники життєдіяльності за допомогою датчиків, мобільний додаток (призначений для перегляду користувачами та їх опікунами показників датчиків, перегляду графіків на основі статистики даних, відправки та отримання сповіщень в разі критичних показників) та веб-серверу (застосовується для зберігання даних користувачів, слугує проміжним вузлом при відправці сповіщень). Основні переваги даної системи в порівнянні з існуючими аналогами – значно нижча ціна (завдяки використанню бюджетного open-source рішення Hexiwear), здатність виявляти непритомність, падіння та проблеми з пульсом, а також можливість сповіщення опікунів при виникненні критичної ситуації. Hexiwear, розроблений компанією MikroElektronika, - це універсальний смарт-годинник з док-станцією та відкритим кодом, який був успішно профінансований на KickStarter. Він поєднує в собі безліч датчиків, таких як 3D-акселерометр, 3-осьовий цифровий гіроскоп, датчик тиску, датчик пульсу та інше. NXP має ARM-мікроконтроллер Kinetis K64x MCU. Він може підключатися до інших пристроїв за допомогою Bluetooth Low Energy та інформувати користувача через 1,1-дюймовий повнокольоровий OLED-дисплей.

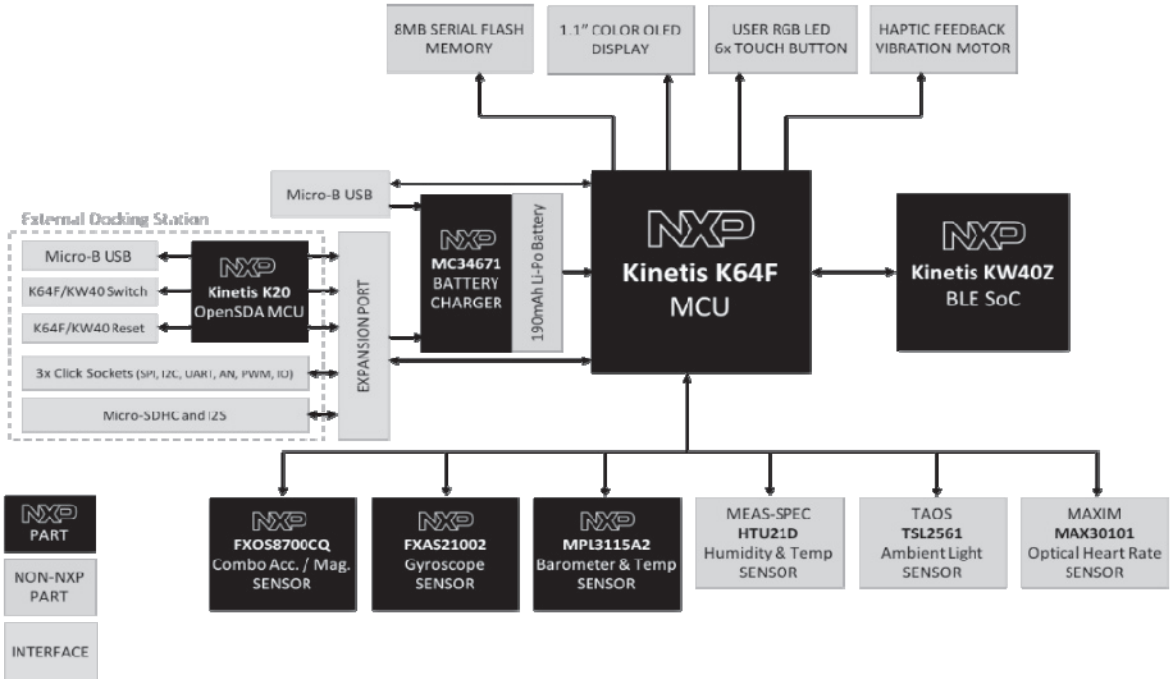


Рис. 1. Апаратна архітектура Hexiwear

За допомогою цих датчиків зчитуватимуться показники пульсу, положення в просторі та прискорення. Система сповіщуватиме опікуна про те, що пацієнт впав або знепритомнів. Спочатку необхідно виявити непритомність. Незважаючи на

наявність деяких комерційних пристроїв та патентів на виявлення падіння, ці пристрої не є задовільними. Основними причинами є високий рівень помилок сигналізації, високі початкові та експлуатаційні витрати пристроїв та їх неергономічний характер. Носимі системи виявлення падіння критикуються переважно тому, що люди можуть забути, знехтувати або не бажати носити їх. Якщо вони працюють від батареї, їх потрібно буде замінити або перезаряджати час від часу [1]. Тому найбільш прийнятний варіант – виявляти падіння за допомогою акселерометра, що дозволить виявляти падіння більш точно завдяки роботі спеціального алгоритму, або за допомогою виявлення змін пульсу. При цьому, непотрібно носити якісь спеціальні пристрої, достатньо зняти показники датчиків Hexiwear. За даними MedicineNet, зміни в серцевому ритмі є найпоширенішою причиною слабкості (непритомності) [2]. Hexiwear вимірює рух користувачів за допомогою акселерометра та гіроскопа. Якщо відбувається раптова зміна, а за нею йде нерухомий стан, то це показує, що користувач втратив свідомість. Значення прискорення всіх трьох напрямків і кутових швидкостей висоти та обертання накладалися, оскільки обертання у вертикальній осі не було значним при падінні. Кінематично, рухи вниз, такі як падіння, демонструють все більше прискорення, тоді як рухи вгору, як стрибки, показують зменшення прискорення у вертикальному напрямку. Таким чином, мінімальне накладене прискорення тривісного прискорення в кожному завданні було використано для виявлення падіння [3].



Рис. 2 Hexiwear

Для перегляду показників датчиків в реальному часі, та моніторингу життєдіяльності було розроблено спеціальний мобільний додаток для ОС Android. Для кожного користувача він відображає показники датчиків в реальному часі, при цьому є можливість перегляду графіків зміни значень пульсометра, акселерометра та гіроскопа.

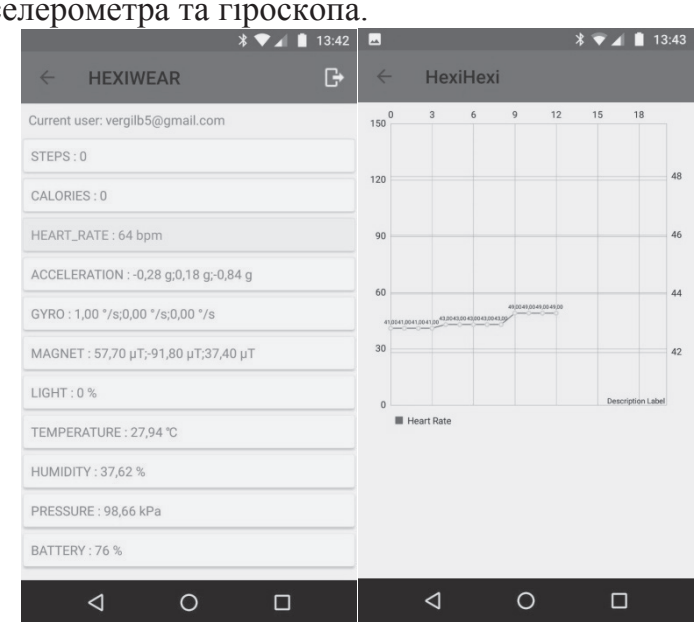


Рис. 3, 4. Інтерфейс програмного додатку

Завдяки реалізації механізму підписки, опікуни автоматично отримуватимуть сповіщення про критичні ситуації тієї людини, за якою вони піклуються.

Таким чином, розроблювана система є досить ефективною завдяки її вигідній ціні, а також завдяки наявності таких функцій, як відстеження падінь та змін в роботі серця та інформуванню опікунів в критичних випадках.Код проекту HexiHexi опубліковано на порталі Github для проектів з відкритим вихідним кодом [4]

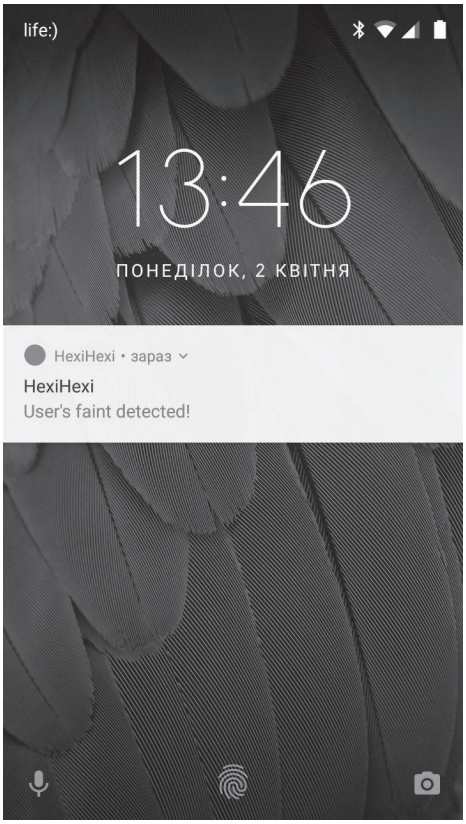


Рис. 5. Приклад сповіщення про небезпеку

Висновки В даній статті було розглянуто існуючі рішення в області технологій для догляду за людьми похилого віку, були виділені їх переваги та недоліки. Також було розглянуту розроблювану систему на базі розподілених обчислень та носимої електроніки та було виділено її переваги над наявними продуктами.

Список використаної літератури

1. Ahmet Turan Özdemir and Billur Barshan. Detecting Falls with Wearable Sensors Using Machine Learning Techniques. ISSN 1424-8220., 18 c.

2. Fainting (Syncope) [Електронний ресурс] 20.02.2018 – Режим доступу:URL: https://www.medicinenet.com/fainting/article.htm#fainting_syncope_facts

3. On Developing a Real-Time Fall Detecting and Protecting System Using Mobile Device [Електронний ресурс] 20.02.2018 – Режим доступу: URL: https://www.jniosh.go.jp/en/publication/doc/ICFPP26.pdf

4. HexiHexionGitHub [Електронний ресурс] 20.02.2018– Режим доступу: URL: https://github.com/RadMushroom/HexiHexi

ДОВІДКА ПРО АВТОРІВ

Гордієнко Юрій Григорович - професор, доктор технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Gordienko Yuri Grygorovych - professor, Doctor of Technical Sciences, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: yuri.gordienko@gmail.com

Бугай Сергій Васильович - студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Buhai Serhii Vasyliovych - student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: sergeybugay94@gmail.com

Serhii Buhai, Yuri Gordienko

METHOD OF INTEGRATION OF DISTRIBUTED COMPUTING BASED ON WEARABLE ELECTRONICS

Relevance of research topic. Nowadays, there are quite a variety of product offerings for the care of the elderly, but most of them can not meet all the needs of this vulnerable population, or they are too expensive. The article examines existing solutions in the field of care for the elderly, and also as the developed budget system based on the Hexiwear smart watch, mobile application and web server.

Formulation of the problem. Developing an effective system for the care of the elderly using distributed computing and wearable electronics

Analysis of existing research. Many products for the care of the elderly are presented on the market, but they either have a high price or can not meet all their needs.

Selection of unexplored parts of the general problem. This article is devoted to the development of a system for elderly care that contains Hexiwear clock, mobile app and web server.

Task objectives. The challenge is to create an effective system for elderly care using modern technologies

Presentation of the main material. Description of the elements of the system, its work, as well as its comparison with analogues

Conclusions. The efficiency of the system in comparison with analogues is shown, possible ways of improvement are highlighted.

Keywords: elderly care, distributed computing, wearable electronics, drop detection

УДК 004.75

М. І. Киричек,
А. М. Волокита

ПОЛІПШЕННЯ ВХОДЖЕННЯ В МЕДИТАТИВНИЙ СТАН ЗА ДОПОМОГОЮ МОБІЛЬНОГО ЗАСТОСУНКУ З ЕЛЕМЕНТАМИ УПРАВЛІННЯ СВІТЛОДІОДНИМИ ЛАМПАМИ.

IMPROVEMENT IN ENTERING INTO A MEDITATIVE STATE BY MEANS OF A MOBILE APPLICATION WITH LED LAMPS MANAGING FEATURE

У даній науковій статі розглядається питання полегшення входження в медитативний стан людини за допомогою управління навколишнім світлом. Після проведеного аналізу існуючих рішень та наукових досліджень у цій сфері на сторінках цією праці можна знайти рішення, яке дозволяє займатися медитативними практиками ефективно та просто. При розробці комп'ютерної системи було використано мобільний додаток iOS та світлодіодні лампочки PhilipsHue.

Ключові слова: інтернет речей, медитація, мобільний застосунок, розумні лампочки, philipshue.

Актуальність теми дослідження. Останнім часом спостерігається стрімкий розвиток у сфері IOT технологій і їх відчутний вплив на якість людського життя. Також на сьогодні близько 20 % людей у світі стикаються з проблемами зі стресом та ментальними захворюваннями [4]. Дієвим методом боротьби є заняття медитацією. Вивчення впливу медитацій при практиці у 7 років показало покращення когнітивних здатностей та покращення задоволенням життям у цілому [5].

Постановка проблеми. На даний момент людям досить складно починати медитативні практики. Причиною цьому є стереотипи зв’язані з релігією, описання практик зв’язані з надмірною езотерикою та достатньо коштовні заняття з учителем. Також на перших етапах заняття є складність сконцентруватися та перевести свій організм у спокійний стан.

Аналіз останніх досліджень і публікацій. Останні роки тема впливу на настрій людини була досить популярна для дослідження. У зв’язку з цим проводились вивчення впливу світла на стан та настрій людини. Прийшли до висновку, що настрій та активність людського тіла залежить типу штучного світла[1]. Так при нормальному циркадному ритму людини, більш яскраве світло, схожу на те, що ми бачимо в день, – буде призводити до активності та навпаки: тьмяне світло, яке по циркадному ритму буде ближчим до вечірніх сутінків, більш розслабляє людину, готуючи до сну [2].

Виділення недосліджених частин загальної проблеми. Попри широкий спектр досліджень медитації та світосприйняття світла на людину, дуже мало рішень які вирішують проблему за допомогою обох чинників одразу. У даній статті представлена гіпотеза, що при входженні в медитативний стан при правильному освітленні, допоможе досягти бажаного ефекту більш ефективно.

Постановка завдання. Ціль статті вирішити проблему складності початку заняття медитативними практиками за допомогою зручного рішення у вигляді мобільного застосунку з елементами управління світла, що знаходиться у середовищі місцеперебування людини.

Викладення основного матеріалу.

Вибір світлодіодних лампочок. При вирішенні питання вибору розумних світлодіодних лампочок, треба врахувати наступні показники: їх дороговизна, популярність на ринку, легкість у використанні та функціональні можливості. При порівнянні були вибрані наступні моделі: PhilipsHue, LIFX's, C by GE.

Таблиця 1

Огляд розумних світлодіодних лампочок

Назва	Ціна	Обхват ринку у світі	Функції
Philips Hue	150$ (50$/шт)	41.23%	Підтримка Wi-Fi та весь спектр кольорів
LIFX's	189$ (47$/шт)	12.3%	Підтрибка Wi-Fi без містка та всього спектру кольорів
C by GE	68$ (17$/шт)	9.2%	Працюють тільки по Bluetooth, обмежений спектр кольорів

Згідно TrustedReviews [6], PhilipsHue показали найкращий результат у надійності підключення та швидкості зміни кольору. Лампи LIFX's мають періодичні проблеми з підключенням, а C by GE складні у керуванні шляхом підключення суто через Bluetooth.

Оскільки ціль цієї статті полгешити заннятя медитацією, дуже важливим фактором є охват ринку у світі. Зважаючи на популярність пристроїв Philips та їх стабільність у використанні, доцільно використовувати саме їх у виконанні завдання.

Розробка технічної системи. Комп'ютерна система має бути легко керована та постійно доступно для користувача. Найзручніший та найзвичніший для користувача девайс – мобільний застосунок на смартфоні. Враховуючи дані StatsCounter, де значиться, що більшість власників PhilipsHue – це мешканці США доцільно буде вибрати операційну систему iOS для розробки мобільного застосунку.

Для максимальної простоти було вибрано інтерфейс, який складається з двох екранів: початковий (Рис. 1) та описуючий процес медитації (Рис. 1). На початку процесу грає запис з аудіогідом по практиці медитації. При цьому задля повного розслаблення застосунок посилає сигнал для переходу світла на жовтий м'який колір з шістнадцятковим кодом #F0C080. По закінченню світло повертає свій попредній відтінок.

Посилання запитів для повільного зміни кольорів відбувається за допомгою HTTP запитів на відкритий APIPhilipsHueBridge.

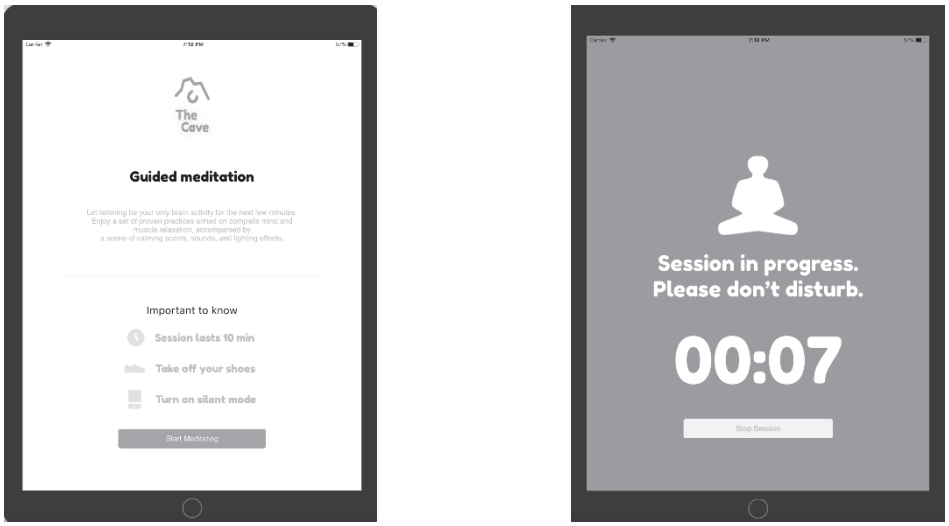


Рис. 1. Інтерфейс програми для медитації

Експеримент з реальними користувачами. При розробці додатку був проведений експеримент для перевірки ефекту спроектованого мобільного застосунку. Для цього була побудована спеціальна тиха кімната розміром 4x5 обладнана лампами PhilipsHue, м'яким простирадлом та подушками, планшетом iPad з операційною системою iOS 11.3. На планшет було попередньо встановлення розроблене рішення з десятихвилинним звукозаписом. Далі на емоційно напруженому заході Hackamackthon 3.0 п'ятьом охочим пропонувалось зменшити стрес та розслабитися у кімнаті. При даному експерименті охочим пропонувалось поміряти пульс до, під час та після досліду.

Таблиця 2

Пульс користувачів за час експерименту

	Пульс до початку експерименту (уд./хв)	**Пульс під час експерименту (уд./хв)**	**Пульс після експерименту (уд./хв)**
Користувач 1	63	52	55
Користувач 2	64	55	59
Користувач 3	82	70	76
Користувач 4	71	63	67
Користувач 5	85	69	74

Експеримент показав, що під час та після практики у користувачів значно знижувався пульс і залишався в середньому на 7 уд./сек меншим, ніж був до цього. Навіть більше, відгуки тестувальників показали, що їх стан значно поліпшився, колір від лампочок був приємним та більша частина мало бажання практикувати медитацію на постійній основі.

Висновки. Стаття показала, як можна покращити наявні рішення за допомогою світосприйняття людини. Розроблене рішення є простим мобільним додатком, що здатне коректно спрямувати людину на заняття медитацією. Проведений експеримент підтвердив ефективність розробленого рішення. Приємне світло дозволяє зменшити активність мозку та більше сконцентруватися на чомусь одному, що і є основним елементом практики. Більше того, маючи таке рішення користувачі мають змогу займатися медитативними практиками індивідуально.

Наступні кроки у цій статті є підключення інших девайсів на виміру активності мозку та пульсу. На основі останніх можливе зміна яскравості світла відносно самопочуття людини.

Список використаної літератури

1. LeGates, Tara A., et al. "Aberrant light directly impairs mood and learning through melanopsin-expressing neurons." *Nature* 491.7425 (2012): 594.

2. Wever, Rütger A., Jan Polášek, and Christina M. Wildgruber. "Bright light affects human circadian rhythms." *Pflügers Archiv* 396.1 (1983): 85-87.

3. Walerczyk, S. T. A. N., CLEP HCLPC, and LC LIGHTING WIZARDS. "Human Centric Lighting." *Architectural SSl* (2012): 20-26.

4. Peterson, Linda Gay, and Lori Pbert. "Effectiveness of a meditation-based stress reduction program in the treatment of anxiety disorders." *Am J Psychiatry* 149.7 (1992): 936-943.

5.Anthony P. Zanesco, Brandon G. King, Katherine A. MacLean, Clifford D. Saron. Cognitive Aging and Long-Term Maintenance of Attentional Improvements Following Meditation Training. *Journal of Cognitive Enhancement*, 2018;

6. http://www.trustedreviews.com [Електронний ресурс] : [Інтернет-портал]. – Електронні дані. – [Philips HueTrustedReviews. – Режим доступу: http://www.trustedreviews.com/reviews/philips-hue (дата звернення 02.05.2018). – Назва з екрану.

ДОВІДКА ПРО АВТОРІВ

Волокита Артем Миколайович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

E-mail: artem.volokita@kpi.ua

Киричек Микита Ігорович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

E-mail: nikita.kirichek@gmail.com

Киричек М. І.,
Волокита А. М.

ПОЛІПШЕННЯ ВХОДЖЕННЯ В МЕДИТАТИВНИЙ СТАН ЗА ДОПОМОГОЮ МОБІЛЬНОГО ЗАСТОСУНКУ З ЕЛЕМЕНТАМИ УПРАВЛІННЯ СВІТЛОДІОДНИМИ ЛАМПАМИ.

Актуальність теми дослідження.

На сьогодні близько 20 % людей у світі стикаються з проблемами зі стресом та ментальними захворюваннями. Дієвим методом боротьби є заняття медитацією. Вивчення впливу медитацій при практиці у 7 років показало покращення когнітивних здатностей та покращення задоволенням життям у цілому.

Постановка проблеми. Нині людям досить складно починати медитативні практики через стереотипи зв'язані з релігією, описання практик зв'язані з надмірною езотерикою та достатньо коштовні заняття з учителем. Також на перших етапах заняття є складність сконцентруватися та перевести свій організм у спокійний стан.

Аналіз останніх досліджень і публікацій. Останні роки проводились вивчення впливу світла на стан та настрій людини. Прийшли до висновку, що настрій та активність людського тіла залежить типу штучного світла і часу, за яким його показують.

Виділення недосліджених частин загальної проблеми. Наразі дуже мало рішень які вирішують проблему за допомогою обох чинників одразу: медитації та світосприйняття світла на людину. У даній статті представлена гіпотеза, що при входженні в медитативний стан при правильному освітленні, допоможе досягти бажаного ефекту більш ефективно.

Постановка завдання. Завданням є вирішення проблеми складності початку занятть медитативними практиками за допомогою зручного рішення у вигляді мобільного застосунку з елементами управління світла, що знаходиться у середовищі місцеперебування людини.

Викладення основного матеріалу. Проведено аналіз чинних засобів керування світлом. Розроблено мобільний застосунок, який керує розумними світлодіодними лампами PhilipsHue та має зручний інтерфейс для початку занять медитацією. Також виконаний експеримент, що показав наскільки це рішення є ефективним.

Висновки. Розроблене рішення є простим мобільним додатком, що здатне коректно спрямувати людину на заняття медитацією. Проведений експеримент підтвердив ефективність розробленого рішення. Приємне світло дозволяє зменшити активність мозку та більше сконцентруватися на чомусь одному, що і є основним елементом практики.

Ключові слова: інтернет речей, медитація, мобільний застосунок, розумні лампочки, philipshue.

УДК 004.8

Ярослав Корнага,
Альона Вітюк

ОЦІНКА ЯКОСТІ СТІЙКОГО ЗАХВАТУ, СПЛАНОВАНОГО НА ОСНОВІ ТРИВИМІРНОЇ РЕКОНСТРУКЦІЇ ЦІЛЬОВОГО ОБ'ЄКТУ ПО ЗОБРАЖЕННЯМ З МОНОКАМЕРИ

QUALITY EVALUATION OF THE STABLE GRASP PLANNED ON THE BASIS OF THE 3D RECONSTRUCTION OF THE OBJECT USING MONOCAMERA IMAGES

У статті розглядається проблема оцінки якості побудови тривимірної моделі об'єкту на основі карти навколишнього середовища мобільного робота, що отримана алгоритмом одночасної локалізації та картографування, для здійснення стійкого захвату робочою кінцівкою робота. Якість отриманих моделей проаналізована на основі успішності захватів у спроектованій системі для обробки візуальних даних з моно камери мобільного робота та планування стійкого захвату цільового об'єкту.

Ключові слова: тривимірна реконструкція, захват об'єктів, одночасна локалізація і картографування

Табл.: 1. Бібл.: 14.

The article deals with the problem of estimating the quality of reconstruction of a 3d object model on the basis of a map of the environment of a mobile robot, obtained by the algorithm of simultaneous localization and mapping, in order to achieve a stable grasp of the robot end effector. The quality of the obtained models is analyzed on the basis of the success of the grasps in the developed system for the processing of visual data from the mobile robot mono camera and the planning of the stable grasp of the target object.

Keywords: 3dreconstruction, object grasping, simultaneous localization and mapping

Table.: 1. Bibl.: 14.

Актуальність теми дослідження. Навігація робота за допомогою візуальної інформації полягає у постійному оновленні свого розташування та карти середовища шляхом здійснення вимірів відстані до оточуючих об'єктів. У техногенних умовах процес тривимірної реконструкції на основі декількох зображень може бути використаний, щоб допомогти роботам визначати їх положення у просторі та побудувати тривимірну карту навколишнього середовища.

Алгоритм одночасної локалізації та картографування (Simultaneous Localization and Mapping -- SLAM) вирішує завдання оцінки позиції мобільного датчика та побудови карти навколишнього середовища в режимі реального часу. Монокулярний SLAM, що обробляє дані з однієї камери, став особливо актуальною темою досліджень останнього десятиліття.

Постановка проблеми. Протягом останніх років, дослідження візуальної одометрії досягли значних результатів по підвищенню точності алгоритму

одночасної локалізації та картографування, надійності та ефективності [1]. Здійснено спроби реалізації різних видів візуальної одометрії, наприклад, пряма та непряма (на ключових точках), щільна чи напівщільна оптимізація та розріджена оптимізація. Дослідженню підлягає застосування алгоритму одночасної локалізації і картографування як методу отримання тривимірної реконструкції об'єкту для захвату робочою кінцівкою мобільного робота.

Аналіз останніх досліджень і публікацій. В останні десятиліття проблема автоматичного захоплення невідомих об'єктів стала відігравати все більшу роль в машинному баченні. Більшість розглянутих алгоритмів використовують певний вид навчання для здійснення стійкого захоплення.

Оскільки захоплення передбачає контакт та сили пальців на поверхню об'єкта, то важливо мати достовірну інформацію як про руку, так і про поверхню цільового об'єкта. Необхідно враховувати обмеження руки робота та конфігурації об'єкта, а також обмеження поставленого завдання [2]. Незважаючи на наявність підходу зі здійснення захоплення лише по двовимірним зображенням [3, 4], більшість технік полягають у обробці тривимірних даних. З огляду на складність поставленої задачі, багато робіт були спрямовані на спрощення тривимірної форми, зокрема плоских представленнях [5] та компланарності, що комбінується з інформацією про колір [6]. Інші підходи передбачають наявність точної моделі об'єкта або такої, що складається з високорівневих примітивів. Прикладом є використання принципу синергії руки [7, 8, 9]. Також існує підхід із застосуванням машинного навчання, зокрема алгоритму Support Vector Machines (SVM) та використанням примітивів форм високого рівня [10]. Алгоритм з навчанням на основі якості двовимірних захоплень, що базується на нейронних мережах та генетичному алгоритму, представлений у роботі "Real-Time Visual Grasp Synthesis Using Genetic Algorithms and Neural Networks" [11].

Задача побудови реконструкції навколишнього середовища роботом є актуальною та активно досліджується. Оскільки камери глибини не є настільки поширеними, як кольорові камери, багато робіт пов'язані зі щільними та напівщільними методами одночасної локалізації та картографування (simultaneous localization and mapping - SLAM) по зображенням з однієї камери [12, 13, 14].

Виділення недосліджених частин загальної проблеми. Незважаючи на таку різноманітність досліджуваних підходів, все ще залишається питання щодо ефективності використання моделі навколишнього середовища для планування стійкого захвату об'єкту як деякої частини цього середовища.

Постановка завдання. Для цілісної оцінки роботи модуля тривимірної реконструкції на основі алгоритму одночасної локалізації та картографування як джерела моделі цільового об'єкту, постає задача аналізу ефективності тривимірної реконструкції з точки зору успішності здійснення захватів моделей цільових об'єктів.

Викладення основного матеріалу.

Механізм обміну даними реалізовано на основі системи ROS. Структура, створена під час дослідження, полягає в контролі емульованого захватного пристрою руки Jaco від одного "Master" комп'ютера. Це зроблено шляхом

створення невеликої підмережі видавців ROS, що зв'язує емулятор руки разом з основною системою. Крім того, система пов’язана з монокамерою, яка закріплена на робочій кінцівці робота.

Поки емулятор обробляє дії руки за допомогою бібліотеки MoveIt!, основний комп'ютер може просто відправляти повідомлення і приймати інформацію про стан через обмін у відповідних темах.

Використання ROS значно спрощує інтеграцію між усіма компонентами, оскільки їх нормальне середовище роботи є неоднорідним. Дійсно, емуляція руки системою MoveIt! контролюється окремим пакетом moveit_setup_assistant, тоді як обробка зображень з камери та побудова тривимірної моделі середовища запрограмовані за допомогою окремих бібліотек C++ (OpenCV та PCL відповідно). Таким чином, ROS дозволяє автоматизувати зв'язок між різними програмами, що працюють одночасно на одній машині та на інших машинах однієї і тієї ж під мережа.

Для оцінки розробленої системи було обрано 4 моделі цільових об’єктів, що отримані з реконструкції сцени у формі хмари точок. Зокрема, було обрано один об’єкт, який є занадто великим для здійснення захвату обраною кінцівкою. Алгоритм має розпізнати, що об’єкт не підходить та максимальний кут розкриття руки є надто малим.

Як бачимо, система побудови тривимірної моделі об’єкта та планування захвату для невідомих об’єктів демонструє хороші результати (табл. 1). Всі тести було проведено на персональному комп’ютері з процесором 3.2GHz на два ядра та середнім часом роботи близько 485.48 сек., планування оптимального захвату потребувало близько 385.67 сек. Для тестової хмари точок значення наведені у таблиці.

Таблиця 1

Тривалість обчислень

Обчислювальний етап	**Час, сек.**
Фільтрація хмари точок	14
Генерація мешу	57
Сегментація	3
Планування захвату	380
Загалом	454

Кожен з обраних об’єктів було протестовано на чотирьох хмарах точок, як показано в табл. В жодному випадку робот не зміг здійснити захват об’єкта номер 2 через невідповідність його розмірів, які є занадто великими для даного захватного пристрою. Таким чином, незадовільний захват був успішно протестований. Для того, щоб захват вважався успішним, рука робота повинна взяти об’єкт, стиснути його та тримати без впускання.

В середньому, робот піднімав невідомий об’єкт у 65% випадків, включно з тестовим об’єктом номер 2, який не є задовільним. Якщо не враховувати 2 об’єкт, рівень успішності становить 87%. Для об’єкту у вигляді коробки-паралелепіпеда алгоритм відпрацював ідеально з рівнем успіху 100%. Хоча, об’єкти такі як чашка та іграшковий автомобіль є більш складними, адже містять кривизну та неоднорідність поверхонь, результати їх захвату є досить успішними з невеликим рівнем помилкових захватів.

Таблиця 2

Успішність спланованих захватів тестових об'єктів

Номер	Об'єкт	Рівень успішності захвату
1	Коробка	100%
2	Кулька	0%
3	Чашка	85%
4	Іграшковий автомобіль	75%
Загалом		65% (87%)

Висновки. У статті проаналізовано ефективність роботи системи тривимірної реконструкції на основі алгоритму одночасної локалізації та картографування як джерела моделі цільового об'єкту з точки зору успішності здійснення захватів моделей цільових об'єктів, що отримані даною системо.

Для цього розроблено систему побудови тривимірної реконструкції середовища на основі даних системи комп'ютерного зору робота для планування стійкого захвату цільового об'єкта. Розробка здійснена на основі таких принципів взаємодії компонентів: засобами системи відбувається обмін даними між підсистемами обробки вхідних зображень з камери, побудови тривимірної реконструкції середовища, отримання тривимірної моделі об'єкту та планування стійкого захвату робочою кінцівкою.

Здійснено загальну оцінку рівня успішності запланованих захватів на наборі тестових об'єктів. При цьому загальний рівень успішності виконання захвату є досить високим – 87%, що свідчить про високу ефективність розглянутої системи.

Перелік посилань

1. Engel, Jakob & Schoeps, Thomas & Cremers, Daniel. (2014). LSD-SLAM: large-scale direct monocular SLAM. Eur. Conf. Comput. Vis.. 8690. 1-16. 10.1007/978-3-319-10605-2_54.

2. Sancho-Bru JL, Mora MC, Leon BE, Perez-Gonzalez A, Iserte JL, Morales A., "Grasp modeling with a biomechanical model of the hand", Comput Methods Biomech Biomed Eng 17(4):297-310, 2014

3. Han, Yuexing, Bing Wang, Hideki Koike and Masanori Idesawa. "Object Recognition with a Limited Database Using Shape Space Theory." Image Processing: Concepts, Methodologies, Tools, and Applications, 2013, pp.181-200

4. Michael Laskey, Jeffrey Mahler, Zoe McCarthy, Florian T. Pokorny, Sachin Patil,"Multi-Armed Bandit Models for 2D Grasp Planning with Uncertainty.", IEEE Conference on Automation Science and Engineering, 2015.

5. Ala R.K., Dong H.K., Shin S.Y., "A 3D-grasp synthesis algorithm to grasp unknown objects based on graspable boundary and convex segments", Inf. Sci. 295, 2015, pp.91-106.

6. Popovic M., Kraft D., Bodenhagen L., Baeski E., Pugeault N., Kragic D., Asfour T., Kruger N., "A strategy for grasping unknown object based on co-planarity and colour information", Robot. Auton. Syst. 58(5), 2010, pp.551-565.

7. M. Bonilla, D. Resasco, M. Gabiccini, and A. Bicchi, "Grasp Planning with Soft Hands using Bounding Box Object Decomposition", IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, pp. 498-503 (2015).

8. M. Bonilla, E. Farnioli, L. Pallottino, A. Bicchi, “Sample-Based Motion Planning for Robot Manipulators with Closed Kinematic Chains”, IEEE ICRA, Seattle, Washington, USA (2015).

9. Tamim Asfour, Júlia Borrás, Christian Mandery, Peter Kaiser, Eren Erdal Aksoy, and Markus Grotz, “On the dualities between grasping and whole-body loco-manipulation tasks“,International Symposium on Robotics Research (ISRR) (2016).

10. R. Pelossof, A. Miller, P. Allen, and T. Jebara, “An SVM Learning Approach to Robotic Grasping,” in IEEE International Conference on Robotics and Automation, 2004, pp. 3512–3518.

11. Huang, Bidan, Sahar El-Khoury, Miao Li, Joanna J. Bryson, and Aude Billard, “Learning a real time grasping strategy” In IEEE international conference on robotics and automation (ICRA), Karlsruhe, 2013, pp.593-600.

12. R. A. Newcombe, S. Lovegrove, and A. J. Davison. Dtam: Dense tracking and mapping in real-time. In IEEE ICCV, pages 2320– 2327, 2011

13. J. Engel, T. Schps, and D. Cremers. LSD-SLAM: LargeScale Direct Monocular SLAM. In European Conference on Computer Vision (ECCV), 2014.

14. R. Mur-Artal, J. M. M. Montiel, and J. D. Tards. Orb-slam: A versatile and accurate monocular slam system. IEEE Trans. Robotics, 31(5):1147–1163, 2015

ДОВІДКА ПРО АВТОРІВ.

Корнага Ярослав Ігорович - кандидат технічних наук, доцент кафедри технічної кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Kornaga Yaroslav - associate Professor of the Department of Technical Cybernetics, PhD of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».

E-mail:slovyan_k@ukr.net

Вітюк Альона Євгеніївна - студентка кафедри Технічної Кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Vitiuk Alona - student of the Department of Technical Cybernetics of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».

E-mail: alyonavityuk@gmail.com

YaroslavKornaga,
AlonaVitiuk

QUALITY EVALUATION OF THE STABLE GRASP PLANNED ON THE BASIS OF THE 3D RECONSTRUCTION OF THE OBJECT USING MONOCAMERA IMAGES

Relevance of research topic. Navigation of the robot using visual information means continually update its location and environment maps by making measurements of distance to surrounding objects. In anthropogenic conditions, the process of 3d reconstruction based on multiple images can be used to help robots determine their position in space and build a three-dimensional environmental map.

The research objective. Research of the application of the simultaneous localization and mapping algorithm as a method for obtaining a 3d reconstruction of the object for a mobile robot grasping.

Actual scientific researches and issues analysis. In recent decades, the problem of automatic grasping of unknown objects began to play an increasingly important role in the machine vision. Most of the considered algorithms use predefined object model for grasping.

Uninvestigated parts of general matters defining. Despite such a variety of research approaches, there is still a question about the effectiveness of using an environmental model to plan a stable grasp of an object as part of this environment.

Target setting. In order to evaluate the effectiveness of the 3d reconstruction module on the basis of the simultaneous localization and mapping algorithm as the source of the model of the target object, the task of analyzing the effectiveness of the 3d reconstruction in terms of the success of the seizure of the target object models is presented.

The statement of basic materials. The system of constructing a 3d reconstruction of the environment based on the data of the system of computer vision work is developed for planning of the sustainable seizure of the target object. The overall assessment of the level of success of the planned grasps in the set of test objects is carried out.

Conclusions. The overall assessment of the level of success of the planned seizures on a set of test objects was performed. At the same time, the overall level of success of the capture is quite high - 87%, which indicates the high efficiency of the considered system.

Key words: 3d reconstruction, object grasping, simultaneous localization and mapping.

УДК 004.7

Олексій Гранюк, Юрій Гордієнко

ПИТАННЯ ПЕРЕДАЧІ ДАНИХ ВІД СЕНСОРІВ НА МОБІЛЬНИЙ ТЕЛЕФОН КОРИСТУВАЧА В КОНТЕКСТІ ІНТЕРНЕТУ РЕЧЕЙ

ISSUE OF TRANSFER OF DATA FROM SENSORS TO A SMARTPHONE IN THE CONTEXT OF INTERNET OF THINGS

У статті розглядається питання передачі даних від сенсорів на мобільний телефон користувача у контексті Інтернету речей. Проаналізовано особливості передачі даних від сенсорів до хмарних сервісів. З'ясовано, що при передачі великого обсягу даних оптимальним рішенням є обробка цих даних локально за допомогою шлюзу.

Ключові слова: Інтернет речей, передача даних, хмарний сервіс, сенсори.

Рис.: 2. Бібл.: 14.

The article deals with the issue of transmitting data from sensors to a user's mobile phone in the context of the Internet of things. The peculiarities of data transmission from sensors to cloud services are analyzed. It is revealed that when transferring a large amount of data, the optimal solution is to process these data locally using the gateway.

Key words: Internet of things, transfer of data, cloud service, smartphone, sensors.

Fig.: 2. Bibl.: 14.

Актуальність теми дослідження. На сьогодні тренд Інтернету речей (*InternetofThings, IoT*) набирає у світі все більшу популярність. Під Інтернетом речей вітчизняні та зарубіжні науковці розуміють концепцію обчислювальної мережі, що складається із взаємозв'язаних фізичних пристроїв, що мають вбудовані сенсори і програмне забезпечення для обміну даними з іншими пристроями [1]. У найзагальнішому вигляді Інтернет речей можна записати у вигляді такої символічної формули: *IoT = Сенсори (давачі) + Дані + Мережі + Послуги* [1; 2].

На основі Інтернету речей можуть бути реалізовані ряд «розумних» додатків (*smart*) [5] у різних сферах діяльності та життя людини: «Розумна планета», «Розумне місто», «Розумний будинок», «Розумна енергетика», «Розумний транспорт», «Розумна медицина», «Розумна освіта» та ін.

При цьому актуальним є питання прийняття рішень для Інтернету речей. Тому, паралельно з популяризацією Інтернету речей розробляються методики мобільних рішень в IoT-проектах та хмарні технології.

Постановка проблеми. Як свідчать доробки провідних вітчизняних та зарубіжних науковців останніх років одним із найпоширеніших варіантів рішення Інтернету речей є використання мобільного телефону для отримання даних від сенсору чи управління підключеннями пристроями [2; 13]. При цьому для передачі та обробки даних використовується хмарний сервіс. Такий підхід має багато переваг, оскільки для передачі даних використовується мережа Інтернет, яка дозволяє отримувати необхідну інформацію навіть коли користувач знаходиться на значній відстані від сенсору.

Аналіз останніх досліджень і публікацій. Питання рішення для Інтернету речей за останнє десятиріччя локалізувалось в центрі дослідницької уваги багатьох вітчизняних і закордонних учених. Зокрема, окремі аспекти

обміну даними у контексті Інтернету речей висвітлені у роботах таких українських дослідників, як: Адріан Наконечний [4], Анжеліка Пархоменко [5], Володимир Скляр [6], Зеновій Верес [4], Олексій Бондарев [1], Сергій Стіренко [11], Юрій Гордієнко [11], Юрій Кошуга [11] та ін. Характеристиці аватарів, сервісів та окремих API-інтерфейсів присвячені доробки таких закордонних учених, як: Бед Партс [8], Генрік Консек [9], Девід Ру [13], Ендрю Фішер [3], Крістофер Еберт [8], Уільям Сталлінгста [10] та ін.

Виділення недосліджених частин загальної проблеми. Залишається недостатньо вивченим питання оптимізації передачі великого обсягу данихвід сенсорів на мобільні телефони користувачів у контексті Інтернету речей.

Постановка завдання. Завданням даної статті є створення моделі оптимізації передачі даних від сенсорів до мобільного телефону користувача з обробкою цих даних у шлюзі у контексті Інтернету речей. Дослідження сфокусовано нааналізі особливостей передачі даних від сенсорів до хмарних сервісів.

Викладення основного матеріалу. Зазвичай, передача усіх даних від сенсору на мобільний пристрій користувача здійснюється напряму в хмару, де вони зберігаються і, за необхідності, обробляються. Але такий підхід до рішення Інтернету речей не є ідеальним у випадку, коли відбувається передача даних від декількох однотипних пристроїв чи групи пристроїв, що належать одному користувачу. В такому разі доцільним буде виконати обробку даних локально за допомогою шлюзу.

Шлюз для Інтернету речей – це пристрій, що дозволяє об'єднувати роботу кількох десятків сенсорів чи інших інтелектуальних пристроїв. Такі шлюзи можуть працювати одночасно у декількох Wi-Fi мережах, підтримують мобільні мережі GPRS, 2G, 3G, LTE та оснащені відносно потужними процесорами і значними об'ємами оперативної пам'яті [7].

Розглянемо питання: навіщо організовувати роботу сенсорів та інтелектуальних пристроїв за допомогою шлюзів? Зауважимо, що сенсори, зазвичай, мають досить слабкі чіпи, малі об'єми пам'яті і обмеження на використання електроенергії [12]. Все це створює проблеми при передачі даних телеметрії в хмару. Сенсори можуть передавати дані з використанням досить простого протоколу UDP (*UserDatagramProtocol* – протокол датаграм користувача, який працює без встановлення з'єднання, забезпечує дуже простийінтерфейсміжмережнимтапрограмним рівнями, але не гарантує доставку повідомлень), а отже може траплятися втрата даних при передачі через глобальну мережу. Водночас цей протокол дуже ефективний при використанні в локальній мережі, як у випадку при підключенні до шлюзу. Сенсори не можуть тривалий час зберігати дані телеметрії, якщо відбувається втрата з'єднання з хмарою. Коли з'єднання втрачено через проблеми у глобальній мережі, то шлюз, за необхідності, зможе зберігати данні до відновлення зв'язку.

Одне з основних переваг використання IoT-шлюзів – це можливість агрегації даних, що надходять від інших пристроїв. Пристрої малого форм-фактору можуть вирішувати багато завдань самостійно, але, якщо навантажити їх надміру, обмеження в обсязі пам'яті і обчислювальної потужності можуть уповільнити весь проект, довести його до майже неробочого стану. Об'єднання безлічі таких пристроїв за допомогою шлюзу дає можливість організувати ефективний збір їхніх даних, дозволяє їм вирішувати виключно ті завдання, на які вони розраховані. При цьому неважливо, пристрої яких саме форм-факторів задіяні в проекті. Шлюз, в будь-якому випадку, здатний значно поліпшити роботу. Більшість шлюзів здатні забезпечити роботу приблизно з десятком

самостійних датчиків. Вони можуть підтримувати відразу декілька робочих Wi-Fi мереж, що дозволяє їм взаємодіяти з датчиками, які спрямовані передавати дані по таких мережах.

Щодо передачі даних через шлюз, то зазначимо, що основні функції шлюзу полягають: у збірці розрізнених показників даних у файли, захисті цих даних та забезпеченні високої швидкості їх передачі через хмарний сервер до користувача, нейтралізуючи можливі затримки.

Для прикладу, розглянемо передачу даних з датчиків температури на мобільний телефон користувача. Передача даних здійснюється згідно схеми, зображеної на рис. 1.

У якості хмарного сервісу розглянемо SamsungArtikCloud. Це спеціальна платформа для розробки рішень для Інтернету речей, що включає в себе як хмарний сервіс так апаратні рішення для IoT. ArtikCloud надає можливість керувати підключеним пристроями через веб-інтерфейс, слідкувати за працездатністю підключених пристроїв. Користувачі також мають можливість переглядати дані, що надходять в хмару в режимі реального часу на графіках, і використовувати вбудований генератор даних. Однак ArtikCloud не надає інструментів для виконання складних обчислень чи машинного навчання. Вважатимемо, що для розробки використовується мова програмування Java. При цьому передачу даних найкраще здійснювати з використанням протоколів CoAP*(ConstrainedApplicationProtocol*) або MQTT (*MessageQueuingTelemetryTransport*).

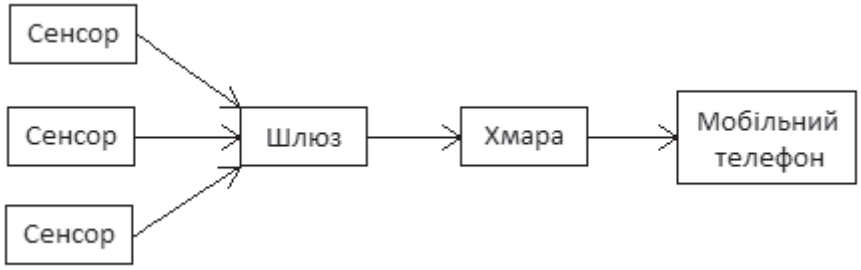


Рис. 1. Схема передачі даних від сенсорів до мобільного телефону користувача

Протокол CoAP використовує модель REST (*RepresentationalStateTransfer*– підхід до архітектури мережевих протоколів, за яким доступ до ресурсів здійснюється через їх представлення) для надсилання повідомлень, що робить його дуже схожим на протокол HTTP (*HyperTextTransferProtocol*– протокол передачі гіпер-текстових документів) з єдиною відмінністю, що CoAP розрахований на міжмашинну взаємодію.

Протокол MQTT працює за принципом «видавець-передплатник». Для надсилання повідомлень використовується програма «брокер». Згідно цієї програми «видавці» надсилають повідомлення брокеру. Дані надсилаються разом із темою. Зауважимо, що розробники можуть самостійно створювати ієрархію тем для роботи великих систем, що може значно спрощувати розробку і підтримку системи. «Передплатники» підключаються до брокера і замовляють тему чи групи тем. Відмітимо, що до переваг протоколу MQTT відносяться: компактність повідомлень, низьке навантаження на мережу, здатність працювати в умовах тимчасових втрат зв'язку. Як досліджено в [14] при втраті пакетів при передачі через мережу на рівні 25% всі повідомлення були вдало надіслані отримувачу при рівні якості обслуговування 1, при цьому, якщо розмір повідомлень складає більше 400 байт, то по протоколу MQTT надсилається на

20-25% менше додаткових даних, ніж при використанні протоколу CoAP при рівні втрати пакетів більше 10%.

Водночас виникає питання, як оптимально реалізувати передачу даних з використанням брокера. Хмарні сервіси надають доступ до власного брокера, розміщеного в хмарі, а це означає, що сенсор буде передавати дані в хмару без попередньої обробки та у великому обсязі. Тому необхідно встановити власний брокер на шлюз. Таким чином дані, які передаються від різних сенсорів, потраплятимуть спочатку до брокера на шлюзі, а далі – передаватимуться програмі обробки даних. Оброблені дані можна надіслати в хмару, використовуючи, наприклад, протокол CoAP, для постійного зберігання даних. Завдяки зручності сучасних хмарних сервісів користувач в будь-який момент зможе отримати поточні дані чи переглянути історію за певний проміжок часу.

Отже, модель оптимізації передачі даних від сенсорів до мобільного телефону користувача з обробкою цих даних на шлюзі має вигляд зображений на рис. 2.

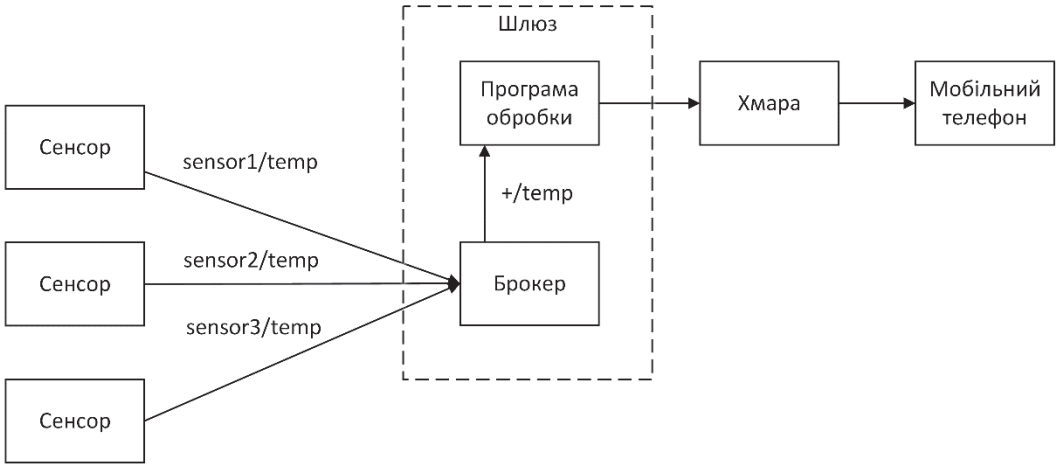


Рис. 2. Модель оптимізації передачі даних від сенсорів до мобільного телефону користувача

Зауважимо, що у представленій моделі поряд із стрілочками написані приклади тем, які можна використати для позначення даних. Символ «+» в назві теми «+/temp» означає, що на цьому рівні може бути вказана будь-яка тема. Такий підхід до вибору назв тем значно спрощує підтримку системи: можна додавати в систему нові сенсори без внесення змін в програму обробки та створювати більш інтелектуальні програми обробки даних.

Висновки. Таким чином,з використанням шлюзу та функціонуванням у ньому власного брокеру можна реалізувати різні варіанти обробки великого обсягу даних. При цьому, з огляду на те, що багато шлюзів оснащені потужними багатоядерними процесорами, навіть досить складні програмні маніпуляції з даними не створюватимуть навантаження на систему передачі даних і не спричинятимуть загрози сповільнення роботи усієї системи. В результаті шлюз виявляється ідеальним пристроєм для підготовки даних малопотужних пристроїв перед відправкою їх, наприклад, в хмару. І як наслідок, запропонована модель оптимізації передачі даних від сенсорів до мобільного телефону користувача з обробкою цих даних на шлюзі, забезпечуватиме оперативне поставлення необхідних даних користувачу. А це, в свою чергу, сприятиме швидкому впровадженню розумного Інтернету речей у різні сфери життєдіяльності людини у нашому суспільстві.

Список використаних джерел

1. Бондарєв Олексій. Що таке інтернет речей і навіщо він потрібен? URL: https://nv.ua/ukr/science/lectures/lektorij-shcho-take-internet-rechej-i-/ navishcho-vin-potriben-1326653.html (дата звернення: 27.04.2018).

2. Интернет вещей (IoT): Пять технологических прорывов 2017 года. URL: https://delo.ua/business/internet-veschej-iot-pjat-tehnologicheskih-proryvov-/ 2017-goda-329735/ (дата звернення: 25.04.2018).

3. Лучшие методики разработки решений для Интернета вещей Эндрю Фишер. URL: https://www.ibm.com/developerworks/ru/library/iot-mobile-/ practices-iot-success/index.html (дата звернення: 25.04.2018).

4. Наконечний А. Й., Верес З. Є. Інтернет речей і сучасні технології // Вісник Національного університету «Львівська політехніка». Авоматика, вимірювання та керування. 2016. № 852. С. 3–9.

5. Пархоменко А. Насколько «умным» должен быть Умный Дом// Карт Бланш. 2017. №1. С. 30–31.

6. Скляр В. Стандарты архитектуры для InternetofThings. URL: https://habr.com/post/307668/ (дата звернення: 27.04.2018).

7. Шлюзы в IoT-разработке: когда, зачем и почему. Блог компанії Intel. URL: https://habrahabr.ru/company/intel/blog/302824/ (дата звернення: 27.04.2018).

8. Ebert Christof, Parts Bad: Looking into the Future, IEEE Software, November/December 2015.

9. Konsek Henryk. Instant Apache ServiceMix How-to. Birmingham : Packt Pub., 2013.

10. Stallings William. Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud. USA: Addison-Wesley Professional, 2015.

11. Stirenko S., Gordienko Y., Shemsedinov T., Alienin O., Kochura Y., Gordienko N., Rojbi A., Benito J., Artetxe González E.User-driven intelligent interface on the basis of multimodal augmented reality and brain-computer interaction for people with functional disabilities // arXiv preprint arXiv:1704.05915, 2017

12. The Elertus Smart Sensor. URL: http://www.elertus.com (дата звернення: 27.04.2018).

13. 7 Big Problems with the Internet of Things David Roe. URL: https://www.cmswire.com/cms/internet-of-things/7-big-problems-with-the-internet-of-things-024571.php (дата звернення: 25.04.2018).

14. Dinesh Thangavel, Alvin Cerdena Valera, Xiaoping Ma, Colin Keng-Yan Tan. Performance evaluation of MQTT and CoAP via a common middleware. 2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing, At Singapore.

ДОВІДКА ПРО АВТОРІВ

Гранюк Олексій Володимирович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Hraniuk Oleksii – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: alekseygranuk@gmail.com

Гордієнко Юрій Григорович – доктор фізико-математичних наук, професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Gordienko Yuri – D.Sc., PhD, professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: yuri.gordienko@gmail.com

OleksiiHraniuk, Yuri Gordienko

ISSUE OF TRANSFER OF DATA FROM SENSORS TO A SMARTPHONE IN THE CONTEXT OF INTERNET OF THINGS

Relevance of research topic. In the most general form, the Internet of things can be written in the form of the following symbolic formula: *IoT = Sensors (sensors) + Data + Networks + Services*. In this case, the question of using a mobile phone for data acquisition or management of connected devices is relevant.

Formulation of the problem. The optimal solution is to process these data locally using the gateway while transmitting a large amount of data.

Analysis of recent research and publications. Over recent years, more and more scientific works have been published by domestic and foreign researchers devoted to the problem of constructing IoT solutions. However, the issue of large-scale data exchange in the context of the Internet of things is still not well understood.

Unexplored parts of the general problem. This article is devoted to the study of the problem of optimizing the transmission of data from sensors to the end devices (smartphones) in the context of the Internet of things.

The research objective. The purpose of this article is to create a model for optimizing the transmission of data from sensors to end devices (smartphones) in the context of the Internet of things using gateways. The research focuses on the analysis of data transfer features from sensors to cloud services.

The statement of basic materials. The advantages of using IoT-gateways in the transmission of data from sensors to end device through the cloud service are analyzed. It was found that data transmitted from sensors might be processed on the gateway before reaching end devices. In case of using MQTT protocol this will lead to additional benefits such as lowered impact on network throughput. The main functions of the gateway are: collection of data, aggregation of data, secure transfer of data to end devices, partial protection of sensors and other connected devices from hacking.

Conclusions. Thus, with the use of the gateway it is possible to implement processing of large amounts of data. The proposed model for transmitting data from sensors to user smartphone with processing these data in the gateway will provide the prompt delivery of the necessary data to the user.

Key words: Internet of things, transfer of data, cloud service, sensors.

УДК 004.056

Марковський О.П.,
Ван Хуаі, Гарасимович Г.В.

ОРГАНІЗАЦІЯ ЗАХИЩЕНОГО ОБЧИСЛЕННЯ МОДУЛЯРНОЇ ЕКСПОНЕНТИ З ВИКОРИСТАННЯМ АДИТИВНОГО МАСКУВАННЯ

ORGANIZE MODULAR EXPONENTIATION SECURE COMPYTATIONBY USING ADDITIVE MASKING

У статті пропонується метод виконання модулярного експоненціювання у хмарних системах із захистом даних та коду експоненти. Запропонований метод дозволяє розділити процедуру модулярного експоненціювання на дві складові: одна, менша частина, виконується користувачем, а інша, більша за обсягом, реалізується на віддалених обчислювальних потужностях. Детально описані математичні принципи пропонованого методу організації захищеного обчислення, викладена відповідна формалізована методика. Наведено числові приклади віддаленого захищеного обчислення модулярної експоненти.

Ключові слова: модулярне експоненціювання, обчислення в хмарі, криптографія, захищені обчислення.

In this paper a method for the performing the calculations required for modular exponentiation using remote or distant computational resources. The proposed method operates by separating the procedure for modular exponentiation in two components. The first component which is computationally simple is performed on the user terminal and the second and computationally complex component, is executed on powerful cloud computational resources. The details of the proposed distributed calculation are presented. Hence, the methodology for the organization of the calculations is analyzed. The calculation is illustrated by means of a simple numerical example.

Key words: modular exponentiation, cloud computing, cryptography, secure computations.

Актуальність теми дослідження. Однією з найбільш знакових подій у розвитку комп’ютерної обробки даних стала поява хмарних технологій. У вирішальній мірі ці прогресивні технології з’явились завдяки динамічному розвитку засобів телекомунікацій та Інтернету. Хмарні технології дозволяють надавати на комерційній основі широкому колу користувачів значні за обсягом обчислювальні потужності, ресурси пам’яті, а також програмне забезпечення. Можливість віддаленого доступу до таких ресурсів істотно підвищує рівень інформатизації в усіх сферах людської діяльності. Разом з тим, розвиток хмарних технологій породжує ряд проблем, одна з яких полягає в потенційній небезпеці надання ресурсів для вирішення завдань, пов'язаних із зломом протоколів криптографічного захисту. Таким чином, впровадження хмарних технологій об'єктивно знижує рівень захищеності існуючих протоколів, що вимагає адекватного підвищення рівня їх захищеності [1].

В основі більшості існуючих протоколів криптографічного захисту інформації лежать незворотні перетворення теорії чисел, базовою математичною операцією яких є модулярне експоненціювання, тобто обчислення $A^E \bmod M$ [1]. Рівень захищеності цих протоколів повністю визначається розрядністю використовуваних в операції модулярного експоненціювання чисел.

На сьогодні в більшості протоколів використовується розрядність 2048 [2], очевидним шляхом підвищення рівня захищеності є збільшення розрядності до 4096 або 8192. Зростання розрядності має наслідком експоненційне зростання обчислювальної складності реалізації мережевих протоколів захисту інформації. Ця проблема стає особливо відчутною для малопотужних термінальних і мобільних пристроїв, що підтримують мережеві протоколи. Виконання на таких пристроях модулярного експоненціювання на розрядностях, більших за 2048, може призвести до порушення часових рамок роботи протоколу.

Виходячи з цього, актуальним стає завдання підвищення продуктивності операції модулярного експоненціювання на малопотужних термінальних і мобільних пристроях, що підтримують протоколи криптографічного захисту.

Постановка проблеми. Один з можливих варіантів вирішення цієї проблеми полягає в залученні для обчислення експоненти значних за обсягом обчислювальних ресурсів хмарних технологій. Необхідною умовою реалізації вказаної можливості є виключення можливості доступу до секретних елементів операції модулярного експоненціювання. Іншими словами, при обчисленні $A^E \bmod M$ необхідно врахувати, що M є частиною відкритого ключа і, відповідно, може передаватися в явному вигляді, компонент E є секретним ключем, компонент A є секретними даними і, відповідно, їхня передача по мережі в явному вигляді має бути виключена. Тобто виникає задача організації віддаленого обчислення модулярної експоненти таким чином, щоб більша за обсягом частина обчислень виконувалася на віддалених обчислювальних потужностях і при цьому не передавалася інформація, яка дає змогу дістати доступ до значень секретних елементів операції – даних та коду експоненти.

Таким чином, наукова задача організації віддаленого обчислення модулярної експоненти, що виключає передачу секретних даних, є актуальною і важливою на сучасному етапі розвитку інформаційних технологій.

Аналіз останніх досліджень і публікацій. Проблемі захисту даних користувача в системах віддаленої обробки присвячено в останні роки значну кількість робіт. Основна проблема полягає в тому, що не існує єдиного підходу до захисту даних у процесі їх обробки. Фактично переважна частина виконаних досліджень вирішує проблему захисту даних тільки для окремих класів обчислювальних задач, наприклад, для лінійної алгебри, обробки зображень і т.п. [2].

Головна проблема, на вирішення якої направлені виконані дотепер дослідження, полягає у використанні спеціальних методів шифрування, які б дозволяли отримувати коректний результат шляхом дешифрування результатів віддаленої обробки зашифрованих даних. З наведеного слідує, що не існує універсальних методів шифрування даних перед їх віддаленою обробкою, які не залежать від операцій обробки сигналів. Це означає, що для кожного виду обробки сигналів слід окремо розробляти метод шифрування та дешифрування.

Фактично опубліковані до теперішнього часу дослідження можна підрозділити на два класи. До першого класу належать роботи, які для окремого класу задач вирішують задачу виключення можливостей доступу до даних, що обробляються віддалено на непідконтрольних обчислювальних потужностях.

Виділення недосліджених частин загальної проблеми. Розроблені дотепер методи захищеного обчислення модулярної експоненти дозволяють близько 70% обчислень реалізувати на віддалених обчислювальних потужностях, а близько 30% - на обчислювальній платформі користувача. Для підвищення ефективності захищених обчислень потрібно підвищити питому вагу обчислень, що виконуються на віддалених комп'ютерних системах.

Постановка завдання. Метою роботи є підвищення ефективності захищеного обчислення модулярної експоненти на віддалених обчислювальних потужностях за рахунок збільшення питомої ваги об'єму обчислень, що виконуються на них і, відповідно, зменшення долі обчислень, які реалізуються користувачем, що дозволяє зменшити час модулярного експоненціювання чисел, розрядність яких значно перевищує розрядність процесора, і тим самим, прискорити обчислювальну реалізацію протоколів криптографічного захисту інформації.

Викладення основного матеріалу. Організація адитивного маскування коду експоненти. В якості найпростішого методу організації захищеного обчислення модулярної експоненти можна використати розкладання коду експоненти на випадкові адитивні складові. Для цього код експоненти E пропонується спочатку трансформувати до вигляду:

$$y = a + b \quad (1)$$

Потім пропонується випадковим чином розділити отриманий код T на h адитивних складових:

$$\delta_0, \delta_1,...,\delta_{h-1}: \; T = \delta_0, + \delta_1 +... +\delta_{h-1}.y = a + b \quad (2)$$

Тоді, за умови, що код A не є секретним (наприклад, у викладених у першому розділі протоколах ідентифікації віддалених користувачів), процедура віддаленого обчислення може бути представлена у вигляді такої послідовності дій:

1. Користувач вибирає випадкове мале ξ та трансформує код експоненти E наступним чином $T=E- \xi$.

2. Користувач випадковим чином розділяє отриманий код T на h адитивних складових $\delta_0, \delta_1,...,\delta_{h-1}: \; T = \delta 0, + \delta 1 +... +\delta h-1$.

3. Користувач посилає в систему код A та набір $\delta_0, \delta_1,...,\delta_{h-1}$.

4. Система обчислює:

$$r_0 = A^{\delta_0} \bmod M, \; r_1 = A^{\delta_1} \bmod M, ..., \; r_{h-1} = A^{\delta_{h-1}} \bmod M$$

5. Одночасно користувач обчислює $D = A\xi \; mod \; M$

6. Користувач отримує від системи обчислені значення $r_0, r_{1, \dots,} r_{h-1}$.

7. Користувач отримує результат у вигляді $R = (r0 \cdot r1, \cdot \dots \cdot, rh-1 \cdot D) \; mod \; M$.

В окремому випадку розкладення коду експоненти на адитивні складові може бути виконано у вигляді суміжних розрядів, що містять у собі, відповідно, $n_0, n_1,...,n_{h-1}$ двійкових розрядів так, що $n_0 + n_1 +...+n_{h-1} = n$. Група δ_0 включає в себе перші n_0 двійкові розряди коду експоненти $\delta_0 = \{e_0, e_1,...,e_{n_0-1}\}$, група δ_1 містить наступні n_1 розрядів: $\delta_1 = \{e_{n_0}, e_{n_0+1},...,e_{n_1-1}\}$; і так далі. Тоді розряди групи δ_0

утворюють число g_0, розряди групи δ_1 утворюють число g_1, розряди групи δ_{h-1} утворюють число g_{h-1}:

$$\forall l \in \{0,1,..., h-1\} : g_l = \sum_{j=0}^{n_l-1} e_{n_0+n_1+...+n_{l-1}+j} \cdot 2^j$$

Відповідно, код експоненти E може бути представлений у вигляді суми адитивних складових [22]:

$$E = g_0 + g_1 \cdot 2^{n_0} + g_2 \cdot 2^{n_0+n_1} + ... + g_{h-1} \cdot 2^{m-n_{h-1}} = \sum_{l=0}^{h-1} g_l \cdot 2^{\sum_{j=0}^{l-1} n_j}$$

Якщо увести позначення $w_0=1$, $w_1 = 2^{n_0}$, $w_2 = 2^{n_0+n_1}$,..., $w_{h-1} = 2^{n_0+n_1+n_2+...+n_{h-2}}$, то експонента E може бути представлена у більш компактному вигляді:

$$E = \sum_{l=0}^{h-1} g_l \cdot w_l$$

Відповідно, операція модулярного експоненціювання $A^E \ mod \ M$ може бути представлена у вигляді добутку:

$$A^E \bmod M = ((A^{g_0} \bmod M)^{w_0} \cdot (A^{g_1} \bmod M)^{w_1} \cdot ... \cdot (A^{g_{h-1}} \bmod M)^{w_{h-1}}) \bmod M =$$

$$= (\prod_{l=0}^{h-1} (A^{g_l} \bmod M)^{w_l} \bmod M) \bmod M$$

Кожен з цих добутків може обчислюватися віддалено, незалежно від іншого, на віддалених обчислювальних потужностях без пересилки їм секретних кодів A та E. Для цього пропонується наступна організація обчислення модулярної експоненти $A^E \ mod \ M$:

1) Користувач обчислює $R_0 = A^{g_0} \bmod M$, $R_1 = A^{g_1} \bmod M$,..., $R_{h-1} = A^{g_{h-1}} \bmod M$.

2) Обчислені значення $R_1,...,R_{h-1}$ і $w_1,...,w_{h-1}$ користувач відсилає в віддалену комп'ютерну систему.

3) Комп'ютерна система паралельно обчислює $D_1 = R_1^{w_1} \bmod M$, $D_2 = R_2^{w_2} \bmod M$ ..., $D_{h-1} = R_{h-1}^{w_{h-1}} \bmod M$, після цього обчислені значення повертаються користувачеві.

4) Користувач обчислює остаточний результат у такому вигляді: $A^E \bmod M = (R_0 \cdot \prod_{i=1}^{h-1} D_i \bmod M) \bmod M$.

Запропонована організація обчислення модулярної експоненти на віддалених комп'ютерних системах може бути проілюстрована таким прикладом для A, E і M малої розрядності.

Нехай модуль M=143, A=96. $E = 93_{10} = 1011101_2$. $A^E \ mod \ M = 96^{93} \bmod 143 = 83$. Нехай 7-розрядний код експоненти E ділиться на три фрагменти: 3-розрядний та два 2-розрядні: n_0= 3, n_1=2, n_2 = 2, δ_0={1,0,1}, δ_1={1,1}, δ_2={1,0}. Тоді g_0=5, g_1=3, g_2=2, w_0=1, $w_1 = 2^3$=8, $w_2 = 2^{3+2}$=32. $R_0 = 96^5 \bmod 143 = 109$, $R_1 = 96^3 \bmod 143 = 138$, $R_2 = 96^2 \bmod 143 = 64$.

Обчислені значення R_1=138, R_2=64, а також значення w_1= 8, w_2= 32, відсилаються у віддалену комп'ютерну систему. Там паралельно обчислюються

$D_1 = 138^8 \bmod 143 = 92$ та $D_2 = 64^{32} \bmod 143 = 92$, які повертаються користувачеві. Останній обчислює $A^E \bmod M = (R_0 \cdot D_1 \cdot D_2) \bmod M = (109 \cdot 92 \cdot 92) \bmod 143 = 83$.

При обчисленні модулярної експоненти $A^E \ mod \ M$ за класичним алгоритмом[4] виконується n кроків, на кожному з яких реалізується модулярне піднесення до квадрату та модулярне множення на A, якщо поточний двійковий розряд коду експоненти дорівнює одиниці. Відповідно, середня кількість N_0 операцій модулярного множення становить $N_0 = 1.5 \cdot n$.

Таким чином, при запропонованій організації обчислення модулярної експоненти кількість операцій модулярного множення, що виконуються користувачем, скорочується приблизно втричі. Це підтверджено експериментальними дослідженнями. Разом з тим, за кодами $R_0, R_1, \ldots, R_{h-1}$, що передаються у віддалену комп'ютерну систему, практично неможливо відновити секретні коди експоненти E та числа A.

Висновки. Доведено, що для маскування секретного коду А в процесі обчислень може бути використане лише мультиплікативне маскування, для якого доцільно використовувати постійне число. Це відкриває можливості для застосування передобчислень, що виконуються лише один раз. Теоретично показано, що для захисту коду експоненти Е найбільш доцільним є використання адитивного маскування, оскільки при застосуванні мультиплікативного маскування на порядки зростає складність обчислень.

Теоретично обґрунтована можливість вибору в якості маски спеціальних кодів, числове значення яких може бути визначено через функцію Ейлера за допомогою узагальнення Ейлера малої теореми Ферма для криптографічних алгоритмів, модуль яких утворюється як добуток двох простих чисел.

Список використаних джерел

1. Boroujerdi N. Cloud Computing: Changing Cogitation about Computing / N. Boroujerdi, S. Nazem. // IJCSI International Journal of Computer Science Issues, - Vol. 9, - Issue 4. -2012.- No 3.- PP. 169-180.

2. Xiaofeng Chen. New Algorithms for Secure Outsourcing of Modular Exponentiations / Xiaofeng Chen, Jin Li, Jianfeng Ma, Qiang Tang, Wenjing Lou // ESORICS 2012, LNCS 7459, - 2012.- PP. 541–556.

3. Can Xiang. Verifiable and Secure Outsourcing Schemes of Modular Exponentiations Using One Untrusted Cloud Server and Their Application. / Can Xiang // IACR Cryptology ePrint Archive 2014: PP.500.- https://eprint.iacr.org/2014/500.pdf

4. Костенко Ю. В. Метод защищенного модулярного экспоненцирования на удаленных компьютерных системах. / Ю.В. Костенко, А.П. Марковский, О.В. Русанова. // Вісник Національного технічного університету України "КПІ" Інформатика, управління та обчислювальна техніка. К.: ТОО „ВЕК+".- № 64.- 2016.- С. 51-54.

ДОВІДКА ПРО АВТОРІВ

Марковський Олександр Петрович – доцент, кандидат технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Markovskyi Oleksandr – associate professor, candidate of engineering sciences, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: markovskyy@i.ua

O. P. Markovskyi ,
Van Huai,
H. V. Harasymovych

ORGANIZE MODULAR EXPONENTIATION SECURE COMPYTATIONBY USING ADDITIVE MASKING

Target setting. The natural way to restore cryptographic resilience of information security systems of this class in similar conditions is to increase the bit depth n of the numbers utilized. However, as mentioned above, increasing of bit depth significantly slows down calculations, related to information security functions. This situation may be overcome by using computing resources of cloud systems for modular exponentiation, in such a manner that that when calculating the A^E mod M, the secret exponent E code and the processed number A are not disclosed.

The principle difficulty in tackling this problem lies in the fact that there exists no general approach for guaranteeing data protection during processing. Methods for data protection depend heavily on the nature of the operations employed during processing. In other words, data security technologies require a homomorphic encryption procedure for the remote data encryption process [4]. For this reason, a number of different solutions for the secure remote implementation of solutions to problems of different categories, such as linear algebra or image processing have been proposed.

Consequently, amongst the targets of algorithms for the remote implementation of modular exponentiation should be the direct coordination of parallel solutions to this problem in systems with multiple processors.

The purpose of the present research is to develop a method for the secure implementation of the modular exponentiation operation in cloud systems, with the capability for parallelism.

As a result of the research proposed, a simple technological approach was proposed for exploiting cloud computing resources in order to increase the speed of calculations of the fundamental operation that is required for most data security protocols. This operation is modular exponentiation. The acceleration of the calculation is achieved by means of the transfer of a significant proportion of the complexity involved to computing resources available via the use of cloud technologies.

The method proposed involves the execution of part of the calculations on user equipment and another part in remote computational resources. During the cloud processing, the secrecy of both the data and the user key is preserved.

Itwasshownbymeansofboththeoreticalandpracticalcalculationsthatthevolumeofth eoperationsrequired to be performed by the users is reduced by approximately a factor of three times. This reduction results in a significant increase in the speed of implementation of cryptographic mechanisms, by use of the significant processing resources made available by the contemporary cloud resources.

The proposed method is oriented to user applications that operate on portable computing terminals with small processing power and connected to the internet.

Key words: modular exponentiation, cloud computing, cryptography, secure computations.

УДК 004.75

Гарасимович Г. В.,
Куц В. Ю.

МЕТОД ЗАХИСТУ ОПЕРАНДІВ МОДУЛЯРНОГО ЕКСПОНЕНЦІЮВАННЯ ВІД ЇХ РЕКОНСТРУКЦІЇ АНАЛІЗОМ ДИНАМІКИ СПОЖИВАННЯ ПОТУЖНОСТІ

METHOD FOR PROTECTION MODULAR EXPONENTIATION OPERANDS RECONSTRATION BY POWER ANALYSIS

Метою представлених в статті досліджень є аналіз небезпеки реконструкції операндів модулярного експоненціювання аналізом динаміки споживання потужності і розробка способів протидії. Показано, що ступінь модулярного експоненціювання, що є ключем алгоритмів RSA, El-Gamal, DSA може бути реконструйована тимчасовим аналізом споживаної потужності. Для протидії розроблений спеціальний алгоритм модулярного експоненціювання. Алгоритм не має умовних операторів і включає неправдиві операції, що утрудняють тимчасової аналіз потужності. Застосування запропонованого підходу збільшує на 25% час виконання модулярного експоненціювання.

Ключові слова: модулярне експоненціювання, аналіз динаміки споживання потужності, смарт-карти, криптографічні алгоритми, протоколи захисту даних, термінальні обчислювальні пристрої.

Бібл.: 6

The goal of presented by article research is to point out the potential vulnerabilities of modular exponentiation operands reconstruction by power dynamic analysis and to elaborate countermeasures. It has been shown that exponent of modular exponentiation which is secret key of RSA, El-Gamal and DSA can be reconstruction by timing power analysis. For countermeasure the special algorithm for modular exponentiation has been worked out. Proposed algorithm does not conditional operators use and include the false operators which inhibit to timing power analysis. It has been shown that implementation of proposed approach demand about 25% more time for modular exponentiation.

Key words: power analysis, smart cards, cryptographic algorithms, data security protocols, terminal computer devises.

Bibl.: 6

Актуальність теми дослідження. Динамічне поглиблення інформаційної інтеграції на основі комп'ютерних мереж відіграє домінуючу роль на сучасному етапі розвитку більшості сфер людської діяльності. Важливою умовою розширення інформаційної інтеграції є рішення проблеми ефективного захисту даних і розмежування прав доступу до інформації в комп'ютерних системах і мережах. Розширення використання комп'ютерних і мережевих технологій в сфері фінансової діяльності, для управління складними технічними системами, пов'язаними з техногенним ризиком вимагає забезпечення високої надійності захисту інформації в комп'ютерних системах управління і мережах. Таким чином, існує об'єктивна необхідність підвищення ефективності засобів захисту інформації з точки зору розширення сфер використання інтегрованих систем. З іншого боку, об'єктивно існує ряд факторів, що знижують надійність засобів захисту даних. Таким чином, наукова задача підвищення ефективності засобів

захисту інформації є важливою і актуальною для сучасного етапу розвитку інформаційних технологій.

Постановка проблеми. В останні роки, однією з найбільш потенційно небезпечних загроз інформаційній безпеці в системах комп'ютерного управління є несанкціонований доступ до даних за допомогою вимірювання і аналізу динаміки зміни параметрів роботи обчислювального пристрою під час реалізації ним програм [1]. Особливо актуальною є проблема захисту даних від їх реконструкції за допомогою вимірювання і аналізу динаміки зміни споживаної потужності. В сучасних комп'ютерних мережах значну частину термінальних пристроїв складають мікроконтролери, для яких вимір споживаної потужності і зіставлення її з командами, що виконуються, здійснюється відносно просто [1]. Разом з тим, такі термінальні пристрої підтримують мережеві протоколи захисту інформації, при реалізації яких використовуються секретні дані - ключі і паролі.

Сучасні технології дозволяють виконувати статистичну реконструкцію цих секретних даних на основі аналізу динаміки споживаної потужності [2]. Аналіз таких технологій свідчить про те, що вони найбільш ефективні для доступу до постійних компонентів систем захисту інформації - ключам і паролів. Це особливо небезпечно для механізмів захисту даних, ключі яких рідко змінюються. До таких механізмів, зокрема, відносяться алгоритми з відкритим ключем, найбільш поширеними з яких є RSA, El-Gamal, DSA. Зміна ключів при кожному сеансі роботи цього алгоритму принципово порушує концепцію відкритості ключів. Тому ключі згаданих алгоритмів найбільш уразливі для несанкціонованого доступу до них шляхом аналізу динаміки споживання потужності. Разом з тим, ці алгоритми лежать в основі мережевих протоколів інформаційної безпеки, тому втрата контролю над їх ключами загрожує серйозними порушеннями в роботі систем захисту інформації в мережах.

Таким чином, проблема захисту від реконструкції аналізом динаміки споживання потужності ключів алгоритмів, в основі яких лежить модулярне експоненціювання є важливою і актуальною на сучасному етапі розвитку технології захисту даних в мережах.

Аналіз останніх досліджень і публікацій. Базовою обчислювальною операцією при програмній реалізації великої кількості алгоритмів захисту інформації з відкритим ключем, в тому числі RSA, є модулярне експоненціювання $X^E \bmod M$, при $X, E < M$. Кількість n двійкових розрядів чисел X, E і M значно перевищує розрядність k процесора. Кожне з чисел X, E і M, можна представити у вигляді s слів, що складаються з k двійкових розрядів $(s=n/k)$, кожне з яких лежить в інтервалі від 0 до 2^k-1: $X=\{x_{s-1},\ldots,x_1,x_0\}$, $E=\{e_{s-1},\ldots,e_1,e_0\}$ і $M=\{m_{s-1},\ldots,m_1,m_0\}$, $\forall j\in\{0,...,s-1\}$, $0\le x_j,e_j,m_j\le 2^k-1$. Для вказання двійкових розрядів слів X, E і M використовуються позначення з двома індексами, наприклад $e_0=\{e_{0,k-1},e_{0,k-2},\ldots,e_{0,1},e_{0,0}\}$; $\forall l\in\{0,...,k-1\}$: $x_{j,l}$, $e_{j,l}$, $m_{j,l}\in\{0,1\}$:

$$X=\sum_{j=0}^{s-1} x_j\cdot 2^{j\cdot k}=\sum_{j=0}^{s-1}\sum_{l=0}^{k-1} x_{j,l}\cdot 2^{j\cdot(k-1)+l}, E=\sum_{j=0}^{s-1} e_j\cdot 2^{j\cdot k}=\sum_{j=0}^{s-1}\sum_{l=0}^{k-1} e_{j,l}\cdot 2^{j\cdot(k-1)+l},$$

$$M=\sum_{j=0}^{s-1} m_j\cdot 2^{j\cdot k}=\sum_{j=0}^{s-1}\sum_{l=0}^{k-1} m_{j,l}\cdot 2^{j\cdot(k-1)+l}$$

Процес модулярного експоненціювання зводиться до послідовного виконання $log_2E = n$ циклів, в кожному з яких здійснюється операція зведення в квадрат отриманого на попередньому циклі результату і, додатково, в залежності від поточного біта ступеня E, виконується операція множення. Виходячи

з порядку, в якому аналізуються розряди ступеня E, існує два різновиди модулярного експоненціювання: справа - наліво і зліва - направо.

При використанні алгоритмів, заснованих на модулярному експоненціюванні, в якості секретного ключа виступає код експоненти E, який практично не змінюється. Відповідно, основним завданням ставлячи під загрозу безпеку є отримання коду E. В якості побічної задачі виступає задача доступу до значення коду X або результату - $A=X^E \bmod M$ при невідомому, але постійному E [2]

Виділення недосліджених частин загальної проблеми. Як уже зазначалося, значну частину термінальних пристроїв мереж складають вбудовані мікроконтролери та мікропроцесори, що використовуються для збору даних або управління технологічним обладнанням. До 2005 року частина таких термінальних пристроїв становила до 60% [2], причому має місце тенденція її зростання з розвитком комп'ютерних та мережевих технологій, а також з розширенням сфер їх застосування.

Цей клас термінальних пристроїв підтримує протоколи захисту інформації в мережах, в процесі реалізації яких використовує постійні секретні параметри - паролі, ключі криптографічних алгоритмів. Для мікроконтролерів і мікропроцесорів щодо просто виміряти під час виконання програм динаміку споживаної потужності і зіставити її з виконуваними командами. Всі процесорні пристрої обчислювальної техніки побудовані на логічних вентилях, основною частиною яких є транзистори. Струми, що протікають через транзистори, залежать від їх робочих точок і змінюються стрибкоподібно при перемиканні вентилів. При цьому, величина споживаного струму залежить від значення бітів даних, які обробляються в поточний момент часу [3]. Для вимірювання струму, споживаного обчислювальним пристроєм, в його ланцюг харчування додається резистор, напруга на якому вимірюється високоточним цифровим мікровольтметром [3]. Дослідницькі цифрові мікровольтметри дозволяють проводити вимірювання з частотою дискретизації 1-2 ГГц. і точністю до 1мкВ. Промислові загальнодоступні зразки забезпечують частоту дискретизації 50 МГц і точність 5мВ.; Більшість мікроконтролерів, вбудованих мікропроцесорів і смарт-карт працює з тактовою частотою 20-50 МГц.

Судячи з результатів виконаного аналізу публікацій [2], що стосуються технологій реконструкції даних з використанням технологій SPA і DPA, в даний час не існує математичної моделі впливу бітів оброблюваної інформації на споживану потужність. Але, це вплив має місце і може бути виявлено статистичними методами, на яких побудована ідея DPA.

Постановка завдання. Таким чином, проведений аналіз показав, що базові алгоритми модулярного експоненціювання не забезпечують захисту від реконструкції коду експоненти E з використанням SPA. Метою роботи є створення способів захисту від відновлення коду експоненти E за результатами аналізу динаміки потужності, споживаної обчислювальним пристроєм в процесі виконання модулярного експоненціювання.

Метод протидії реконструкції коду експоненти аналізом часовим аналізом динаміки споживання потужності

Наведений аналіз відомих способів протидії реконструкції експоненти, що здійснюється з використанням SPA, показав, що вузловий проблемою є виключення можливості відстеження засобами SPA умовних операторів, що виконують

послідовне тестування розрядів E. Для захисту від SPA необхідно модифікувати алгоритм модулярного експоненціювання таким чином, щоб виключити з нього умовні оператори. В якості основи модифікованого алгоритму вибрано модулярне експоненціювання справа-наліво, оскільки воно не використовує постійних елементів і, відповідно, більш стійкий до DPA. Нижче представлена пропонована модифікація алгоритму модулярного експоненціювання:

1. $A[0] = 1$; $A[1]=1$; $S = X$; $y = 1$;
2. for ($j=0$; $j<s$; j++)
   2.1. for ($l=0$; $l<k$; l++)
   {
   2.1.1. $q = e_j \& y$;11
   2.1.2. $e_j = e_j >> 1$;
   2.1.3. $A[q] = S \cdot A[1] \bmod M$;
   2.1.4. $S = S \cdot S \bmod M$;
   }

Результат $A[1] = X^E \bmod M$.

У запропонованій модифікації відсутні умовні оператори, на кожному циклі виконується одна операція зведення в квадрат і множення, що не дозволяє засобами SPA визначити значення оброблюваного в цьому циклі розряду експоненти. Виняток умовних операторів виконано за рахунок введення паразитного операції множення, результат якого фіксується в A[0], а накопичення правильного результату здійснюється в A[1]. Цілком очевидно, що за рахунок введення паразитного операції множення, обчислювальна складність модулярного експоненціювання зростає на 25%.

В якості другого, ефективного способу протидії SPA-реконструкції експоненти E пропонується виконання модулярного експоненціювання з використанням адитивних ланцюжків.

Адитивним ланцюжком V довжиною L для позитивного цілого числа E називається послідовність $u_0,u_1,\ldots,u_s$ додатних чисел і пов'язана з ними послідовність: $w_1,w_2,\ldots,w_s$ пар $w_i=<i_1,i_2>$, $0 \le i_1, i_2 < i$, які мають наступні властивості:

1. $u_0 = 1$ и $u_L = E$;
2. $\forall u_i$, $1 \le i \le L$, $u_i = u_{i_1} + u_{i_2}$

Алгоритм модулярного експоненціювання $X^E \bmod M$ на основі адитивних ланцюжків має наступний вигляд:

1. $G[0] = X$;
2. for($j=1$; $j<=L$; j++) $G[j] = G[j_1] \cdot G[j_2] \bmod M$;

Результат: $G[L] = X^E \bmod M$.

Наприклад, для $E=12$, можна сформувати адитивний ланцюжок у вигляді багатьох чисел {1,2,3,6,12}, який відповідають пари: : $w_1=<0,0>$, $w_2=<0,1>$, $w_3=<2,2>$, $w_4=<3,3>$. Для тієї ж експоненти $E=12$, можна сформувати інший адитивний ланцюжок, який задається безліччю чисел {1,2,4,5,10,12}, яким відповідають пари: $w_1=<0,0>$, $w_2=<1,1>$, $w_3=<0,2>$, $w_4=<3,3>$, $w_5=<1,4>$. Відповідно до наведеного вище алгоритмом, обчислення X^{12} з використанням останньої ланцюжка виконується в наступній послідовності: G[0]=X;

$G[1]=G[0] \cdot G[0] \bmod M = X^2 \bmod M$;

$G[2]=G[1] \cdot G[1] \bmod M = X^4 \bmod M$;

$G[3]=G[0]\cdot G[2] \bmod M = X^5 \bmod M$;
$G[4]=G[3]\cdot G[3] \bmod M = X^{10} \bmod M$;
$G[5]=G[1]\cdot G[4] \bmod M = X^{12} \bmod M$.

Оскільки наведений алгоритм не містить умовних операторів, то часовий аналіз динаміки споживання потужності при його реалізації практично не дозволить відновити значення *E*, навіть, якщо вдасться відрізнити операції множення і зведення в квадрат. Оскільки для використовуваної на практиці розрядності *n* експоненти (близько 1024) варіантів її розкладання на адитивні ланцюжка існує дуже багато, то відновити конкретно використаний ланцюжок не представляється можливим за результатами аналізу динаміки споживаної потужності.

Ефективність використання адитивних ланцюжків при модулярному експоненціюванні, як засобу протидії SPA, визначається тим, що на практиці експонента *E*, будучи закритим ключем RSA абонента не змінюється . Це дозволяє згенерувати, з використанням випадкових чисел, адитивний ланцюжок, відповідно до якого скомпілювати програму, яка буде безпосередньо виконуватися.

Недоліком використання адитивних ланцюжків є необхідність в додатковому обсязі пам'яті для масиву *G* і зберігання пар *w*, які задають зв'язком компонентів ланцюжків. При цьому необхідний обсяг пам'яті визначається довжиною *L* ланцюжка. Якщо вважати, що середня кількість одиниць в n-розрядній експоненті *E* рівне *n / 2*, то нижня межа значення довжини *L* ланцюжка складає *n + log2n -2*. Для реалізації наведеного вище алгоритму потрібно зберігання відносно невеликого числа елементів масиву *G*: в принципі можна виконати алгоритм при зберіганні всього 2-х чисел зазначеного масиву. Для зберігання пар *w*, потрібний обсяг V_w пам'яті, який визначається формулою:

$$V_w = 2 \cdot (n + \log_2 n - 2) \cdot \left\lfloor \log_2 (n + \log_2 n - 2) \right\rfloor \quad (1)$$

Наприклад, для тих що часто зустрічаються в практиці застосування RSA значення n=1024, V_w = 2838 байт, що не перевищує об’єм вбудованої пам’яті більшості сучасних вбудованих мікроконтролерів.

Висновки. Проведені дослідження, спрямовані на розробку способів протидії відновленню експоненти, здійснюваного шляхом вимірювання та аналізу динаміки споживаної обчислювальної платформою потужності при модулярних експоненціюваннь дозволяють зробити висновки:

1. Операція модулярного експоненціювання лежать в основі ряду криптографічних алгоритмів, таких як RSA, El-Gamal, DSA, які широко використовуються в протоколах захисту інформації в мережах. Експонента зазначеної операції є частиною закритого ключа цих алгоритмів, який змінюється дуже рідко.

2. Найбільшу небезпеку для реконструкції експоненти представляє тимчасової аналіз споживання потужності обчислювальної платформою при виконанні на ній операції модулярного експоненціювання.

3. Ефективним способом протидії реконструкції експоненти тимчасовим аналізом споживання потужності є виняток умовних операторів. Запропоновано алгоритм, що задовольняє цій вимозі, що включає паразитні операції, які ускладнюють ефективний аналіз динаміки споживання потужності. Реалізація запропонованого алгоритму вимагає на 25% більше часу в порівнянні зі звичайним модулярним експоненціюванням.

4. Запропоновано підхід до вирішення завдання протидії тимчасового аналізу споживання потужності для реконструкції секретного коду експоненти. Підхід заснований на використанні технології модулярного експоненціювання на основі адитивних ланцюжків

Запропоновані способи протидії доступу до закритих ключів алгоритмів захисту інформації, в основі яких лежить операція модулярного експоненціювання, можуть бути ефективно використані для підвищення інформаційної безпеки комп'ютерних мереж.

Список використаних джерел.

1. Cohher P. Timing Attack on Implementations of Diffie-Hellman, RSA, DSS and other systems / P.Cohher. // Proceeding of Advances in Cryptology-"CRYPTO-96". LNCS-1882. Springer-Verlag. – 1996. - PP. 104-113.

2. Messerges T.S. Power Analysis Attacks of Modular Exponentiation in Smartcards. / T.S. Messerges., E.A. Dabbish, R.H. Sloan. // Proceeding of 1-th International Workshop "Cryptographic Hardware and Embedded Systems" (CHES-1999), LNCS-1717. Springer-Verlag. - 1999. - P. 145-157.

3. Akkar M.L. Power analysis, what is now possible. / M.L. Akkar., C. Bevan, P. Dischamp, D. Moyart. // Proceeding of International Workshop "Asiacrypt-2000". LNCS-1976. Springer-Verlag. - 2000. - P. 489-502.

4. Mayer-Sommer R. Smartly analyzing the simplicity and power of simple power analysis on smartcards. / R. Mayer-Sommer. // Proceeding of 2-th International Workshop "Cryptographic Hardware and Embedded Systems" (CHES-2000), LNCS-1965. Springer-Verlag. - 2000. - P. 78-92.

5. Kocher P. Differential Power Analysis. / P. Kocher, J. Jaffe, B. Jun. // Proceeding of CRYPTO'99 . - Springer-Verlag. - 1999. - PP.388-404.

6. Марковский А. П. Способ защиты ключей алгоритма ГОСТ 28.147-89 от реконструкции анализом динамики потребляемой мощности. / А. П. Марковский, О. А. Абабне, Ияд Мохд Маджид Ахмад Шахрури. // Вісник національного технічного університету України "КПІ". Інформатика, управління та обчислювальна техніка. – 2007. - № 46. - С.128-138.

ДОВІДКА ПРО АВТОРІВ

Куц Володимир Юрійович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kuts Volodymyr Yuriiovych – associate professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv PolytechnicInstitute".

Гарасимович Галина Володимирівна – студентка, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Harasymovych Halyna Volodymyrivna - student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: harasymovych.hv@gmail.com

Harasymovych Halyna Volodymyrivna,
Kuts Volodymyr Yuriiovych

METHOD FOR PROTECTION MODULAR EXPONENTIATION OPERANDS RECONSTRATION BY POWER ANALYSIS

Topicality of the research. The dynamic intensification of information integration based on the computer networks takes a leading role at the present stage of development of most spheres of human activities. There is an objective need to increase the effectiveness of information security means in terms of expanding the scope of integrated systems. Thus, the scientific task aimed at increasing the effectiveness of information security means is important and topical for the current stage of information technologies development.

Problemstatement. Protection of the keys of algorithms, based on modular exponentiation, against reconstruction by power analysis is important and topical problem at the current stage of the development of technologies for data protection in networks.

Analysis of recent researches and publications. Modular exponentiation X^E mod M, where X, $E < M$ serves as the basic computing operation for the program realization of a large number of public key algorithms, including RSA.

Uninvestigated parts of general matters defining. Following the analysis of publications related to technologies of data reconstruction using SPA and DPA technologies, as of today there is no mathematical model of the impact of bits of the processing information on power consumption. But this impact can be detected by using static method on which the idea of DPA is based.

The research objective. Therefore, the analysis showed that the basic algorithms of modular exponentiation do not protect against reconstruction of the code of exponent $E$by using SPA. The purpose of the research is to create methods of protecting from the restoration of code of the exponent E following the analysis of power consumed by the computing device in the process of modular exponentiation.

Main body. The present analysis of the known methods of counteraction against the reconstruction of the exponent that is performed by using SPA has showed that the key problem is the lack of opportunity of tracking by SPA the conditional operators which perform consecutive testing of E.

Conclusions. Proposed methods of counteracting access to the closed keys of information protection algorithm, based on the operation of modular exponentiation, can be used in order to increase information security in computer networks.

Key words: power analysis, smart cards, cryptographic algorithms, data security protocols, terminal computer devises.

УДК 004.75

Микола Серпученко,
Ірина Клименко

ОГЛЯД СУЧАСНИХ ТЕХНОЛОГІЙ СТВОРЕННЯ ЦИФРОВИХ ПРИСТРОЇВ НА БАЗІ МІКРОПРОЦЕСОРНИХ СИСТЕМ

THE EXPLORATION OF MODERN TECHNOLOGIES OF THE CREATION OF DIGITAL DEVICES BASED ON MICROPROCESSOR SYSTEMS

У цій статті розглядається сучасний ринок мікроелектроніки та досліджуються архітектури сучасних мікроконтролерів та мікропроцесорних систем. В статті визначається поняття сучасної мікропроцесорної системи та розглядаються їх технологічні можливості і області застосування.

Ключові слова: мікропроцесорна система, система на кристалі, *Arduino*.

This article analyzes the modern market of microelectronics and explores the architecture of modern microcontrollers and microprocessor systems. The article defines the concept of a modern microprocessor system and examines its technological capabilities and application areas.

Key words: microprocessor system, System on a Chip, Arduino.

Актуальність теми дослідження. Тенденції розвитку сучасних технологій виготовлення процесорів і їх застосування з кожним роком набирають все більших обертів. Застосовуються нові технології, збільшується кількість ядер на одному кристалі, зростає розрядність процесорів, збільшується кеш пам'ять усіх рівнів, застосовуються нові набори інструкцій і багато іншого. Метою даної статті є огляд сучасного ринку мікроелектроніки та дослідження архітектур сучасних мікроконтролерів та мікропроцесорних систем. В статті визначається поняття сучасної мікропроцесорної системи та розглядаються їх технологічні можливості і області застосування. Також, в статті наведений приклад програмування обчислювальної платформи *Arduino*.

Постановка проблеми. На сьогоднішній день існує багато різних мікропроцесорних систем і їх досить важко класифікувати.Сучасний ринок мікроелектроніки швидко змінюється, що ускладнює завдання його дослідження.

Аналіз останніх досліджень і публікацій. За останні роки з’являються все більше нових різноманітних статей, присвячених мікропроцесорним системам та побудованих на них цифрових пристроях. Однак, сучасні технології використання таких систем досі розкриті не повністю.

Постановка завдання. Завдання полягає в аналізі сучасного ринку мікроелектроніки та вивченні архітектури сучасних мікроконтролерів та мікропроцесорних систем. Іншою метою статті є визначення поняття сучасної мікропроцесорної системи та вивчення її технологічних можливостей.

Викладення основного матеріалу. Сучасні тенденції розвитку мікропроцесорних систем можна представити через узагальнену класифікацію процесорів за областю їх застосування в комп’ютерних системах (рис. 1).

Представлена класифікація є результатом сучасної тенденції універсалізації комп'ютерних систем. Така універсалізація призвела до широкого поширення кластерних суперкомп'ютерів, або так званих високопродуктивних кластерів. Серед очевидних переваг, обумовлених особливостями архітектури та програмного забезпечення кластерних систем – їх відносна дешевизна, простота програмування, простота адаптації програмного забезпечення, відпрацьованого на персональних комп'ютерах.

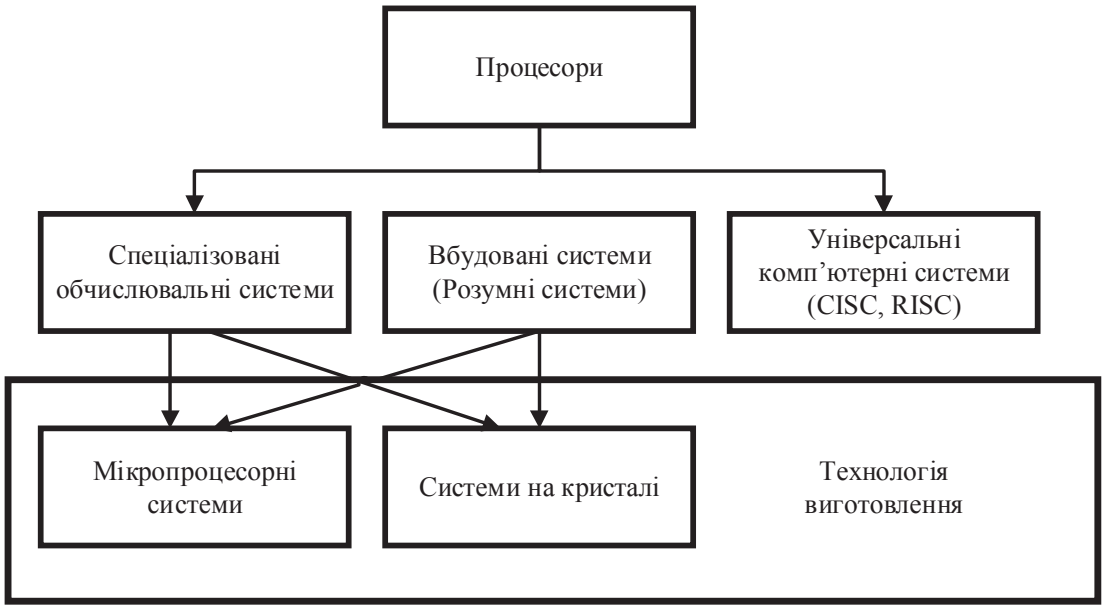


Рис. 1. Узагальнена класифікація процесорів

Універсальні обчислювачі або просто процесори це мікросхеми, які дозволяють виконувати дії над інформацією, такі, як логічну обробку та арифметичні обчислення. Сучасні універсальні процесори обробляють інформацію, представлену на вхідних мовах високого рівня та використовуються в ПК. Універсальні процесори характеризуються широкою областю застосування. Серед них розрізняються процесори з універсальним набором команд (CISC) та скороченим набором команд (RISC).

У CISC (Complex Instruction Set Computer) процесорах використовуються складні команди змінної довжини з великим різноманіттям режимів адресації операндів. Такі команди нагадують оператори мов високого рівня. Такі процесори базуються на фон-Неймановській архітектурі, яка передбачає принцип спільного зберігання команд і даних в пам'яті комп'ютера.

У RISC (Reduced Instruction Set Computer) процесорах розмір команди постійний і, як правило, обмежений одним словом оперативної пам'яті. У таких процесорах усі операції виконуються тільки над операндами, які знаходяться в регістрах процесора. Для завантаження операндів у регістри з оперативної пам'яті існують спеціальні команди. Такі процесори мають гарвардську архітектуру, в якій пам'ять даних і пам'ять команд розділені. Так само, шина даних та шина команд – це два окремі канали.

В умовах сучасного науково-технічного прогресу, вимоги завдань до комп'ютерних систем безперервно зростають, вимагаючи їх ускладнення і

удосконалення. Це підвищує вартість, складність розробки і експлуатації задаче-орієнтованих комп'ютерних систем і ставить під питання економічну і практичну доцільність їх реалізації. Враховуючи, також, що при вирішенні задач інших класів продуктивність обчислень різко знижується.

Спеціалізовані та убудовані обчислювальні системи (ОС) є об’єктно-орієнтованими, що забезпечують найкращу продуктивність за рахунок орієнтації архітектури на рішення специфічної задачі. В результаті абсолютної відповідності цільового призначення, кінцевий користувач завжди отримує максимальну продуктивність обчислень. Сучасна тенденція розробки спеціалізованих та убудованих обчислювальних систем – це широке втілення програмованих логічних інтегральних схем, що надає нових можливостей для удосконалення таких систем – розміщення більш складних пристроїв на меншій площині кристала, економія внутрішніх ресурсів ПЛІС, адаптація до різних класів задач, універсалізація, мультизадачний режим функціонування.

*Спеціалізовані та убудовані обчислювальні системи*використовують для створення різного роду інтелектуальних зовнішніх пристроїв, здатних самостійно виконувати ряд функцій з обробки інформації. Важливою особливістю таких ОС є те, що на них може реалізуватися система команд, найбільш зручна для вирішення певного класу задач. Так, ОС, яка використовується в інтелектуальному пристрою зв'язку з об'єктом, може мати систему команд, орієнтовану на роботу з модулями вводу-виводу аналогових і дискретних сигналів, а також на первинну обробку інформації і реалізацію законів регулювання. Це дозволяє робити такі пристрої особливо ефективними саме для даного застосування.

Фізично, спеціалізовані та убудовані системи можуть бути малими як, наприклад, портативні пристрої (цифрові годинники або MP3-плеєри), або великими, як світлофори, пристрої керування виробництвом та системи керування ядерними електростанціями. Відповідно, складність може бути як низькою – з одним чипом керування, так і дуже високою – з багатьма пристроями, периферійними приладами, комп'ютерними мережами тощо.

*Системи на кристалі*або *SystemonaChip* (*SoC*)будуються на програмованих логічних інтегральних схемах (ПЛІС). На відміну від звичайних цифрових мікросхем, логіка роботи ПЛІС не визначається цілком при виготовленні, а створюється шляхом програмування.

Розумні системи, які також називаються «*Embedded Systems*» – це спеціалізована комп'ютерна система, повністю убудована всередині керованого пристрою, призначена для виконання обмеженої кількості функцій [1].

Технології виготовлення спеціалізованих та убудованих обчислювальних систем

Існує два широко застосовуваних поняття в досліджуваній області – це мікроконтролери або мікропроцесори і мікропроцесорні системи.

Мікропроцесор (*Мр*) або мікроконтролер являє собою пристрій, який реалізує основні елементи комп'ютерної системи на єдиній інтегральній схемі.

Інтегральна схема являє собою напівпровідниковий кристал, який містить набір електронних компонентів, з'єднаних між собою для виконання певної цифрової функції.

Мікропроцесорна система(*MPU*) створюється на одній або на декількох інтегральних схемах діючих як єдине ціле.

Сучасні мікропроцесорні системи включають в себе функціональні можливості годинника, центрального процесору (*CPU*), арифметико-логічного блоку (*ALU*), блоку з плаваючою комою (*FPU*), блоку керування (*CU*), блоку керування пам'яттю (*MMU*), переривань, інтерфейсів вводу-виводу, і кеш-пам'яті. Спеціалізовані мікропроцесори можуть також використовуватися в якості графічних процесорів (*GPUs*), блоків обробки сигналів (*DSPs*), блоку нейронної обробки (*NPU*) і мікроконтролерів.

Система на кристалі(*SoC*) – обчислювальна система, яка містить всі компоненти комп'ютерної системи, включаючи додаткові функції, які зазвичай надаються допоміжними мікросхемами. Вони можуть включати в себе такі речі, як *Wi-Fi*, *Ethernet*, *SD*-карта, *ADC*, *DAC*, *LCD* драйверів і ПЛІС. *SoCs* здатні працювати з повноцінними сучасними операційними системами і з усіма їхніми функціями.

Функціональність мікропроцесорних систем

У найширшому розумінні базова функціональність мікропроцесора полягає в:

– безперервному зчитуванні цифрових даних, що складаються з інструкцій і, можливо, значень;

– виконанні їх шляхом інтерпретації інструкцій; виконанні певної операції над значеннями;

– виведення результату.

Узагальнена базова функіональність мікропроцесорної системи зображена на рис. 2.

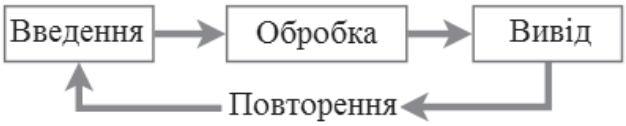


Рис. 2. Узагальнена базова функіональність мікропроцесорної системи

У той час як базові функції властиві усім мікропроцесорам, вони сильно розрізняються за типом і розміром даних, які вони обробляють, за типами операцій, які вони підтримують, за те, як вони виконують ці операції, за їх призначенням і за їх характеристикам продуктивності.

Технічні характеристики мікропроцесорних систем

Технічні характеристики мікропроцесорів засновані на мікроархітектурі убудованого процесора, задіяної напівпровідникової технології і властивості всієї системи. Нижче наведені деякі загальні специфікації:

– технологія – напівпровідникова технологія, яка використовується для створення MPU (наприклад, CMOS, BiCMOSiTTL)

– процес – розмірі правила дизайну. Розмір функції сам по собі (наприклад,1 0 мкм).

– мікроархітектура – функціональний опис базової схеми мікропроцесора.

– розмір слова – розмір слова мікропроцесора зазвичай відноситься конкретно до використовуваному розміру слова даних – тобто максимальній

довжині операнда, використовуваного для керування стандартними цілочисельними значеннями.

– кількість ядер – кількість фізичних ядер, убудованих в чіп.

– порядок байтів – порядок байтів, який використовується мікропроцесором при роботі з багатобайтовими значеннями.

– базова частота – внутрішня робоча частота ядра ЦП. Це один з багатьох параметрів, які використовуються для оцінки продуктивності мікропроцесора.

– пакет – це фізичний корпус мікропроцесора. Найчастіше поєднується з гніздом, що представляє собою з'єднання, які знаходяться на самій платі, де вставлений пакет.

Архітектура мікропроцесорних систем

Мікропроцесори виготовляються з кремнієвого матеріалу, який містить крихітні електричні компоненти, убудовані на поверхню мікрочіпа. У стандартний мікропроцесор зазвичай входять наступні компоненти:

Арифметичний і логічний блок (*ALU*) виконує математичні обчислення, такі як віднімання, складання, ділення і булеві функції. Булеві функції – це тип логіки, використовуваної для функціональних елементів. *ALU* також виконує порівняння і логічне тестування. Процесор передає сигнали в *ALU*, який інтерпретує інструкції і виконує обчислення.

Регістри – тимчасові місця зберігання даних. Ці області пам'яті підтримують дані, такі як інструкції комп'ютера, адреси сховища, символи та інші дані. У деяких командах комп'ютера може знадобитися використання певних регістрів як частина команди. Кожен регістр має певну функцію, таку як регістр команд, лічильник, регістр акумулятора і регістр адреси пам'яті.

Блоки керування приймають сигнали від ЦП, які інструктують блок керування переміщати дані з одного мікропроцесора на інший. Блок керування також направляє арифметичний і логічний блок. Блоки керування складаються з декількох компонентів, таких як декодер, схеми синхронізації і логіки керування. Працюючи разом, ці пристрої передають сигнали в певні місця на процесорі. Наприклад, декодер отримує команди від додатка. Декодер інтерпретує інструкції і вживає заходи. Він відправляє сигнали в *ALU* або направляє регістри для виконання певних завдань. Блок логіки керування передає сигнали в різні секції мікропроцесора і регістрів, які інформують ці компоненти про виконання дій. Годинники посилають сигнали, які синхронізують і забезпечують своєчасне виконання команд і процесів.

*Шини.*Мікропроцесори мають систему шин, які переміщують дані. Шини відносяться до класифікацій провідників, які мають конкретні завдання і функції. Шина даних передає дані між центральним процесором і оперативною пам'яттю (ОЗУ) – основний пам'яттю комп'ютера. Шина керування відправляє інформацію, необхідну для координації і керування декількома завданнями. Адресна шина передає адресу між ЦП і ОЗП для оброблюваних даних.

*Кеш пам'ять***.** Деякі вдосконалені мікропроцесори мають кеші пам'яті, які зберігають останні дані, які використовуються ЦП. Накопичувачі пам'яті прискорюють обчислювальний процес, тому що ЦПУ не повинен звертатися до більш повільного ОЗУ для вилучення даних. Багато комп'ютерів мають кеші 1

або 2 рівня, деякі системи мають кеші рівня 3. Рівень кеша вказує порядок, в якому ЦП перевіряє дані, починаючи з рівня 1. Виробники часто інтегрують кеші рівня 2 і рівня 3 в мікропроцесор, що збільшує швидкість обробки.

Сімейства мікропроцесорних систем

Сімейство мікропроцесорів є групою мікропроцесорів, мікроконтролерів, мікросхем або компонентів, розробленої конкретними виробником, які поділяють ту ж загальну архітектуру і набір функцій. У конкретному сімействі чіпи можуть бути додатково класифіковані відповідно до їх моделлю, яка може мати різні можливості, тактові частоти і додатки. Моделі можуть бути далі розділені на ступені, які є незначними уточненнями і налаштуваннями для виправлення і підвищення стабільності і надійності моделі.

Мікропроцесорні системи сімейства *Arduino*

Arduino – це обчислювальна платформа, яка представляє собою комплекс апаратних і програмних компонентів. Ця платформа дає змогу розробляти інтерактивні цифрові системи, які приймають дані від різних комутаторів, датчиків та індикаторів для керування двигунами та іншими фізичними об'єктами [1].

Основними компонентами платформи *Arduino* є мікроконтролер та плата з компонентами периферії, такої як елементи вводу/виводу. Програмна частина складається з середовища розробки *Processing/Wiring* на мові програмування, що є підмножиною C/C++. *Arduino* може використовуватися як для створення автономних інтерактивних цифрових систем, в тому числі таких, що працюють в режимі реального часу, так і підключатися до програмного забезпечення, яке виконується на комп'ютері (наприклад: *Processing, Adobe Flash, Max/MSP, Pure Data, SuperCollider*). Інформація про плату (рисунок друкованої плати, специфікації елементів, програмне забезпечення) знаходяться у відкритому доступі і можуть бути використані тими, хто бажає створювати плати власноруч.

Апаратна частина плат *Arduino*

Плата *Arduino* складається з мікроконтролера *AtmelAVR*, а також елементів обв’язки для програмування та інтеграції з іншими пристроями. На багатьох платах наявний лінійний стабілізатор напруги +5В або +3,3В. Тактування здійснюється на частоті 16 або 8 МГц кварцовим резонатором. У мікроконтролер записаний завантажувач (*bootloader*), тому зовнішній програматор не потрібен.

На концептуальному рівні усі плати програмуються через *RS*-232 (послідовне з’єднання), але реалізація даного способу різниться від версії до версії. Новіші плати програмуються через *USB*, що можливо завдяки мікросхемі конвертера *USB-to-SerialFTDI FT232R*. У версії платформи *Arduino* Uno як конвертер використовується контролер *Atmega8* у *SMD*-корпусі. Дане рішення дозволяє програмувати конвертер таким чином, щоб платформа відразу розпізнавалася як миша, джойстик чи інший пристрій за вибором розробника зі всіма необхідними додатковими сигналами керування. У деяких варіантах, таких як *ArduinoMini* або неофіційній *Boarduino*, для програмування потрібно підключити до контролера окрему плату USB-to-Serial або кабель.

Плати *Arduino* дозволяють застосовувати значну кількість виводів мікроконтролера як вхідні/вихідні контакти у зовнішніх схемах. Наприклад, у платі *Decimila* доступно 14 цифрових входів/виходів, 6 із яких можуть

генерувати ШІМ-сигнал, і 6 аналогових входів. Ці сигнали доступні на платі через контактні площадки або штирові розніми. Також існує багато різних зовнішніх плат розширення, які називаються «*shields*» («щити»), які приєднуються до плати *Arduino* через штирові роз'єми.

Незважаючи на те, що платформу *Arduino* не вийде використовувати в якості комп'ютера, її низькі вимоги потужності дозволяють їй ефективно контролювати інші пристрої.

Raspberry Pi

Raspberry Pi – це недорогий комп'ютер розміром з пластикову карту, який можна підключати до монітора комп'ютера або телевізора. Керувати ним можна за допомогою стандартної клавіатури і миші. З його допомогою можна робити все, що і на звичайному комп'ютері – переглядати веб-сторінки, програвати відео високої роздільної здатності, створювати електронні таблиці, редагувати текстові документи і грати в ігри [2].

Beaglebone

Комп'ютер *Beaglebone* аналогічний *Raspberry Pi* в розмірах, вимогах потужності і застосуванні. *Beaglebone* перевершує *Raspberry Pi* в обчислювальній потужності, тому є оптимальним вибором для додатків з більш високими вимогами [3].

Приклад програмування мікропроцесорної системи *Arduino*

Програми для систем *Arduino* розробляються на *C++* подібній мові. Будь-яка така програма складається, що найменше, з двох обов'язкових функцій:***setup()*** та ***loop()***.

Функція ***setup()*** запускається один раз, після кожного включення живлення або скидання плати *Arduino*. У тілі даної функції пишеться код для створення змінних, установки режиму роботи цифрових портів, та інших ініціалізацій. Цей механізм буде розглянуто у подальших прикладах.

Функція ***loop()*** в нескінченному циклі послідовно раз по раз виконує команди, які описані в її тілі. Тобто після завершення функції знову відбудеться її виклик.

Розглянемо, для прикладу, просту програму, що змушує світлодіод проблискувати:

```
void setup()
{
pinMode(13, OUTPUT);
}
void loop()
{
digitalWrite(13, HIGH);
delay(100);
digitalWrite(13, LOW);
delay(900);
}
```

Кожен вираз – це наказ процесору зробити щось. Вирази в рамках одного блоку виконуються один за одним, суворо по порядку без пауз і перемикань. Тобто кожен конкретний блок коду читати зверху вниз, щоб зрозуміти що відбувається. Самі блоки також мають свій порядок виконання:

1. Після ввімкнення *Arduino*, або натискання кнопки *RESET*, викликається функція *setup*. Тобто вирази в ній виконуються послідовно підряд.

2. Як тільки робота *setup* завершується, відразу ж викликається функція *loop*.

3. Після того, як робота *loop* завершується, відразу функція *loop* викликається ще раз і так до нескінченності.

Розглянемо докладніше кожен вираз даного коду і з'ясуємо чому виконання цієї програми призводить до мерехтіння світлодіода. Як відомо, піни *Arduino* можуть працювати і як виходи і як входи. Коли є необхідність чимось керувати, тобто видавати сигнал, потрібно перевести керуючий пін в стан роботи на вихід. У нашому прикладі керування світлодіодом відбувається на тринадцятому піні, тому тринадцятий пін перед використанням потрібно зробити виходом. Це робиться виразом в функції *setup*:

pinMode(13, OUTPUT);

Вирази бувають різними: арифметичними, деклараціями, визначеннями та умовними. В даному випадку в виразі здійснюється виклик функції. В *setup* викликається функцію з ім'ям *pinMode*. Вона встановлює заданий за номером пін в заданий режим: вхід або вихід. Номер піна та режим вказується в круглих дужках, через кому, відразу після імені функції. У нашому випадку небхідно, щоб тринадцятий пін працював як вихід. *OUTPUT* означає вихід, *INPUT* – вхід.Уточнюючі значення, такі як 13 і *OUTPUT* називаються аргументами функції.

Розглянемо функцію loop:

```
voidloop()
{
digitalWrite(13, HIGH);
delay(100);
digitalWrite(13, LOW);
delay(900);
}
```

Функція *loop*, як зазначалося, викликається відразу після *setup*. І викликається знову і знову після свого завершення. Функція *loop* називається основним циклом програми і ідеологічно призначена для виконання корисної роботи. У нашому випадку корисна робота – це миготіння світлодіодом.

Дослідимо вирази по порядку. Отже, перший вираз – це виклик убудованої функції *digitalWrite*. Вона призначена для подачі на заданий пін логічного нуля (*LOW*, 0 вольт) або логічної одиниці (*HIGH*, 5 вольт) У функцію *digitalWrite* передається 2 аргументи: номер піна і логічне значення. У підсумку, в першу чергу запалюється світлодіод на тринадцятому піні, після подання на нього 5 вольт. Як тільки це зроблено процесор моментально приступає до наступного виразу. В даному випадку – це виклик функції *delay*. Функція *delay* – це убудована функція, яка змушує процесор зробити затримку на певний час. Ця функція приймає лише один аргумент: час в мілісекундах, на який слід зробити затримку. У цьому випадку це 100 мс. Під час затримки стан системи не змінюється, тобто світлодіод продовжує горіти. Як тільки 100 мс закінчуються, процесор одразу переходить до наступного виразу. У цьому прикладі це знову виклик убудованої функції *digitalWrite*. Однак, цього другим аргументом передається значення *LOW*. Тобто на тринадцятому піні встановлюється логічний нуль, подається 0 вольт, та

світлодіод згасає. Після того, як світлодіод згас виконується наступний виразу. Це знову виклик функції *delay*. Цього разу робиться затримка на 900 мс. Як тільки затримка закінчується, функція *loop* завершується. Після завершення вона викликається ще раз і все відбувається знову.

Висновки

На сьогоднішній день мікропроцесорні системи широко використовуються в багатьох пристроях керування та системах з, як правило, малим споживанням енергії. Вони є актуальними у пристроях, що повинні мати відносно невеликі розміри, оскільки мають широкий функціонал можливостей, при своїй помірній ціні. На прикладі програмування мікропроцесорної системи *Arduino* була показана доступність таких систем для користувачів, що не мають поглиблених знань у схемотехніці та мікропроцесорній архітектурі.

Список літератури

1. В. Н. Кинягин, М. С. Липецкая, И. Б. Андреев «Умные» среды, «Умные» системы, «Умные» производства, 2012. – с. 9-10.

2. Arduino official site, 2018. [Електронний ресурс] – Режим доступу: https://www.arduino.cc/en/Guide/Introduction. – Дата звернення: 28.05.2018.

3. Офіційний сайт обчислювальної платформи Raspberry Pi, 2018. [Електронний ресурс] – Режим доступу: https://www.raspberrypi.org/about/. Дата звернення: 27.05.2018.

4. Офіційний сайт обчислювальної платформи Beaglebone, 2018. [Електронний ресурс] – Режим доступу: http://beagleboard.org/about. Дата звернення: 25.05.2018.

ДОВІДКА ПРО АВТОРІВ

Серпученко Микола Вадимович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Serpuchenko Mykola – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: nikolayserpuchenko@gmail.com

Клименко Ірина Анатоліївна – д.т.н., доцент, професор кафедри обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Klymenko Iryna –Doctor of Technical Sciences, Professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: ikliryna@gmail.com

MykolaSerpuchenko,
IrynaKlymenko

THE EXPLORATION OF MODERN TECHNOLOGIES OF THE CREATION OF DIGITAL DEVICES BASED ON MICROPROCESSOR SYSTEMS

Relevance of the research topic. The tendency of the development of modern processor manufacturing technologies and their applying evolves every year. There are used new technologies; the number of cores on one crystal increases, the processor runs faster, cache memory of all levels increases, new sets of instructions are used, and more. This article analyzes the modern market of microelectronics and explores the architecture of modern microcontrollers and microprocessor systems. The article defines the concept of a modern microprocessor system and examines its technological capabilities and application areas. Also, there is given an example of the programming of the Arduino computing platform.

Formulation of the problem. There are too many different microprocessor systems and it is quite difficult to categorize them all.

The analysis of recent research and publications. Over the past years more and more articles are devoted to exploration of the microprocessor systems, in particular, based on them digital devices. However, the modern technologies of usage these systems are not covered enough.

Setting objectives. The task is to analyze the modern market of microelectronics and explore the architecture of modern microcontrollers and microprocessor systems. One more aim of the article is to define the concept of a modern microprocessor system and examine its technological capabilities.

Presentation of the main material. The article has showed the generalized classification of the modern computing systems. There have been also explained the functionality of microprocessor systems and their generalized structure and architecture. The example of programing microprocessor system Arduino has clarified the way of usage such systems.

Conclusions. The modern microprocessor systems are widely used in many control systems and systems with generally low energy consumption. They are relevant in devices that should be relatively small in size, since they have a wide range of functionality at a moderate price. In the programming example of the Arduino microprocessor system, there has been shown the availability of such systems for users without advanced knowledge in circuit design and microprocessor architecture.

Key words: microprocessor system, System on a Chip, Arduino.

УДК 004.75

Соляніков Павло

ДОСВІД ЗАПУСКУ РОБОТИЗОВАНОГО ЗВАРЮВАЛЬНОГО КОМПЛЕКСУ

EXPERIENCE OF THE LAUNCH OF A ROBOTIC WELDING COMPLEX

У цій статті я поділюся своїм досвідом роботи з промисловим роботом, розповім про актуальність роботизації на сьогоднішній день та можливості робота с машинним зором.

Ключові слова: роботизація, машинний зір, lasertracking, ABB.

In this article, I will share my experience working with an industrial robot, I will talk about the relevance of robotics and capabilities of a robot with machine vision.

Key words: robotics, machine vision, ABB.

Актуальність теми дослідження та постановка проблеми.

Автоматизація - широке поняття, яке описує всі процеси, які виконуються автоматично, за допомогою ПО або роботів. Тобто, це ті завдання, які виконуються без участі людини. Роботизація ж охоплює лише ту частину цієї практики, коли на заміну людям приходять фізичні механізми.

Уже зараз людину оточує велика кількість технологій і розробок, які автоматизують процеси. Цей підхід також активно починає використовувати бізнес. Останнім часом все більше компаній переходять до автоматизації, а також починають залучати роботів для виконання деяких завдань. Завдяки цьому ефективність бізнесу зростає, а витрати на персонал знижуються.

Аналіз останніх досліджень і публікацій.Четверта промислова революція прискорюється: до 2018 року в світі введено в експлуатацію близько 1,3 мільйона промислових роботів. Інвестиції в промислових роботів в високоприбутковому автомобільному секторі збільшилися з 2013 по 2014 рр на рекордні 43%.

За даними Міжнародної федерації робототехніки (IFR), в 2015 році обсяг світового ринку роботехнічних систем оцінювався приблизно в 32 млрд доларів.

Щільність роботизації є ключовим показником ефективності для оцінки поточного ступеня автоматизації виробництва. Середня щільність роботизації в світі складає 66 роботів на 10 000 співробітників. У двадцяти однієї країни цей показник перевищує середньосвітове значення. Більша частина з них знаходяться в Європейському Союзі. У топ-список також входять азіатські країни - Південна Корея, Японія, Тайвань, а також США і Канада.

Нинішній світовий лідер в промислової роботизованою автоматизації - Південна Корея. Щільність роботизації в цій країні перевищує середньосвітові показники в цілих сім разів -478 роботів на 10 000 співробітників. Також до лідируючої трійки входять Японія (314) і Німеччина (292).

Виділення недосліджених частин загальної проблеми. Недослідженим залишається те, що незрозуміло, чи вигідний процес роботизації для сучасної людини.

Все більше компаній модернізують свої підприємства роботами, люди лишаються своїх робочих місць, адже в них немає вибору. Роботи - надійні робітники, які не потребують вихідних чи відпочинку.

Професор Стівен Хокінг, один з найбільш шанованих і відомих британських вчених, заявив, що зусилля по створенню мислячих машин можуть привести до того, що існування людства опиниться під загрозою. Також це стосується і глобальної роботизації.

Але якщо не рухатися вперед, то не буде сенсу всієї цивілізації і, на мою думку, процес роботизації та створення штучного інтелекту неминучий.

Викладення основного матеріалу.



Рис. 1. Маніпулятор IRB 2400.

Історія знайомства робота зі мною.

Влітку 2016 року мій батько замислився над питанням створення власного виробництва і було прийнято рішення придбати у Німеччині зварювальний б\у робот ABBirb 2400.

Батько звернувся до спеціалістів по роботах abbщодо налаштування робота таотримав відповідь, що відновлення робота неможливе.

Після остаточного рішення я запитав у батька чому робот стоїть без роботи. І, дізнаючись проблему, я вирішив запустити робот власноруч. Влітку 2017 року я почав вивчати систему та розділив роботу на 5 етапів, на які я витратив близько 7 місяців.

Етап 1: Підключення робота до мережі та встановлення софту.

Блок керування роботом складається з чотирьох частин :

- Підключення до мережі ,Input\Output та Safety сигнали.
- Керування робота частотними перетворювачами.
- RobotComputer.
- Зварювальний апарат.

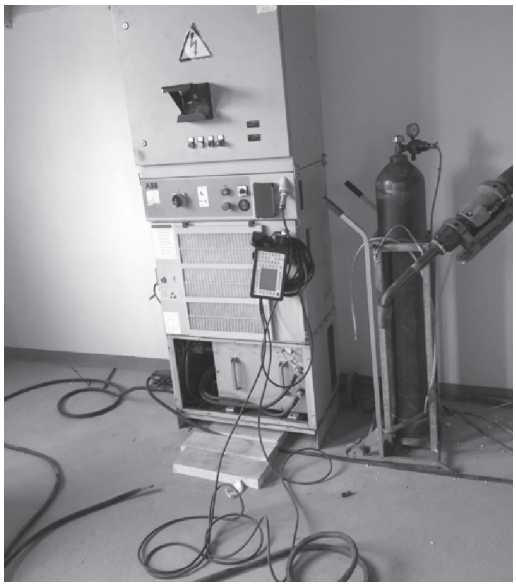


Рис. 2. Блок керування IRB 2400.

Через те, щоцей робот везли з Німеччини, всі з’єднання були розрізані іприйшлось розбиратися, як правильно всепід'єднати, щоб нічого не спалити.

З’єднань було так багато, що нам знадобилося чимало часу, щоб все правильнопід'єднати.

Далі я почав вивчати етапи встановленнясофту та основи ручної маніпуляції роботом. З цим мені допомоглипонад десятьрізних мануалів ABB.

Етап 2: Програмування робота та керування Input/Output сигналами, Калібрування.

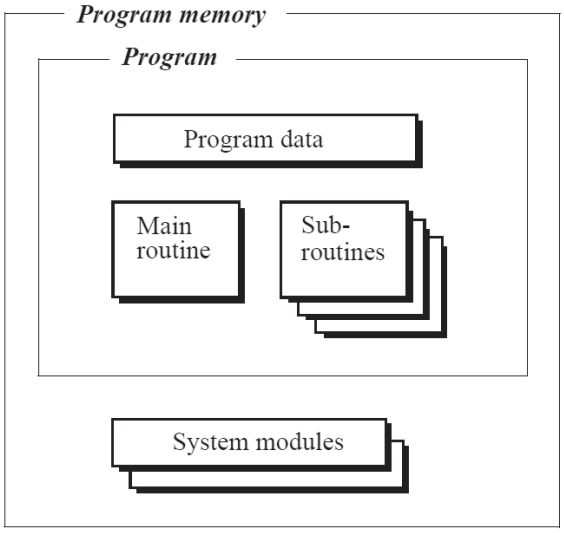


Figure 1 The program instructions control the robot and robot peripherals.

Рис. 3. Схема програмування робота.

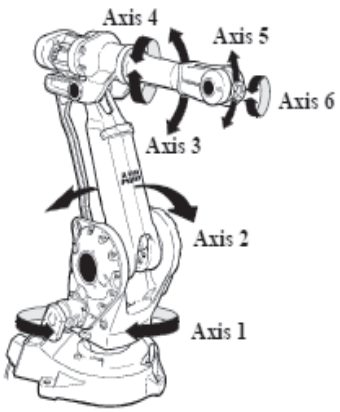


Рис. 4. 6 ступенів свободи робота.

Роботи ABB програмуються за допомогою мови програмування RAPID.

Я почав програмувати на RAPID та створював свої перші програми, які робот вже виконував. Але потрібно обов'язково було зробити калібрування робота (для точності роботи).

Далі я почав вивчати програмування сигналами Input/Output.

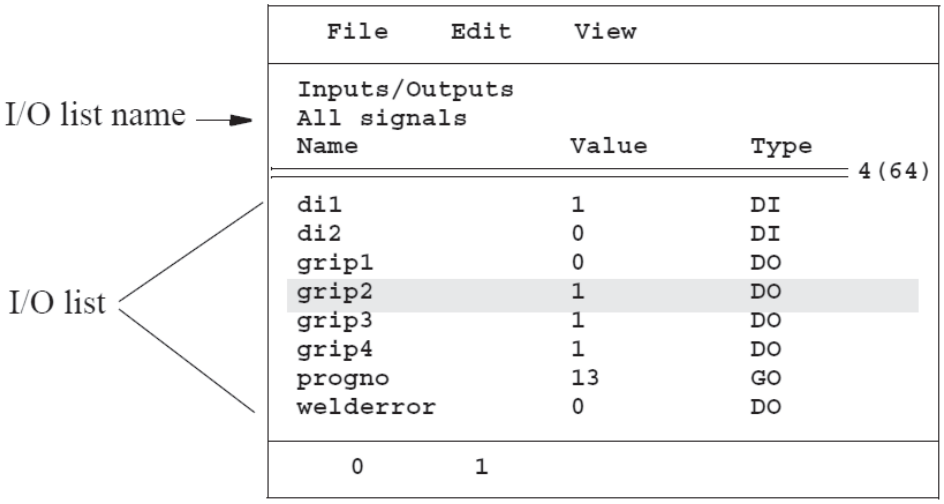


Рис. 5. Меню програмування I/O сигналів.

Наприклад:
IFdi1 = 1
DO
…
(якщо на комірку do1 подається 24V, то роби наступні дії)

Етап 3: Принцип роботи частотних перетворювачів, створення позиціонера. Після тривалої роботи маніпулятора я виявив дуже неприємну несправність: Третя вісь робота перестала рухатися. Було 4 теорії: поганий контакт силових кабелів, проблеми із софтом, несправність мотору, несправність частотного перетворювача.

Я перевірив всі контакти, перевстановив софт, дістав та протестував мотор, але проблема залишилась. Залишилося тільки зрозуміти принцип роботи частотних перетворювачів і спробувати полагодити його.

І я з батьком полагодив робот.

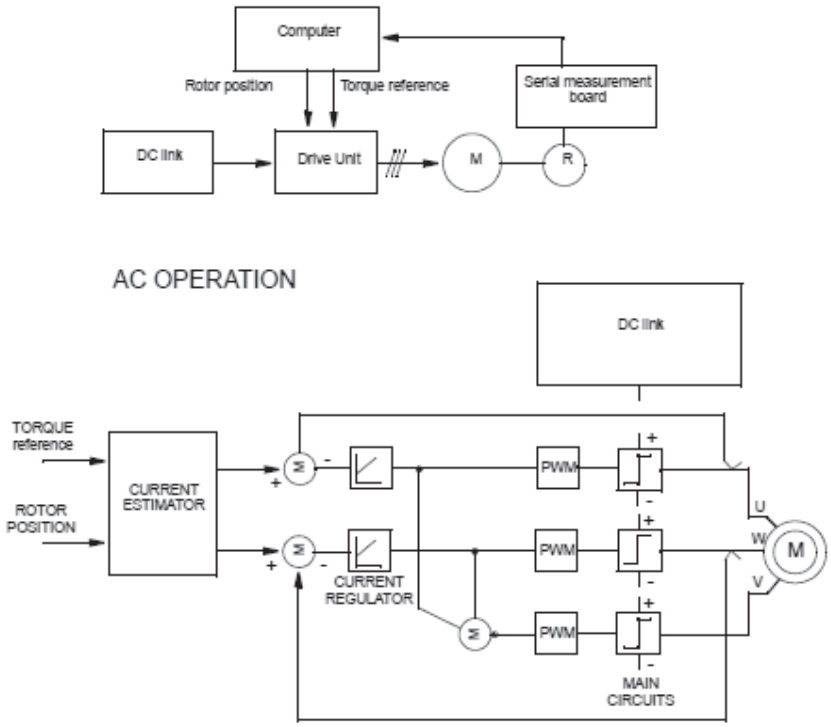


Рис. 6. Схема частотних перетворювачів робота ABBIRB 2400.

Важливо те, що, зрозумівши цей матеріал, у мене з'явилась можливість підключити додаткову сьому вісь (позиціонер).

Етап 4: Софт для сварки такерування зварювальними процесами.
Кожен робот розробляється для конкретної задачі.
І для зварювального робота йшов додатковий софт.

За допомогою нього комп'ютер спілкується з платою зварювального апарату і відбувається зварювальний процес.

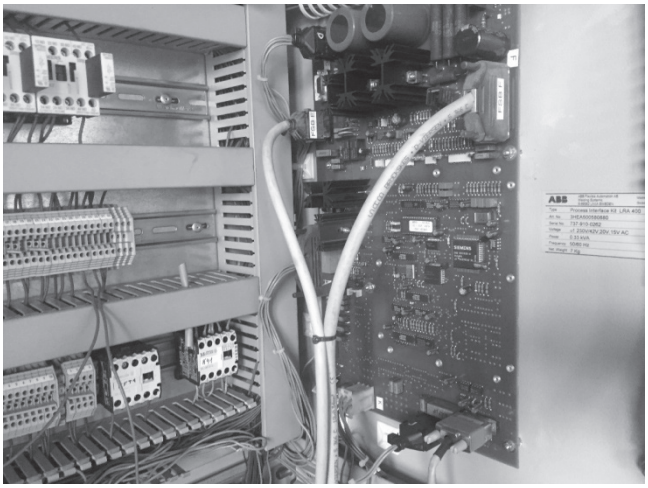


Рис. 7. Плата зварювального апарату.

Етап 5: Запуск системи, DeviceNetта машинний зір.

На даний момент, робот вже працює. Я створив сайт, де пропонується послуги зварювального робота та можливість модернізації підприємства роботом.

http://rws-k.com

Зараз я займаюся над проектом підключення міні-комп'ютера raspberrypідо системи за допомогою протоколу DeviceNet.

До raspberrypi буду підключено камеру, яка передає зображення міні-комп'ютеру, а він, в свою чергу, обробляє зображення на наявність заданих предметів (реалізація на мові Pythonз використанням бібліотеки OpenCV).

Якщо на зображенні є потрібний предмет, то raspberrypi відправляє координати до системи робота за допомогою DeviceNet і робот рухається у задану точку та, наприклад, забирає об'єкт.

Системи машинного зору дуже часто використовують на підприємствах, наприклад:

Weldguide IV - це потужний сенсор відстеження шву, заснований на запатентованій технології та призначений для робототехнічних систем зварювання ABB.

Weldguide IV забезпечує функціональність відстеження шляхом читання справжніх значень імпедансу, близьких до дуги, на частоті 25 кГц, а потім направляє робота на правильний шлях. Weldguide IV призначений для відстеження важких варіантів зварювального з'єднання, що виникає внаслідок складових компонентів або інших проблем, що виникають перед процесом.

Висновки.

Нещодавно ми тільки запитали себе, що таке роботизація (AutomationofRoboticProcessAutomation, RPA) і яка користь може отримати бізнес від впровадження цієї технології? Мало хто чекав, що РПА буде так стрімко розвиватися і проникнути в різні сфери бізнесу, а також, що багато компаній вже найближчим часом будуть мати досвід використання можливостей цієї технології поряд із штучним інтелектом (Штучний інтелект, AI).

Великий потенціал AI, належним чином інтегрований з програмними роботами, робить майбутнє RPA дуже багатообіцяючим і, безперечно, виведе бізнес на новий рівень операційної ефективності. AI, ML і BI дозволять автоматизувати через роботизацію ще багато процесів, включаючи і ті, які вимагають прийняття рішень, суджень, обробки нетривіальних сценаріїв і неструктурованих даних. Для компаній, які хочуть досягти значного ефекту від роботизації бізнес-процесів, дуже важливо вже зараз розпізнати вигоди комбінування RPA з різними типами AI, ML, BI і когнітивних обчислень, а також почати впроваджувати когнітивні технології на ранній стадії розгортання RPA, щоб досягти успіху і бути лідером на ринку.

Список літератури

1. ABB Robotics Products AB. (2004). *Product-Specification-2400* (pp. 5 - 45).

ДОВІДКА ПРО АВТОРІВ

Соляніков Павло Олександрович – студент 2го курсу, група ІП-64, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Solyanikov Pavlo- 2nd year student, IP-64 group, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: pashasolyanikov@gmail.com

Solyanikov Pavlo

EXPERIENCE OF THE LAUNCH OF A ROBOTIC WELDING COMPLEX

Relevance. Robots in antiquity and through the Middle Ages were used primarily for entertainment. However, the 20th century featured a boom in the development of industrial robots. Through the rest of the century, robots changed the structure of society and allowed for safer conditions for labor. In addition, the implementation of advanced robotics in the military and NASA has changed the landscape of national defense and space exploration. Robots have also been influential in the media and profitable for toy manufacturers.

Problem Statement. Replacement of human work, improved production, lower costs, increased productivity

Analysis of recent research and publications. The Boston Consulting Group predicts that the share of robot-solved tasks will increase from 8% today to 26% by 2025. The leaders of the robotic production will be China, Germany, Japan, South Korea and the United States. Together, 80% of all robots' purchases will be needed for their share. Unlike people who can double productivity over 10 years, robots are able to double their every four years - this is Sirkin's analyst from BCG.

General model structure. Robots will soon be everywhere, in our home and at work. They will change the way we live. This will raise many philosophical, social, and political questions that will have to be answered. In science fiction, robots become so intelligent that they decide to take over the world because humans are deemed inferior. In real life, however, they might not choose to do that.

Robots will be commonplace: in home, factories, agriculture, building & construction, undersea, space, mining, hospitals and streets for repair, construction, maintenance, security, entertainment, companionship, care.

Conclusions. It is worthwhile to understand that even today robots play an important role in production, they constantly increase their influence on all spheres of industry, they show special efficiency and the opportunity to increase production rates.

Key words: robotics, machine vision, ABB.

УДК 004.315

Жабін В. І., Токар А. Г.

АНАЛІЗ ЕФЕКТИВНОСТІ РЕАЛІЗАЦІЇ ДОДАВАННЯ В НЕАВТОНОМНОМУ РЕЖИМІ

ANALYSIS OF EFFICIENCY OF ADDITION OPERATION REALIZATION IN ON-LINE MODE

Розглядається алгоритм та апаратна реалізація операції додавання різної кількості операндів в неавтономному режимі. Такий режим дозволяє з використанням надлишкової системи числення поєднувати процеси порозрядного введення, обробки і порозрядного виведення даних, починаючи зі старших розрядів. Показано, що застосування багатоопераднної функції додавання дозволяє зменшити затримку формування розрядів результату та апаратні витрати.

Ключові слова: неавтономні обчислення, надлишкові системичислення, суміщення залежних операцій.

Рис.: 3. Табл.: 1. Бібл.: 6.

The article deals with an algorithm and hardware implementation of two-place and multi-place functions of addition in on-line mode, which allows to combine the processes of bit-by-bit input, processing and bit-by-bit output of information, starting with the high-order bit with the use of redundant number system. It is shown that the implementation of a multi-place function enables to reduce delay in the result digit obtaining, as well as scope of equipment.

Key words: on-line computations, redundant number system, dependent operations overlap.

Fig.: 3. Tabl.: 1. Bibl.: 6.

Актуальність теми дослідження. Рішення складних завдань обробки інформації в реальному часі обумовлює застосування паралельних обчислювальних систем. Час рішення задач в таких системах залежить не тільки від часу виконання паралельних гілок алгоритмів, але і від витрат часу на обмін інформацією між обчислювальними модулями (ОМ). Зменшити витрати часу на обмін даними дозволяє використання потокових систем з безпосередніми зв'язками (ПСБЗ) між ОМ [1, 2]. Підвищення ефективності зазначених систем є актуальним завданням, що вимагає додаткових досліджень.

Постановка проблеми При побудові систем на базі ПЛІС важливою проблемою є економне використання ресурсу мікросхем (внутрішніх функціональних блоків і зовнішніх висновків) з метою зменшення числа мікросхем для реалізації обчислювальної системи. У роботі розглядається можливість вирішення цього завдання за рахунок зменшення кількості з'єднань між компонентами системи.

Аналіз останніх досліджень і публікацій. Виконання операцій методами паралельної арифметики вимагає передачі даннях між ОМ багаторозрядними кодами, що не дозволяє економити елементи комутації ПЛІС. Для скорочення числа зв'язків розроблені методи квазіпараллельний арифметики [1, 2]. Обчислення в цьому випадку в кожному ОМ виконуються при порозрядному введенні операндів і порозрядному формуванні результату зі старших розрядів. Для виключення переносів в старші розряди використовуються надлишкові

системи числення. Завдяки цьому при виконанні ланцюжків залежних за даними операцій виникає можливість часткового суміщення операцій на рівні обробки розрядів операндів, що показано в ряді робіт [3-6].

Виділення недосліджених частин загальної проблеми. При виконанні ланцюжків залежних операцій кожен ВМ працює в неавтономному (on-line) режимі, приймаючи на кожному кроці чергові розряди операндів від попередніх ОМ і передаючи розряди результату наступного в ланцюжку ОМ. Зі збільшенням довжини ланцюжка операцій збільшується час обробки даних. Для скорочення кількості операцій в ланцюжку доцільно використовувати ОМ, які виконують багатомісні операції. Розробка і дослідження алгоритмів багатомісних операцій є важливим завданням підвищення ефективності виконання залежних операцій.

Постановка завдання. У даній роботі на прикладі операції додавання досліджується ефективність реалізації багатомісних операцій при порозрядної обробці операндів. Розглядається апаратна реалізація операції при різному числі операндів. Операція складання кількох чисел може бути корисна при цифровій обробці сигналів, реалізації штучних нейронів, обробці таблиць даних.

Викладення основного матеріалу. У роботі [2] запропонований метод складання двох чисел поданих в надлишковій системі числення з природним порядком ваг, основою $k = 2$ і цифрами з множини {-1,0,1}.При виконанні ланцюжка залежних за даними операцій ОМ працюють синхронно наступним чином [2]. На кожному кроці ОМ приймає по одному розряду операндів з попереднього в ланцюжку ОМ і формується один розряд результату для наступного ОМ. Операція виконується зі старших розрядів. Отриманий розряд результату не вимагає корекції, оскільки в надлишкової системі числення переноси в старші розряди результату не поширюються.

Показано [2], що розряди результату формуються с затримкою на $p = 2$ циклів, тобто ОМ виконує двомісну операцію $Z = 2^{-p} \cdot (X + Y)$, причому, операнди є правильними r-розрядними дробовими числами:

$$X = \sum_{i=1}^{r} x_i 2^{-i}, \quad Y = \sum_{i=1}^{r} y_i 2^{-i}$$

де $x_i, y_i \in \{-1,0,1\}$.

Для отримання n розрядів функції $Z^* = (X + Y)$ необхідно сформувати $m=r+p$ розрядів функції $Z = 2^{-p} \cdot (X + Y)$. Після виконання m кроків буде отримана функція з похибкою, яка за модулем не перевищує k^{-m-1}.

Алгоритм виконання i-го циклу ($i = 1,..., m$) зводиться до виконання наступних обчислень [2].

По формулі

$$H_i = 2R_{i-1} + 2^{-2} x_i + 2^{-2} y_i \tag{1}$$

знаходиться проміжна перемінна H_i (початковим є значення $R_0 = 0$).

Цифра результату формується за правилом

$$z_i = \begin{cases} -1, \text{ якщо } H_i < -2^{-1}; \\ 0, \text{ якщо } -2^{-1} \le H_i < 2^{-1}; \\ 1, \text{ якщо } H_i \ge 2^{-1}. \end{cases} \tag{2}$$

Для виконання наступного циклу обчислюється

$$R_i = H_i - z_i. \quad (3)$$

Розглянемо модифікацію алгоритму для додавання в неавтономному режимі K r-розрядних чисел $X_j (j = 1,...,K)$, тобто обчислення функції $Z = 2^{-p} \sum_{j=1}^{K} X_j$, де $X_j = \sum_{i=1}^{r} x_{ji} 2^{-i}$.

Якщо в кожному i-му циклі буде виконуватися співвідношення

$$Z_i - 2^{-i-1} \le 2^{-p} \sum_{j=1}^{K} X_j < Z_i + 2^{-i-1}, \quad (4)$$

то після виконання $m = n + p$ циклів буде одержано результат з похибкою, що за модулем не перевищує 2^{-m-1}.

Ввівши заміну

$$R_i = (2^{-p} \sum_{j=1}^{K} X_j - Z_i) 2^i, \quad (5)$$

вираз (4) можна подати у формі

$$-2^{-1} \le R_i < 2^{-1}. \quad (6)$$

Нерівність (6) виконується у вихідному стані при $R_0 = 0$. Будемо вважати, що на (i–1)-му кроці нерівність (6) також виконується, та визначимо при якому мінімальному значенні p це співвідношення буде виконуватись на будь-якому наступному кроці.

Виділимо в (5) значення R_{i-1} і зробимо заміну

$$H_i = 2R_{i-1} + \sum_{j=1}^{K} x_{ji}. \quad (7)$$

Тоді вираз (6) приймає вигляд

$$z_i - 2^{-1} \le H_i < z_i + 2^{-1}. \quad (8)$$

Аналізуючи вирази (6), (7), (8) та враховуючи допустимі значення цифр операндів, отримуємо значення затримки

$$p \ge 1 + \log_2 K, \quad (9)$$

при якій виконується умова (6) в будь-якому циклі, тобто процес обчислень буде збіжним. Звичайно, треба вибрати мінімальне ціле значення затримки.

Підставляючи в рівність (8) допустимі значення $z_i \in \{-1, 0, 1\}$, отримаємо правило формування цифри результату, яке співпадає з правилом (2) для двомісної функції додавання.

Отже, при додаванні K чисел в кожному циклі потрібно по формулі (7) знайти значення H_i, визначити цифру функції за допомогою правила (2), та знайти значення R_i для наступної ітерації за формулою (3).

Розглянемо апаратну реалізацію додавання в надлишкових системах числення.

Для побудови ВМ використовується звичайна двійкова апаратура. Цифри –1, 0, 1 кодуються двома двійковими розрядами, наприклад, відповідно 10, 00, 01 та пересилаються між ВМ по двом провідникам.

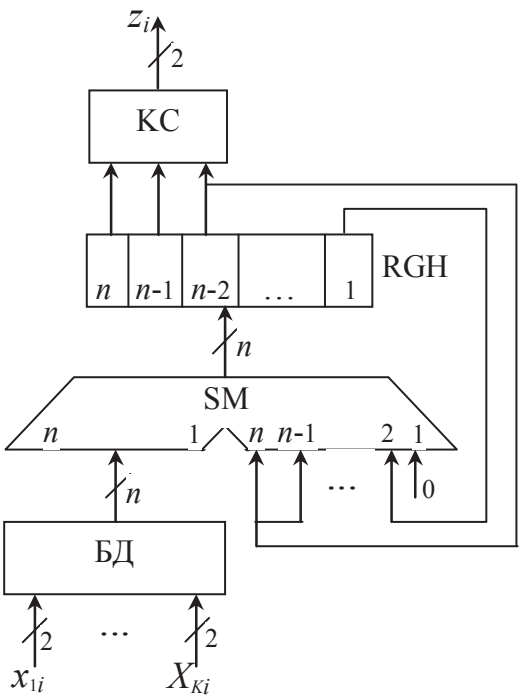


Рис. 1. Структура обчислювального модуля

Схема пристрою (рис. 1) містить регістр RGH для зберігання H_i. Розрядність регістра визначається формулою $n = 2 + p = 3 + \log_2 K$, де $\log K$ відповідає максимальній розрядності прирощення для H_i. Регістр має 2 знакових розряди. Значення H_i формується в доповняльному коді відповідно з формулою (7). Блок додавання (БД) формує суму i-х розрядів всіх операндів, що подаються в ВМ в i-му циклі. Кінцеве значення H_i формується на суматорі (SM). На рис. 2 для випадку $K = 2$ показано перетворення цифр операндів -1 та 1 в n-розрядний доповняльний код, що подається на вхід суматора.

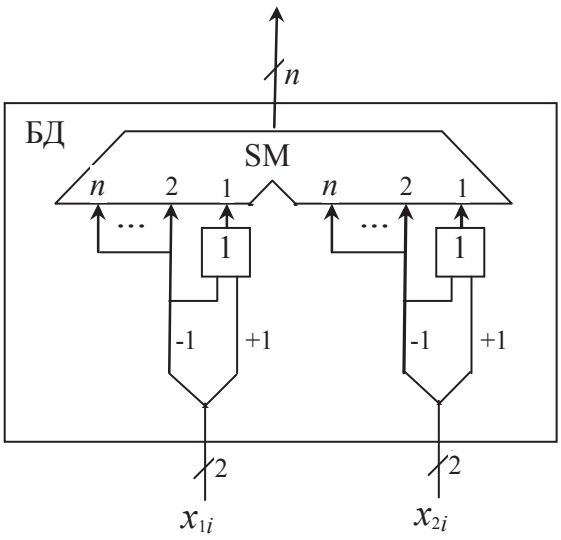


Рис. 2. Блок додавання

Значення розряду результату z_i визначається комбінаційною схемою (КС), яка аналізує три старші розряди коду H_i відповідно з таблицею істинності (табл. 1).

Таблиця 1.

Таблиця істинності функцій

H_i	**Функція z_i=-1**	**Функція z_i=1**
10,1… 11,0…	1	0
00,0… 11,1…	0	0
00,1… 01,0…	0	1
01,1… 10,0…	-	-

Віднімання z_i відповідно з формулою (3) замінюється просторовим функціональним зсувом на регістрі RGH. Як показано на рис. 1, вихід (n-2)-го розряду переписується водночас в n-й та (n-1)-й розряди регістру.

Розглянемо два варіанти реалізації додавання декількох чисел в неавтономному режимі.

В першому варіанті сума K чисел визначається за допомогою бінарного дерева ВМ, призначених для підсумування двох чисел. Дерево має $N=\log_2 K$ ярусів однакових ВМ, що працюють в неавтономному режимі. тобто ланцюжок ВМ, що відповідає максимальній затримці формування розрядів результату, також складає $N=\log_2 K$ модулів. Загальна кількість ВМ у бінарному дереві, як відомо, складає суму перших $N=\log_2 K$ членів геометричної прогресії із знаменником 2 та дорівнює $M = 2^{\log_2 K} - 1 = K - 1$.

В другому варіанті всі K чисел обробляються в одному ВМ, що реалізує K-місну функцію додавання.

Можна порівняти приблизно апаратні витрати на реалізацію двох варіантів додавання. В першому випадку M блоків БД в складі ВМ містять по одному суматору (рис.2), а в другому варіанті один БД містить дерево із M суматорів, тобто на БД витрачається приблизно однаковий склад апаратури.

Вся інша апаратура, що є в одному ВМ за другим варіантом, в системах за першим варіантом повторюється в кожному ВМ, тобто M разів. Слід зауважити, що розрядність суматорів і регістрів в різних ВМ може несуттєво відрізнятися, але це не впливає на висновок, що перший варіант програє другому за складністю.

Порівняємо час виконання обчислень в системах за двома варіантами у передбаченні, що різні ВМ працюють в складі системи в синхронному режимі, тобто цикл обчислень в різних ВМ однаковий. Такий режим є звичайним для ПСБЗ, що реалізовані на ПЛІС.

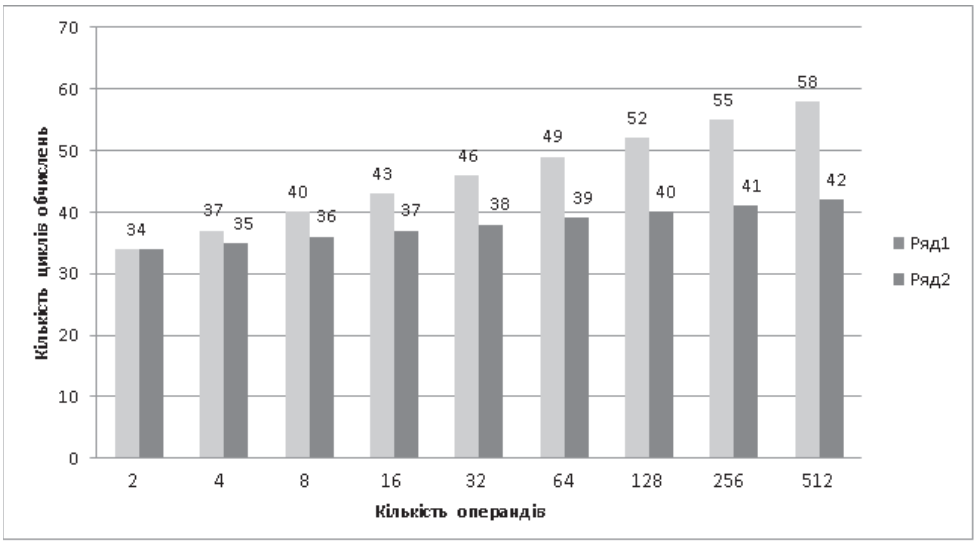


Рис. 3. Залежність кількості циклів обчислень від кількості операндів: ряд 1 – для використання двомісних функцій; ряд 2 – для K -місних функцій

Час виконання послідовності з M операцій в неавтономному режимі складає [4]

$$T=\left[m-1+M(p+1)\right]t_0, \tag{10}$$

де t_0 – тривалість циклу.

При $m=32$ для двомісної та K -місної функції додавання формула (10) приймає відповідно вигляд:

$$T=(31+3\log_2 K)t_0, T=(32+1+\log_2 K)t_0$$

З гістограми (рис. 3) видно, що для будь-якого $K>2$ використання багатомісної функції додавання забезпечує виграш в швидкодії відносно двомісних функцій.Виграш в швидкодії зростає із збільшенням місності функції додавання.

Висновки. Запропоновано алгоритм та реалізація багатомісної функції додавання чисел в неавтономному режимі з використанням надлишкової симетричної системи числення з цифрами {-1, 0, 1} та основою $k=2$. Операція виконується порозрядно зі старших розрядів із суміщенням процесів введення, обробки та виведення розрядів чисел. Це дозволяє частково поєднувати залежні операції в різних модулях, тобто виконувати ланцюжки залежних операцій в режимі суміщення на розрядному рівні.

Показано, що реалізація багатомісної функції додавання дозволяє прискорити обчислення та зменшити апаратурні витрати. Обмін даними між вузлами системи послідовним кодом зменшує необхідні функціональні та комутаційні ресурси ПЛІС (функціональні комірки та комірки введення-виведення), що, в свою чергу, підвищує надійність систем.

Список використаних джерел

1. Zhabin V.I. Design of High-Speed Specialized Computers for the Realization of Many-Place Expressions / V.I.Zhabin, V.I.Korneichuk, V.P.Tarasenko // Automatic Control and Computer Sciences. – 1981. – vol. 15, no 6. – P. 15-18.

2. Zhabin V.I. Computation of Rational Functions for Many Arguments / V.I.Zhabin, V.I.Korneichuk, V.P.Tarasenko // Automation and Remote Control. – 1978. – vol. 38, no 12. – P. 1864-1871.

3. Дичка И.А. Совмещение зависимых операций на уровне обработки разрядов операндов / И.А.Дичка, В.В.Жабина // Искусственный интеллект. – 2008. – №3. – С. 649-654.

4. Жабин В.И. Выполнение последовательностей зависимых операций в режиме совмещения / В.И.Жабин // Вісник Національного технічного університету України "КПІ". „Інформатика, управління та обчислювальна техніка". – 2007. – №46. – С. 226-233

5. Максфилд К. Проектирование на ПЛИС. Архитектура, средства и методы / К. Максфилд. – М.: Издательский дом «Додэка-XXI», 2007. – 408 с.

6. Жабин В.И. Реализация неавтономных вычислений в избыточных системах счисления на ПЛИС / В.И. Жабин, В.В. Жабіна, А.В. Скоріченко // Вісник НТУУ "КПІ". Інформатика, управління та обчислювальна техніка. – 2016. – №64. – С. 150-155.

ДОВІДКА ПРО АВТОРІВ

Жабін Валерій Іванович – професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Valerii Zhabin – professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: viz.kpi@gmail.com

Токар Андрій Геннадійович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Andrii Tokar – student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: andrey-naum@ukr.net

Zhabin V.I.,
Tokar A.H.

ANALYSIS OF EFFICIENCY OF ADDITION OPERATION REALIZATION IN ON-LINE MODE

Relevance of the research topic. The solution of complex tasks of information processing in real time determines the use of parallel computing systems. In such systems the time of solving problems depends not only on the time of execution of parallel branches of algorithms, but also on the time spent on the exchange of information between computing modules. Reducing the time spent on data exchange is an urgent task.

Problem statement. When building systems based on FPGA, an economical use of the microcircuit resource (internal functional blocks and external pins) is important in order to reduce the number of chips for the implementation of the computing system. The paper considers the possibility of solving this problem by reducing the number of connections between components of the system.

Analysis of recent researches and publications. Performing operations using parallel arithmetic requires data transfer between blocks with multi-bit codes. This does not save the switching elements of the FPGA. To reduce the number of links, quasi-parallel arithmetic methods have been developed. Calculations in this case in each block are performed with bitwise introduction of operands. To exclude transfers to the upper digits, redundant number systems are used. When performing chains dependent on data operations, it becomes possible to overlap operations at the processing level of the operand bits. This is shown in a number of works. **Singling out of unexplored parts of the general problem**. When performing the chain of dependent operations, each module works in on-line mode. With the increase in the length of the chain of operations, the processing time increases. To reduce the number of operations in a chain, it is advisable to use modules that perform multiple operations.

Task definition. In the present paper, the efficiency of realizing a multi-operand function in the case of redundant data representation is shown.

Statement of the main material. The algorithm and hardware implementation of the addition function in on-line mode using redundant number systems is considered. A comparative analysis of the speed and complexity of devices with different number of operands is performed. The result of comparison showed greater efficiency of devices with a large number of operands.

Key words: on-line computations, redundant number system, dependent operations overlap

Секція 3. AI (Інтелектуальні системи. Машинне навчання, bigdata). Section 3. AI (Intelligent Systems. Machine learning, big data.).

УДК 004.93'12

Юрій Тимошин, Сергій Орленко

ВПЛИВ ПЛЯМИ СПРИЙНЯТТЯ НА ШВИДКІСТЬ ТА ТОЧНІСТЬ РОЗПІЗНАВАННЯ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ

INFLUENCE OF THE RECEPTIVE FIELD ON THE SPEED AND ACCURACY OF RECOGNITION OF THE CONVOLUTIONAL NEURAL NETWORK

У статті розглянута проблема плями сприйняття та способу завдання її розмірів за допомогою зміни топології загорткової нейронної мережі. Наглядно були показані способи досягнення одного розміру плями сприйняття різними способами, виконаний ряд досліджень та розроблені програми для наборів даних MNIST та Sifar, що відображають досягнуту точність розпізнавання та затрачений час на навчання.

Ключові слова:згорткові нейронні мережі, пляма сприйняття, точність, час навчання.

Рис.: 6. Бібл.: 3.

The paper deals with the receptive field problems and methods of its definition with the help of changing topology of convolutional neural network. With examplesshown ways how to achieve the same receptive field in different ways, also performed a number of studies and developed programs for datasets, MNIST and Cifar, which reflects the achieved recognition accuracy and time spent on training.

Key words: convolutional neural network, receptivefield, accuracy, time of study.

Fig.: 6. Bibl.: 3.

Вступ. При побудові системи розпізнавання образів заснованої на штучній нейронній мережі (НМ) вибір типу мережі найчастіше вирішений наперед – це згорткова нейронна мережа (ЗНМ). Запропонована в 1988 році Яном Лекуном вона розроблена на подобі біологічного зору і зорової кори та складається з простих і складних клітин, викликала справжній фурор і відродила інтерес до НМ.

Назва мережа отримала через наявність операцій згортки. ЗНМ має деяку специфіку – вона полягає у перестановці загорткових шарів і шарів субдискретизації (пулінг), що і відсилає її до біологічного зору. Так у людини є прості клітини, які реагують на зміни градієнта, такі як кути і лінії в певній точці – аналог згортки; складні ж клітини інваріантні до зсувів і реагують при активації набору простих клітин – аналог пулінг [1]. На більш пізніх шарах ядра згортки поступово змінюються і формують певне семантичне значення. Оскільки структура мережі багатошарова – однонаправлена і закінчується вона декількома повнозв’язними шарами, втрачається просторова структура пікселів [2]. Хоча

саме процес вилучення абстрактних понять з вихідного зображення і забезпечує ЗНМ таку високу точність.

Постановка задачі. Як уже було згадано, виділення семантичних ознак на пізніх етапах є результатом ряду згорток і субдискретизацій, однак слід розуміти, що всі ці операції проводяться над певною початкової області – плямою сприйняття (receptivefield) [3]. В цілому для ЗНМ вхідна картинка і є зображенням що сприймається, однак не все так просто. Кожна операція (або набір операцій) згортки перед пулінгом захоплює лише частину зображення і це визначає подальшу обробку.

Дане питання стосується архітектури ЗНМ, і оскільки її вибирає розробник, виходячи з кращих прикладів, розглянемо операції, що стосуються перетворення плями сприйняття. Також слід вказати про важливість зазначених нижче методів – всі вони зменшують число параметрів НМ, що покликане знизити обсяг обчислень і прискорити навчання/перенавчання ЗНМ. Завдання полягає у відстеженні отриманої точності і часу навчання при кожному підході.

Підходи до зміни плями сприйняття та число параметрів. До таких підходів, як не дивно, відноситься і пулінг. «Навіщо потрібен пулінг, якщо є фільтр побільше?» – таке питання може поставити будь-який уважний інженер. Субдискретізація, подібно до великого фільтру, покликана збільшити пляму сприйняття (хоча при цьому втрачається локалізація) і при цьому посилити виділення шуканої особливості. Так при пулінгу 2 на 2 пляма сприйняття збільшується в 2 рази.

Припустимо, у нас є зображення розміру 8x8 і ми хочемо отримати з нього якісь особливі властивості. І тут постає вибір – застосувати великий фільтр або фільтр поменше і операцію пулінгу. Припустимо, зображення було піддано згортці з ядром 3x3 і для отриманої карти ознак 6x6 ми застосували субдискретізацію вікном 2x2. В результаті отримали нову карту ознак розміром 3x3, яка містить найосновніші значення особливих властивостей і може використовуватися і далі. Наприклад, взяти ще одну згортку з ядром три на три, вийде послідовність (3x3 | pooling | 3x3), а результат буде одиничним значенням (Рис. 1. *а*). Тепер уявімо, що ми використовували фільтр розміром 8x8 (Рис. 1. *б*) – отримали таке ж єдине значення як і після двох попередніх перетворень.

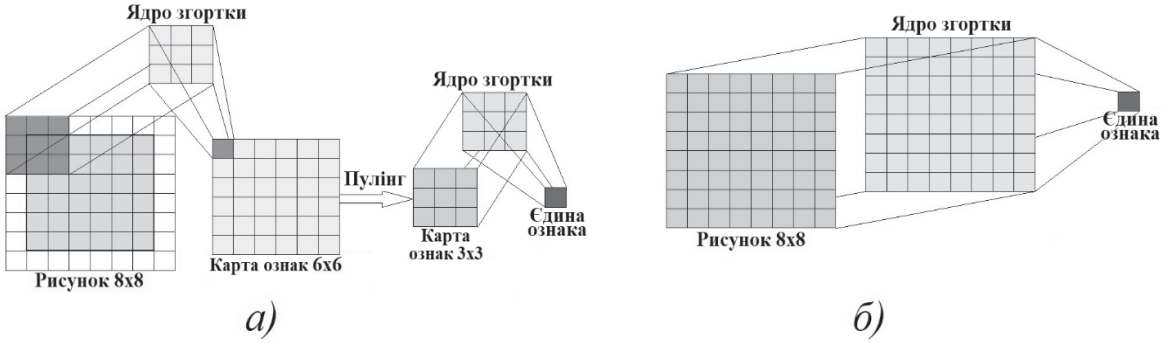


Рис. 1. Використання пулінга замість великого фільтра

Виходить, що при першому підході можна отримати узагальнений результат, який зручно використовувати і далі, а при другому – лише єдине значення.

Тепер розглянемо кількість параметрів. У першому підході при першій згортці використовувалися *3×3+1* параметрів, а для отримання одиничного значення ще *3×3+1* параметрів – разом 20, в той час при другому: *8×8+1=65*.

Однак, при очевидному зменшенні параметрів спостерігається зворотна картина з обчисленнями при прямому проходженні через мережу – для даного рішення кількість арифметичних операцій для другого підходу в 5.2 рази менше ніж для першого.

Наступний підхід до вибору плями сприйняття вельми простий. Він полягає в послідовному використанні двох або більше невеликих фільтрів замість одного великого при збереженні розмірів початкової області. Розглянемо на простих прикладах.

Візьмемо фільтр розміру 5x5 і застосуємо його до зображення 7x7 пікселів – результатом буде карта ознак розмірністю 3x3 (Рис. 2. *а*). Досягти збереження плями сприйняття при заміні початкового ядра згортки можна за допомогою застосування поспіль двох фільтрів розмірністю 3x3 (Рис. 2. *б*).

Кількість параметрів при одному фільтрі дорівнює: *5×5+1=26*, при двох маленьких: *3×3+1+3×3+1=20*; кількість арифметичних операцій для обчислення однієї результуючої карти ознак 50 і 180 відповідно.

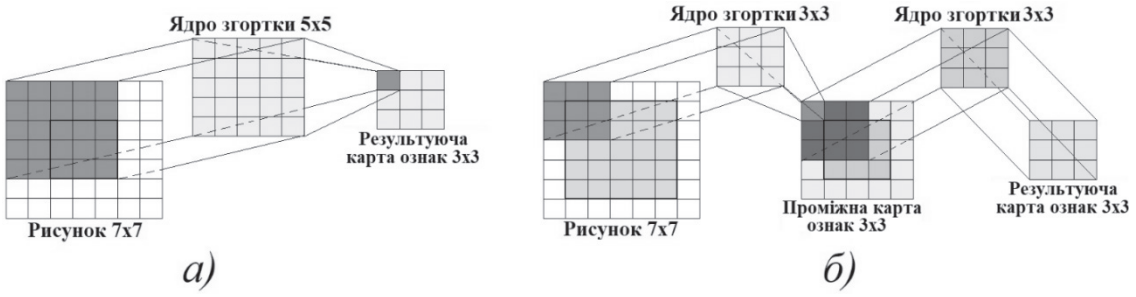


Рис. 2. Заміна фільтра 5x5 на комбінацію двох фільтрів 3x3

Подібну заміну можна провести і при початковому розмірі фільтра 7x7 за допомогою послідовного застосування вже трьох фільтрів розмірністю 3x3. І для такої заміни у одного фільтра буде 50 параметрів та 98 арифметичних операцій, а у комбінації трьох фільтрів – 30 параметрів, але 630 операцій.

Також зазначимо можливість заміни ядра 5x5 послідовним використанням чотирьох односпрямованих векторів: 5x5≈1x3→3x1→1x3→3x1. Ілюстрація даної заміни показана на рис. 3.

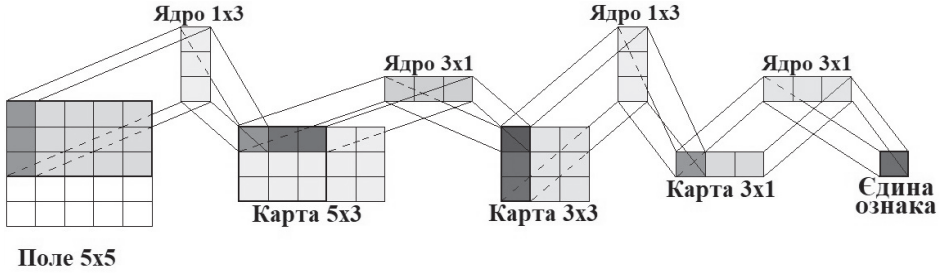


Рис. 3. Заміна фільтра 5x5 на набір одно направлених фільтрів

Властивості фільтра 5x5 залишилися незмінними: 26 параметрів і 50 операцій, для набору фільтрів кількість параметрів складає: *(3+1)×4=16*; кількість арифметичних операцій для обчислення останньої ознаки дорівнює 150.

Експериментально-дослідне дослідження задачі. Для перевірки ефективності наведених вище методів маніпуляції зоною сприйняття проведемо ряд експериментів. Почнемо з початкової моделі: в її якості була обрана мережа, яку легко можна було б переналаштовувати для двох навчальних вибірок. Так, до вхідного зображення послідовно застосовувалися: згортка 3x3 і пулінг, глибина карт активації – 16, згортка 3x3 і пулінг, глибина – 32 і наостанок згортка 3x3 з глибиною 64; для всіх згорток функція активації ReLU, а вікно пулінгу – 2x2. Далі дані по мережі передавалися трьом повнозв'язним шарам (функції активації перших двох шарів – тангенс гіперболічний, а третьої – softmax) по 256, 64 і 10 нейронів відповідно. ЗНМ навчалася за допомогою методу ADAM (adaptive-emomentestimation) з початковим темпом навчання 0.001. Так як навчальні вибірки були різних розмірів – змінювалося число параметрів, що надходять на повнозв'язні шари мережі, схожа картина спостерігалася при зміні вікна сприйняття, але оскільки ЗНМ була реалізована за допомогою бібліотеки PyTorch мови Python – мережу було легко і зручно перебудовувати.

Почнемо з збільшення першої згортки до розміру 5x5. А потім замінимо її послідовним використанням двох згорток 3x3 шляхом додавання ще одного шару – Рис. 4 та Рис. 5. Аналогічну заміну проведемо для фільтра 7x7 двома 5x5, паралельно перевіривши ефективність пулінгу: прибравши його після другого згрорткового шару – Рис. 6.

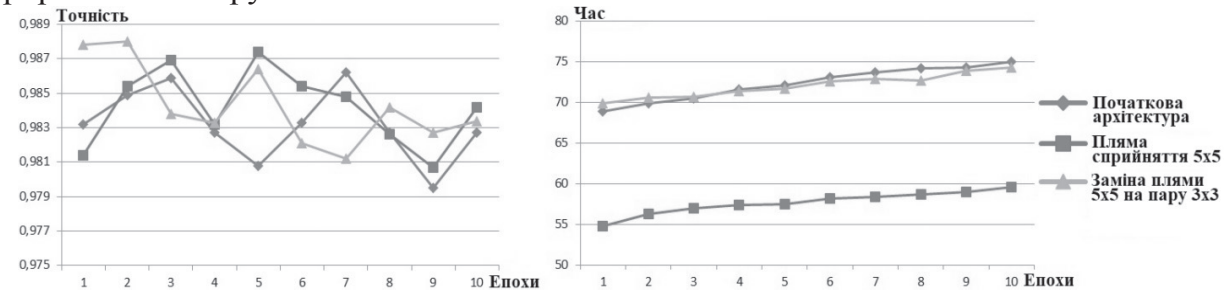


Рис. 4. Дослідження плями сприйняття 5x5 для набору MNIST

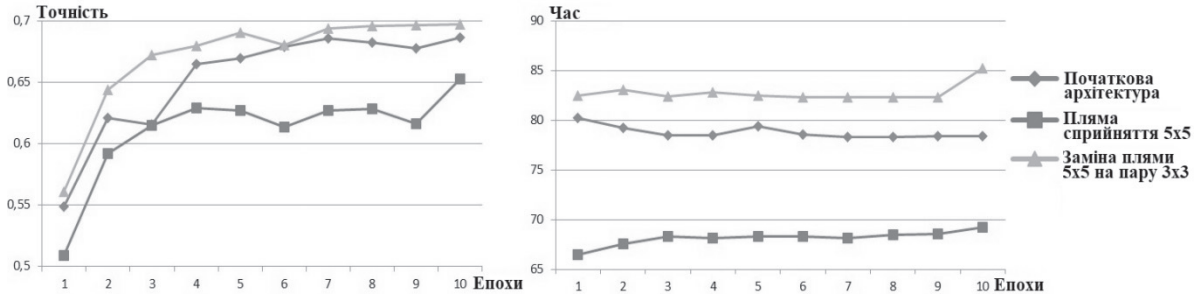


Рис. 5. Дослідження плями сприйняття 5x5 для набору CIFAR10

На обох графіках простежується чітка закономірність – просте збільшення плями сприйняття призводить до збільшення швидкості навчання ЗНМ завдяки зменшенню числа арифметичних операцій, проте на прикладі навчальної вибірки CIFAR10 видно, що при такому підході точність розпізнавання падає. Зворотна ситуація відбувається при збільшенні плями сприйняття шляхом заміщення великого фільтра двома маленькими шляхом додавання ще одного згорткового шару. Так при загальному зменшенні числа параметрів в порівнянні

з великим фільтром все ж вдається виділити більше ознак, що сприяє підвищенню точності розпізнавання, однак час витрачений на навчання зростає в зв'язку зі збільшенням обчислень.

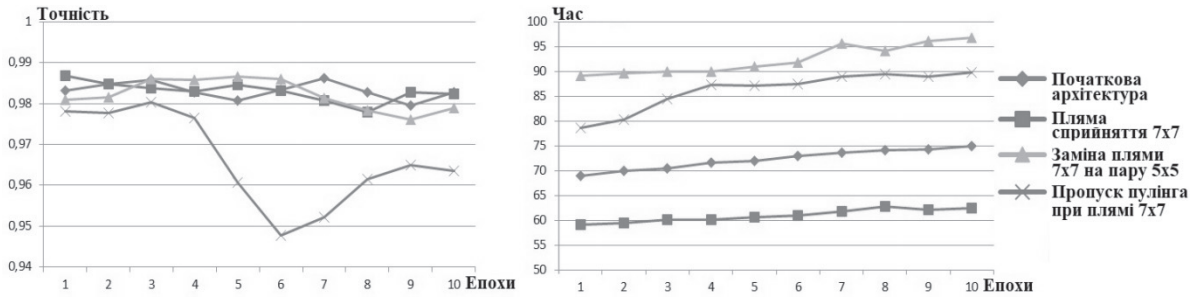


Рис. 6. Дослідження плями сприйняття 7x7 набір MNIST і пропуск шару пулінгу

Наведені вище результати підтверджуютьсяі для вікна розміром 7x7, хоча не так очевидно зменшення точності. Пропуск субдискретізації призводить до істотного збільшення арифметичних операцій і знижує здатність ЗНМ виділяти семантичні ознаки, що і призводить до зниження точності.

Висновки. Отриманий результат дозволяє стверджувати, що вибір плями сприйняття є важливим моментом при проектуванні ЗНМ, а вибір підходу завдання плями безпосередньо впливає на точність і час навчання мережі. Збільшення шарів, як правило, призводить до збільшення точності, для цієї ж мети слід намагатися підібрати оптимальний метод навчання, функції активації мережі і початковий коефіцієнт навчання. Всі ці підходи здатні вплинути на точність і час навчання ЗНМ, тому бажано завжди тестувати і порівнювати спроектовані топології ще й на практиці.

Список використаних джерел

1. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. HubbardandL. D. Jackel: BackpropagationAppliedtoHandwrittenZipCodeRecognition, NeuralComputation, 1(4):541-551, +Winter 1989.

2. Krizhevsky, Alex, and Geoffrey Hinton. Learning multiple layers of features from tiny images [Електронний ресурс] – 2009.– С.60. – Режим доступу: http://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf

3. K. Fukushima, Neocognitron for handwritten digit recognition. Neurocomputing – April 2003 – (Vol. 51, pp. 161-180).

ДОВІДКА ПРО АВТОРІВ

Тимошин Юрій Афанасійович – к.т.н., доцент кафедри технічної кібернетики НТУУ “КПІ ім. Ігоря Сікорського”.

Наукові інтереси: інформаційно-управляючі системи, обробка інформації, складні об'єктів управління, телекомунікаційні системи та бази даних.

Timoshin Yuri – Ph.D., Associate Professor, Department of Technical Cybernetics NTUU "Igor Sikorsky Kyiv Polytechnic Institute".

Scientific interests: information-operating system, information processing, complex facilities management, telecommunications systems and databases.

E-mail: y.timoshin@gmail.com

Орленко Сергій Петрович – студент кафедри технічної кібернетики НТУУ“КПІ ім. Ігоря Сікорського ”

Наукові інтереси: теорія цифрової обробки сигналів та зображень, розпізнавання образів, штучний інтелект

Orlenko Sergey – student of the Department of Technical Cybernetics NTUU "Igor Sikorsky Kyiv Polytechnic Institute".

Scientific interests: theory of digital signal processing and image pattern recognition, artificial intelligence.

E-mail: orlenko_sergey@ukr.net

YuriTimoshyn, SergeyOrlenko

INFLUENCE OF THE RECEPTIVE FIELD ON THE SPEED AND ACCURACY OF RECOGNITION OF THE CONVOLUTIONAL NEURAL NETWORK

Introduction. The convolution neural network was developed by Yann LeCun in 1988 in likeness of biological vision and visual cortex and consists of simple and complex cells. The network has some specificity - it consists in rearranging the convolutional layers and layers of sub-sampling (pooling), which sends it to the biological view.

Formulation of the problem. The separation of semantic signs occurs at later stages, but it should be understood that all these operations are carried out over a certain initial region - a receptive field. In general, there is an image that is perceived, but it all is not so simple. Each convolutional operation before the pooling captures only part of the image and this determines the further processing. This question relates to the architecture of the convolution neural network so we will consider the operations concerning the transformation of the receptive field. The task will be to track the obtained accuracy and time of training with each approach.

Approaches to change receptive field and number of parameters. We considered pooling from the point of increasing the absorption of signs from the larger field, considered the possibility of replacing the 5x5 filter with two 3x3 or by a sequence of 1x3 → 3x1 → 1x3 → 3x1, and replacing 7x7 with two 5x5. We have calculated the number of parameters and arithmetic operations for each approach, also created illustration for them.

Experimental-research study of the problem. We built and changed the convolution neural network according to the methods described above. As the data set we were used traditional MNIST and CIFAR10. In parallel, we recorded the resulting accuracy and elapsed training time and depicted them on the corresponding graphs.

Conclusions. Choice of receptive field is an important point in the design of CN, and the choice of approach to the problem of receptive field directly affects the accuracy and time of learning the network. Increasing the layers, as a rule, leads to increased accuracy, for the same purpose should try to choose the best method of training, network activation functions and the initial learning factor. So it is advisable to always test and compare projected topologies even in practice.

Key words: convolutional neural network, receptivefield, convolution, topology,parameter, accuracy, time of study.

УДК 004.93'12

Пилипенко Павло,
Пантелєєв Олександр,
Олійник Володимир

МЕТОД ВИРІШЕННЯ ЗАДАЧІ РОЗПІЗНАВАННЯ ЛЮДЕЙ В РЕАЛЬНОМУ ЧАСІ В УМОВАХ УСКЛАДНЕНОЇ ВИДИМОСТІ

METHOD OF SOLVING THE PROBLEM OF PEOPLE RECOGNITION AT REAL TIME IN CONDITIONS OF OCCLUDED VISIBILITY

У статті пропонується метод вирішення задачі розпізнавання людей в реальному часі в умовах ускладненої видимості. Для детектування використовується комбінація детекторів людини та обличчя. Для моделювання обличь використовується метод ЛБШ та розпізнавання класифікатором SVM. Для підтвердження розпізнавань використовується трекер MCNF.

Ключові слова: розпізнавання обличчя, детектор людей, підтвердження через відстеження.

Рис.: 2. Бібл.: 3.

This paper proposes a method of solving the problem of recognizing people in real time in conditions of occluded visibility. A combination of human and facial detectors used for detecting. For facial modeling, the LBP method and the SVM classifier recognition are used. MCNF tracker is used to confirm the recognition.

Key words: face recognition, human detector, confirmation through tracking.

Fig.: 2. Bibl.: 3.

Актуальність теми дослідження. Проблема розпізнавання людей стає більш актуальною в останні дні у зв'язку зі зростаючим попитом на автоматичні системи контролю доступу, системи безпеки та системи відеоспостереження тощо. Таким чином, виникає необхідність системи, що надійно та швидко розпізнає людей навіть при поганих умовах видимості. Дана робота присвячена проблемі розпізнавання людей в реальному часі в умовах ускладненої видимості.

Аналіз останніх досліджень і публікацій. Протягом останніх років з’являється все більше статей присвячених розпізнаванню людей у різних середовищах, зокрема, завдяки появі нових методів знаходження та розпізнавання. Проте підходи з використанням відстеження як засобу підтвердження все ще недостатньо вивчені.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена вивченню та аналізу запропонованого підходу для розпізнавання людей в реальному часі в умовах ускладненої видимості. Дослідження сфокусовано на вивченні комбінації детекторів людини та обличчя та підтвердження через відстеження.

Постановка завдання. Завданням є створити модель з базою даних обличь, що може у реальному часі знаходити людей на відео та зіставляти їх з базою даних чи позначати як невідомих та зберігати розпізнавання при відсутності людини у прямої видимості.

Викладення основного матеріалу. Метод рішення можна розділити на 3 етапи. На першому етапі проводиться виявлення об'єктів, схожих на обличчя, на кожному кадрі заданій послідовності. На другому етапі виділяються унікальні ознаки виявлених осіб і порівнюються з особами з бази даних.

Фінальним етапом є побудова траєкторії виявленого особи для підтвердження розпізнавань при відсутності виявлення на одному або декількох кадрах.

Детектування обличь. Для виявлення обличь ми пропонуємо використовувати зв'язку детектора людей і детектора обличь. Детектор людей заснований на ймовірнісної мапі зайнятості. Цей детектор вимагає початкового встановлення площини поверхні у вигляді дискредитованої сітки. Після виявлення детектором людей людини в якійсь клітинці сітки вираховується відстань від поверхні до голови людини і створюється зона пошуку обличчя (ЗПО), на якій проводиться пошук детектором обличь. Це дозволяє значно зменшити зону пошуку і підвищити швидкість виявлення, а також знижує кількість помилкових розпізнавань.

Детектор обличь заснований на методі бінарних ознак яскравості. Цей метод схожий з методом Віоли і Джонса, але замість пошуку ознак на прямокутних областях, цей метод порівнює окремі значення пікселів. Це дозволяє використовувати прості функції, що дає приріст до швидкості.

Для усунення помилкових спрацьовувань виконуються наступні кроки:

1. Шукати тільки в області, що цікавить (ЗПО) навколо очікуваної позиції голови

2. Відхилити виявлення в межах ЗПО, які більше / менше очікуваного розміру голови

3. Обмежити максимальну кількість виявлень в межах ЗПО до одного

Моделювання та розпізнавання. Для створення моделей обличь ми пропонуємо метод гістограм локальних бінарних шаблонів, який був обраний з-за його високої швидкості і надійності щодо пози обличчя і освітленості. Інформація про дані обличчя зберігається у вигляді гістограми, яка представляє вектор ознак ЛБШ.

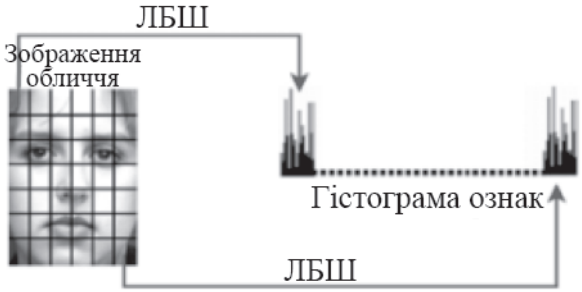


Рис. 1. Дескриптор ознак ЛБШ.

Отримавши невідомий вектор ознак ЛБШ, будь-який алгоритм машинного навчання може бути використаний для порівняння його з векторами в базі даних. Ми пропонуємо використовувати многокласовий SVM. Ознаки гістограми ЛБШ не схильні до змін освітлення і виразів обличчя, а класифікатор SVM навчається на обличчях на зображеннях з різними позами голови, роблячи остаточний класифікатор також стійким до пози голови.

Підтвердження через відстеження. Останнім етапом є інтеграція розпізнавання обличь з трекером людей. Відстеження дозволить знаходити траєкторії руху кожної людини зі збереженням особистості, що дозволить підтвердити або спростувати наявність людини в кадрі при відсутності прямої видимості.

Ми пропонуємо використовувати трекер багатоточечної транспортної мережі (MCNF). Площина сцени представлена у вигляді ймовірнісної карти зайнятості, що містить ймовірності присутності індивідів у кожному осередку сітки. Траєкторії людей моделюються як безперервні потоки, що проходять через осередки сітки. Особистості присвоюються кожної траєкторії за допомогою даних зовнішнього вигляду, різні потоки не повинні змішуватися. Кожен вузол представляє просторове розташування в певному кадрі для конкретної групи ідентичностей. Грані між вузлами представляють допустимий рух між місцями розташування. Оскільки кількість відслідковуються людей може змінюватися з часом, введені 2 віртуальних вузла ($vsource$ і $vsink$), які пов'язані з усіма точками, де людина може увійти або вийти з сцени.

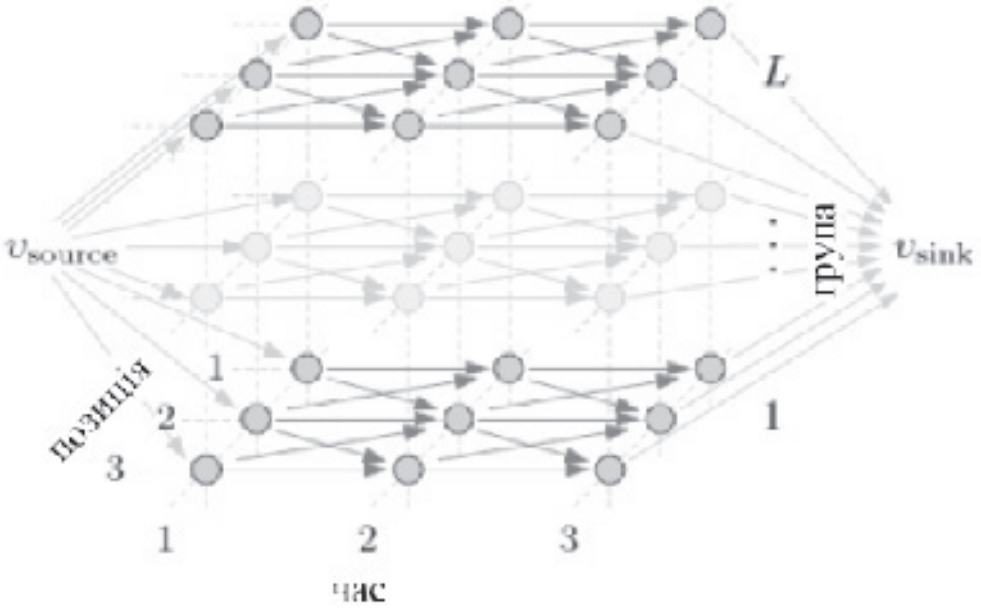


Рис. 2. Граф відстеження MCNF.

Експерименти.Тестування було проведено на відеоданих з набору Terrace. Відео складається з 5250 кадрів, знято зі швидкістю 25 кадрів в секунду в дозволі 720x576 пікселів. Відео містить 21 людини, дані про 18 збережені в базі даних. Швидкість виявлення склала 85 кадрів в секунду, точність розпізнавання досягла 99.3833%.

Висновки. В роботі запропоновано підхід до вирішення задачі розпізнавання людей в реальному часі в умовах утрудненою видимості, який дозволяє підвищити швидкість, точність і надійність розпізнавання людей. Це досягається за рахунок застосування зв'язки детектора людей і детектора осіб, а також введення підтвердження розпізнавань на основі відстеження трекером багатоточечної мережі. Це дозволило досягти продуктивності, достатній людних сценах і на сценах з частковим перекриттям видимості.

Список використаних джерел

1. P. Viola and M. Jones, Robust real-time face detection, International Journal of Computer Vision, vol. 57, no. 2, pp. 137-154, 2004.

2. T. Ojala, M. Pietikainen and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, 2002.

3. H. Ben Shitrit, J. Berclaz, F. Fleuret and P. Fua, Tracking multiple people under global appearance constraints, in IEEE International Conference onComputer Vision (ICCV), Barcelona, 2011.

ДОВІДКА ПРО АВТОРІВ

Олійник Володимир Валентинович – к.т.н., старший викладач, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Oliynyk Volodymyr – Ph.D., associate professor, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: oliinyk.volodymyr@gmail.com

Пилипенко Павло Геннадійович – студент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Pylypenko Pavlo –student, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: pylypenko.pg@gmail.com

Пантелєєв Олександр Сергійович – студент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Pantelieiev Oleksandr –student, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: spn220v@gmail.com

Pylypenko Pavlo,
Pantelieiev Oleksandr,
Oliynyk Volodymyr

METHOD OF SOLVING THE PROBLEM OF PEOPLE RECOGNITION AT REAL TIME IN CONDITIONS OF OCCLUDED VISIBILITY

Target setting. The problem of recognizing people is becoming more relevant in recent days due to the increasing demand for automatic access control systems, security systems and video surveillance systems, etc. Thus, there is a need for a system that reliably and quickly recognizes people even in poor visibility conditions. This work is devoted to the problem of recognizing people in real time in conditions of complicated visibility.

Actual scientific researches and issues analysis. In recent years, there are more articles devoted to recognizing people in different environments, in particular, due to the emergence of new methods for finding and recognizing. However, approaches using tracking as a means of verification are still underdeveloped.

Uninvestigated parts of general matters defining. This article is devoted to the study and analysis of the proposed approach for recognizing people in real time in conditions of complicated visibility. The study focuses on the study of a combination of human and person detectors and confirmation through tracking.

The research objective. The objective is to create a model from a database of faces that can live to find people on the videos and compare them with a database or marked as unknown and store in the absence of recognition of rights in direct line of sight.

The statement of basic materials. The analysis of joint use of human and person detectors is carried out. Descriptions of approaches for face modeling, facial recognition, tracking and building trajectories of people movement and confirmation of recognition through tracking are described. The results were very accurate and reliable, and the speed was sufficient.

Conclusions. The content, structure and parameters of the models that showed the best results of the recognition were analyzed. The approach proved to be good for recognition in conditions of complicated visibility. The results of experiments are presented and the analysis of the following steps is presented.

Key words: face recognition, human detector, confirmation through tracking.

UDC 004.8

Дмитро Смішний,
Олексій Алещенко

РОЗПІЗНАВАННЯ МОВИ З ВИКОРИСТАННЯМ ТЕХНОЛОГІЇ ГЛИБОКОГО НАВЧАННЯ

SPEECH RECOGNITION WITH DEEP LEARNING TECHNOLOGY USING

В даній статті розглядається проблема розпізнавання мови користувача за допомогою технології глибокого навчання. Для вирішення поставленої задачі було також застосовано підхід з рекуретними нейронними мережами та перетворенням Фур'є у ролі математичної моделі. Текст диктується через мікрофон BehringerC-1U. Словом для навчання мережі виступає «Привіт».

Ключові слова: рекурентні нейронні мережі, глибоке навчання, перетворення Фур'є, дискретизація.

Рис.: 4. Бібл.: 5.

This article examines the problem of recognizing speech using deep learning technology. I was also used an approach with recurrent neural networks and Fourier transformation as a mathematical model to solve the problem. The text is dictated by the Behringer C-1U microphone. The word for training the network is Ukrainian hello – "pryvit".

Key words: recurrent neural networks, deep learning, Fourier transformation, sampling.

Fig.: 4.Bibl.: 5.

Актуальність теми дослідження.В нашому інформаційному світі все частіше постає проблема комунікації з тими, чи іншими пристроями одночасно. Оскільки для кожного пристрою потрібно придумувати власний користувацький інтерфейс. Куди як зручніше, для користувача, використовувати голосові команди для керування різноманітними девайсами. Адже це позбавляє додаткових витрат для розробки того ж користувацького інтерфейсу. Та й інструкції до таких пристроїв будуть набагато зрозуміліші і простіші. Ну а використання глибокого навчання, в даному випадку, здешевлює і уніфікує розробку технології голосового керування для різних мов і країн. Дана стаття присвячена проблемі голосового керування саме українською, адже, на жаль, таких систем саме для нашої мови поки дуже мало.

Постановка проблеми. Практично повна відсутність систем розпізнавання голосу для українського сегменту користувачів Інтернету, створених на основі глибокого навчання. А рішення, які вже готові, недосконалі.

Аналіз останніх досліджень і публікацій. За останні роки можна побачити значне пожвавлення розвитку систем розпізнавання голосу. Серед найвідоміших прикладів це AmazonAlexa, GoogleNow, Siri, Алиса від Yandex. Всі вони засновані на нейронних мережах. Однак, варто зауважити, що вони або не підтримують українську, або підтримка даної мови вкрай обмежена. Тому в статті розглядається специфіка розпізнавання тексту українською.

Виділення недосліджених частин загальної проблеми. В даній статті освітлена проблематика розпізнавання голосу, якщо людина говорить українсь-

кою. При цьому, використовується технологія глибокого навчання. Дослідження сфокусовані на генерації тексту на основі голосу, як такого, та застосування даної технології по відношенню до української мови.

Постановка завдання. Завданням даної статті є створення моделі, що буде працювати з українськими словами (в даному випадку, слово «привіт»), розпізнавати їх, та видавати правдивий текст відповідно до надиктованого.

Викладення основного матеріалу.Голос, як і будь-який звук, представляє собою неперервну хвилю з різною амплітудою в різні моменти часу. Оскільки, сучасний комп'ютер - машина цифрова, а не аналогова, то йому для розуміння інформації, що передається голосом, потрібне представлення останньої у вигляді цифрового сигналу. Для цього використовують дискретизацію. В загальному, це технологія, яка дозволяє отримувати данні в певний момент часу з певною частотою [1]. Для людської мови достатньо частоти дискретизації в 16 кГц (16000 разів за секунду отримуємо значення амплітуди). Тепер оцифруємо наше«Привіт» 16000 разів на секунду. Перші 100 точок наведені на рисунку:

```
[-1274, -1252, -1160, -986, -792, -692, -614, -429, -286, -134, -57, -41, -169, -456, -450, -541, -761, -1067, -1231, -1047, -952, -645, -489, -448
, -397, -212, 193, 114, -17, -110, 128, 261, 198, 390, 461, 772, 948, 1451, 1974, 2624, 3793, 4968, 5939, 6057, 6581, 7302, 7640, 7223, 6119, 5461,
 4820, 4353, 3611, 2740, 2004, 1349, 1178, 1085, 901, 301, -262, -499, -488, -707, -1406, -1997, -2377, -2494, -2605, -2675, -2627, -2500, -2148, -
1648, -970, -364, 13, 260, 494, 788, 1011, 938, 717, 507, 323, 324, 325, 350, 103, -113, 64, 176, 93, -249, -461, -606, -909, -1159, -1307, -1544]
```

Рис.1. Кожне число показує амплітуду звукової хвилі
як 1/16000-ої від секундного інтервалу

Можна припустити, що внаслідок такої дискретизації втрачається частина інформації. Тому, для більш точного відображення звуку скористаємося теоремою відліків Віттакера — Найквіста — Котельникова — Шеннона:

$$x(t) = \sum_{k=-\infty}^{\infty} x(k\Delta) sinc\left[\frac{\pi}{\Delta}(t - k\Delta)\right], \text{де } sinc(x) = \frac{\sin(x)}{x}, \quad (1)$$

$$\text{а інтервал дискретизації } 0 < \Delta \leq \frac{1}{2f_c}.$$

Таким чином, ми отримаємо результат, зображений на рисунку 2.

Для полегшення роботи нейронної мережі, ми розділили запис на 20 мілісекундні інтервали, і навчатимемо мережу на них.

Ми розкладаємо звукову хвилю на простіші звукові хвилі, з яких вона складається. Маючи окремі звукові хвилі, ми додаємо потужність звуку для кожної з них.

```
[-1274, -1252, -1160, -986, -792, -692, -614, -429, -286, -134, -57, -41, -169, -456, -450, -541, -761, -1067, -1231, -1047, -952, -645, -489, -448
, -397, -212, 193, 114, -17, -110, 128, 261, 198, 390, 461, 772, 948, 1451, 1974, 2624, 3793, 4968, 5939, 6057, 6581, 7302, 7640, 7223, 6119, 5461,
 4820, 4353, 3611, 2740, 2004, 1349, 1178, 1085, 901, 301, -262, -499, -488, -707, -1406, -1997, -2377, -2494, -2605, -2675, -2627, -2500, -2148, -
1648, -970, -364, 13, 260, 494, 788, 1011, 938, 717, 507, 323, 324, 325, 350, 103, -113, 64, 176, 93, -249, -461, -606, -909, -1159, -1307, -1544,
-1815, -1725, -1341, -971, -959, -723, -261, 51, 210, 142, 152, -92, -345, -439, -529, -710, -907, -887, -693, -403, -180, -14, -12, 29, 89, -47, -
398, -896, -1262, -1610, -1862, -2021, -2077, -2105, -2023, -1697, -1360, -1150, -1148, -1091, -1013, -1018, -1126, -1255, -1270, -1266, -1174, -10
03, -707, -468, -300, -116, 92, 224, 72, -150, -336, -541, -820, -1178, -1289, -1345, -1385, -1365, -1223, -1004, -839, -734, -481, -396, -580, -52
7, -531, -376, -458, -581, -254, -277, 50, 331, 531, 641, 416, 697, 810, 812, 759, 739, 888, 1008, 1977, 3145, 4219, 4454, 4521, 5691, 6563, 6909,
6117, 5244, 4951, 4462, 4124, 3435, 2671, 1847, 1370, 1591, 1900, 1586, 713, 341, 462, 673, 60, -938, -1664, -2185, -2527, -2967, -3253, -3636, -38
59, -3723, -3134, -2380, -2032, -1831, -1457, -804, -241, -51, -113, -136, -122, -158, -147, -114, -181, -338, -266, 131, 418, 471, 651, 994, 1295,
 1267, 1197, 1291, 1110, 793, 514, 370, 174, -90, -139, 104, 334, 407, 524, 771, 1106, 1087, 878, 703, 591, 471, 91, -199, -357, -454, -561, -605,
-552, -512, -575, -669, -672, -763, -1022, -1435, -1791, -1999, -2242, -2563, -2853, -2893, -2740, -2625, -2556, -2385, -2138, -1936, -1803, -1649,
 -1495, -1460, -1446, -1345, -1177, -1088, -1072, -1003, -856, -719, -621, -585, -613, -634, -638, -636, -683, -819, -946, -1012, -964, -836, -762,
 -788]
```

Рис. 2. Двадцяти мілісекундний фрагмент дискретизованого запису слова «привіт»

Це можна зробити за допомогою перетворення Фур'є. Ми розкладаємо звукову хвилю на простіші звукові хвилі, з яких вона складається. Маючи окремі звукові хвилі, ми додаємо потужність звуку для кожної з них [3].

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ix\omega} dx. \quad (2)$$

В результаті, отримаємо оцінку важливості кожного частотного діапазону, від низьких частот до високих:



Рис. 3. Представлення кількості енергії в 50 Гц частотних відрізках

Тепер можна скласти спектрограму, що ще полегшить роботу нейронної мережі, адже їй буде простіше знаходити шаблони в таких даних, ніж в простих записах звуку:

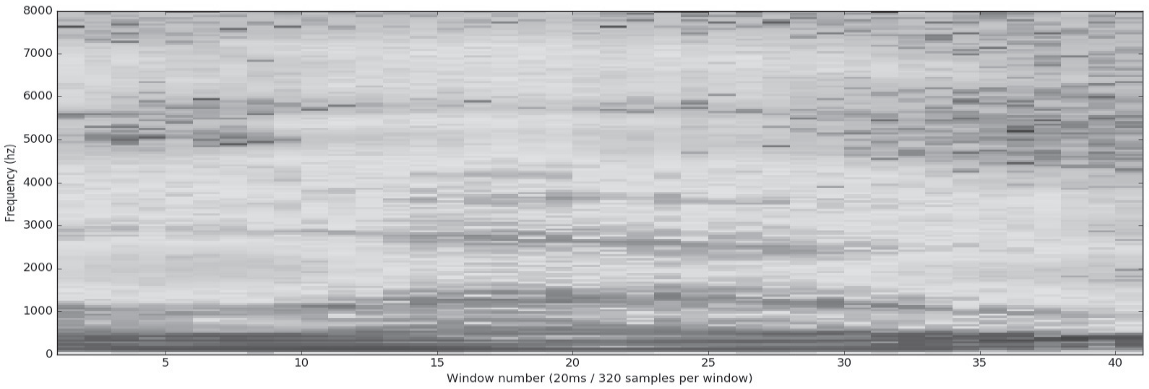


Рис. 4. Повна спектрограма аудіозапису слова «привіт»

Використання ж рекурентних нейронних мереж значно пришвидшить навчання самої мережі. Рекурентні нейронні мережі – це клас штучних нейронних мереж, у якому, з'єднання між вузлами – орієнтований цикл [4]. Це, в свою чергу, створює внутрішній стан мережі, що дозволяє їй проявляти динамічну поведінку в часі. А це означає, що дана мережа на кожному наступному кроці враховує результати попередніх кроків [5]. Звідси, кожна буква, визначена мережею, повинна впливати на ймовірну наступну букву. Наприклад, якщо ми сказали «прив», то швидше за все, далі скажімо «іт», щоб закінчити слово

«привіт». Ймовірність, що ми скажемо щось невимовне, наприклад «шхщ», надзвичайно мала. Таким чином, запам'ятовуючи попередні результати, наша мережа зможе робити більш точні прогнози в майбутньому.

Досліди. Внаслідок проведених дослідів, ми отримали результат «ППРРРРР__ИИ____ВВ_І_ТТТТТ». На першій ітерації, видаляємо зайві літери: «ПР__И____В_І_Т». На 2-ій ітерації, видаляємо пропуски – « ПРИВІТ».

Висновки. Проведено дослідження технологій розпізнавання голосу на основі рекурентних нейронних мереж. Запропоновано модель розпізнавання голосу українською. Наведені результати дослідів. Отримані результати показали, що застосований алгоритм для розпізнавання голосу користувача, є ефективним. Однак, навчання мережі має такі недоліки, як потреба у великому обсягу однотипної інформації для кращого навчання мережі та точнішого розпізнавання голосу. Тобто, для досконалого розпізнавання слова «привіт» потрібно близько двадцяти голосових записів цього слова з різним тембром, швидкістю вимовляння слова та амплітудою звукової хвилі.

Список використаних джерел

1. Colletti, Justin (February 4, 2013). "The Science of Sample Rates (When Higher Is Better—And When It Isn't)". Trust Me I'm A Scientist. Retrieved February 6, 2013.

2. Bengio, Y.; Courville, A.; Vincent, P. (2013). "RepresentationLearning: AReviewandNewPerspectives". IEEETransactions on Pattern Analysis and Machine Intelligence. 1798–1828. arXiv:1206.5538. doi:10.1109/tpami.2013.50

3. Bracewell, R. N. (2000), The Fourier Transform and Its Applications (3rd ed.), Boston: McGraw-Hill, ISBN 0-07-116043-4.

4. Bengio, Yoshua; LeCun, Yann; Hinton, Geoffrey (2015). "Deep Learning". Nature. 436–444. doi:10.1038/nature14539

5. Alex Graves, Santiago Fernandez, Faustino Gomez, Jürgen Schmidhuber Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks [Electronic source] http://www.cs.toronto.edu/~graves/icml_2006.pdf

ДОВІДКА ПРО АВТОРІВ

Смішний Дмитро Миколайович – студент IV курсу, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Smishnyi Dmytro – bachelor student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: sdmitriy3758@gmail.com

Алещенко Олексій Вадимович – асистент, кафедра обчислювальної техніки, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського".

Aleshchenko Oleksii – assistant professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: alexey.aleshchenko@gmail.com

Dmytro Smishnyi, Oleksii Aleshchenko

SPEECH RECOGNITION WITH DEEP LEARNING TECHNOLOGY USING

Task urgency. Problem of communication with a large set of devices at the same time increases too fast in our information world. It`s because each device requires its own user interface. Wherever convenient, for the user, use voice commands to manage various devices. This reduces the cost of developing UI. The instructions for such devices will be much clearer and simpler. Well, using of deep learning, in this case, cheapens and it unifies the development of voice control technology for different languages and countries. This article is devoted to the problem of voice control in Ukrainian, after all, unfortunately, there are very few such systems for this language.

Formulation of the problem. Practically complete absence of voice recognition systems for the Ukrainian segment of Internet users. The solutions that are already ready are imperfect.

Actual scientific researches and issues analysis. In recent years, you can see a significant boost in the development of voice recognition systems. The most famous examples are Amazon Alexa, Google Now, Siri, Alice from Yandex. All of them are based on neural networks. However, it is worth noting that they either do not support Ukrainian, or support for this language is extremely limited. Therefore, the article deals with the specificity of the recognition of the text in Ukrainian.

Uninvestigated parts of general matters defining. In this article the issue of recognition of voice is covered. The technology of deep learning is used. Research focuses on the generation of text based on voice, as such, and the use of this technology in relation to the Ukrainian language.

The research objective. The purpose of this article is to create a model that will work with Ukrainian words (in this case, the word Ukrainian "hello"), to recognize them, and to publish true text in accordance with the dictation.

The statement of basic materials. This article describes a method for generating text based on the Ukrainian speech. The analysis of the described stages of speech recognition based on the technology of deep learning is carried out.

Conclusions. The study of voice recognition technologies based on recurrent neural networks. A voice recognition model for Ukrainian language was proposed. The results of the experiments are presented.

Key words: recurrent neural networks, deep learning, Fourier transformation, sampling.

UDC 004.8

Денис Смішний, Олексій Алещенко

ОБРОБКА ВІЗУАЛЬНИХ ДАНИХ ЗА ДОПОМОГОЮ ТЕХНОЛОГІЇ COMPUTER VISION

VISUAL DATA PROCESSING BY COMPUTER VISION TECHNOLOGY

В даній статті розглядається питання розпізнавання об'єктів та образів з використанням технологій комп'ютерного зору. В ході дослідження описуються популярні алгоритми, такі як VGG16, DeepMask. Проводиться короткий огляд існуючих рішень в цій галузі. Визначаються перспективи використання комп'ютерного зору, зокрема на ринку мобільних пристроїв. Для цього проведено огляд можливостей бібліотеки OpenCV, та, як приклад, використання її у промислових продуктах.

Ключові слова: комп'ютерний зір, VGG16, OpenCV, розпізнавання об'єктів та образів.

Рис.: 3. Бібл.: 5.

In this article the question of recognition of objects and images using computer vision technologies is considered. The research describes popular algorithms such as VGG16, DeepMask. A brief overview of existing solutions in this area is conducted. The prospects for using computer vision, in particular in the mobile devices, are determined. To do this, reviewed the capabilities of the OpenCV library, and, for example, use it in industrial products.

Keywords: computer vision, VGG16, OpenCV, object and image recognition.

Fig. 3. Bibl.: 5.

Актуальність теми дослідження. В зв'язку зі зростаючими темпами впровадження інформаційних технологій в повсякденне життя користувачів, технології, що потребують залучення комп'ютерного зору набувають дедалі більшої популярності. Також подібні технології знаходять активне застосування в галузі криміналістики та пошукових операцій. Дані тенденції спонукають до широкого аналізу проблематики, пов'язаної з темою комп'ютерного зору та проведення певних досліджень в цій області.

Постановка проблеми. Недостатня кількість оглядових матеріалів з існуючих імплементацій машинного бачення. Дефіцит аналітичних статей з порівняння різних алгоритмів розпізнавання в контексті реальних прикладних задач. Заплутаність існуючих прикладів реалізації можливостей бібліотеки OpenCV.

Аналіз останніх досліджень та публікацій. Наразі у широкому доступі можна знайти величезну кількість матеріалів, присвячених тематиці комп'ютерного зору. Зокрема, доволі корисними в цьому плані будуть публікації на таких ресурсах як Хабрахабр [1] та Medium [2]. З іншого боку, подібні статті не мають акредитаційного підтвердження і можуть містити помилкову інформацію. Корисною для початківців буде стаття Learning to Refine Object Segments від науковців центру досліджень штучного інтелекту компанії Facebook [3]. Серед існуючих рішень корисним в плані розуміння існуючих алгоритмів можна виділити документацію бібліотеки з відкритим кодом OpenCV [4].

Виділення недосліджених частин загальної проблеми. На даний момент в науковій літературі спостерігається невелика кількість ґрунтовних публікацій з аналізом та порівнянням алгоритмів машинного бачення. У цих публікаціях рідко приділяється увагаіснуючим бібліотекам реалізацій алгоритмів та написанню прикладних програм з їх використанням.

Постановка завдання. Завданням є огляд основних алгоритмів реалізації технології комп'ютерного зору, наведення прикладу навчання нейронної мережі для визначення об'єктів на картинці та огляд бібліотеки OpenCV.

Викладення основного матеріалу. Комп'ютерним зором називають складну систему визначення та аналізу об'єктів та образів, отриманих системою ззовні. Вивчення можливостей та потенціалу комп'ютерів в галузі машинного бачення розпочалися ще в 70-х роках минулого століття, але активного розвитку зазнали саме час. Зокрема, було створено багато алгоритмів для розпізнавання образів: VGG16, VGG32, VGG29, ResNet, DenseNet, група алгоритмів Inception, тощо. Розглянемо детальніше кілька з них.

VGG16. VGG розшифровується як Visual Geometry Group, тобто група науковців з факультету інженерних наук Оксфордського університету, а 16 - число слоїв нейронної мережі, необхідних для розпізнавання образів. Суть даного алгоритму наступна:

1) поділ зображення пополам, доки не дійдемо до глибини зображення 14x14 пікселів (мінімальна можлива глибина - 3x3 пс);

2) в процесі поділів отриманих шматків зображення намагаємось отримати маски шуканого образу та коефіцієнти ймовірності знаходження шуканого образу в даній ділянці зображення;

3) на основі аналізу зменшених частин зображення будуються вихідні незалежні описи. Потім результат усереднюється для всіх масок.

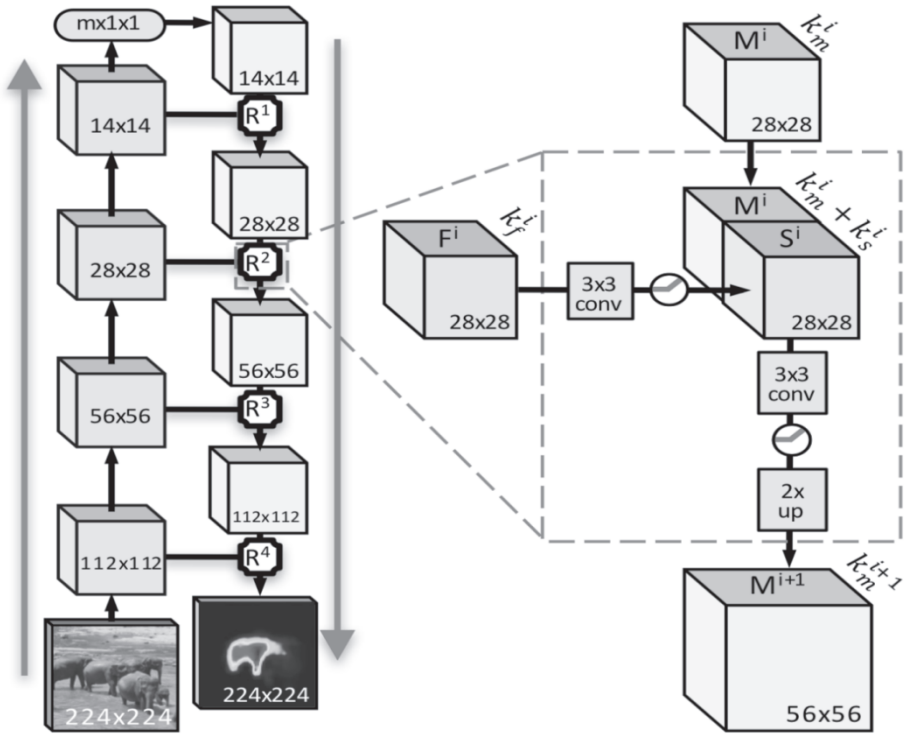


Рис 1. Схема роботи алгоритму

Даний підхід добре працює для знаходження образів, що контрастують з фоном, але не годиться для вилучення одноманітних об’єктів (Наприклад, знайти кількість овець на фото) або знаходження об’єктів, колір яких збігається з кольором фону. Для цього слід використовувати вдосконалений алгоритм згорткової нейронної мережі - DeepMask.

DeepMask. По факту, даний алгоритм є вдосконаленим алгоритмом VGG16. Суть в тому, що на кожному кроці ми додаємо не тільки маску образу, але й використовуємо інформацію, отриману від нижніх шарів, а також дані з розпізнаних верхніх шарів.

Кожний етап зменшення вихідного зображення дозволяє отримати кодовану маску, згенеровану при проходженні від загального зображення до зменшеного. Разом з цим у вдосконаленій згортковій нейронній мережі відбувається рух від меншого зображення до більшого з отриманням характерних функцій (точок). Підсумком є маска, отримана шляхом двонаправленого злиття функцій і опорних точок.



Рис.2. Результат роботи DeepMask

Головна проблема DeepMask в тому, що ця модель використовує просту мережу прямого поширення, яка успішно створює маски з відносно низьким рівнем дискретизації, але не виконує сегментацію з точністю до пікселя [3].

OpenCV. OpenCV (Open Source Computer Vision Library) - це бібліотека програмного забезпечення з відкритим кодом для комп'ютерного зору та машинного навчання. OpenCV була побудована, щоб забезпечити загальну інфраструктуру для програм комп'ютерного зору та прискорити використання машин у комерційних продуктах. OpenCV, що поширюється за ліцензію BSD, дозволяє підприємствам легко використовувати та змінювати код [5].

Бібліотека реалізує понад 2500 оптимізованих алгоритмів. Вони можуть бути використані для виявлення та розпізнавання облич, визначення об'єктів, класифікації людських дій у відео, відстеження руху камери, відстеження рухомих об'єктів, витягування 3D-моделей об'єктів, створення 3D-точкових хмар із стереокамер, зшивання зображень разом для отримання високої роздільної здатності зображення цілісної сцени, пошуку схожих зображень з тими, що знаходяться в базі даних, корекції ефекту “червоних очей”, що виникає при використанні спалаху під час фотографування, відстежування рухів очей, визначення декорації та встановлення маркерів для подальшого накладання їх на доповнену реальність тощо.

Дослідження. Використовуючи бібліотеку OpenCV реалізуємо детекцію QR-коду з урахуванням кутових відхилень. Для цього скористаємось теоремою про співвідношення.

Визначення: співвідношення площ N числа двомірних полігонів залишається інваріантним до рівномірного геометричного перетворення:

Якщо a_r - це площа багатокутника r, то:

$a_1 : a_2 \ldots : a_N = R$ (1)

в 2D площині X. Тоді:

$a'_1 : a'_2 \ldots : a'_N = R$ (2)

для будь-якої іншої 2D-площини Y, яка є геометричним перетворенням X:

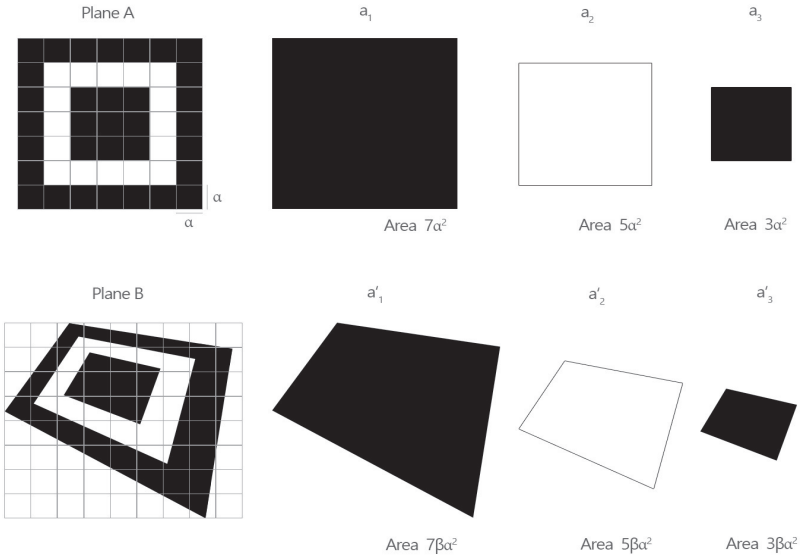


Рис. 3. Співвідношення площ полігонів

Вихідний код проекту доступний за посиланням https://github.com/denGlob/CppDemo.

Висновки. В ході проведеної роботи було розглянуто перспективи використання комп'ютерного зору. Оглянуто основні алгоритми виявлення образів та об'єктів на зображеннях. Зазначається, що алгоритм VGG16 добре працює для об'єктів, що контрастують з фоном, але є дуже неефективним в типових практичних задачах, де фон часто зливається з шуканим об'єктом. Для цього краще використовувати модернізації алгоритму такі як DeepMask та SharpMask. Оглянуто бібліотеку OpenCV. На її основі створено тестовий додаток для детекції QR-кодів.

Список використаних джерел

1. ComputerVision и с чем его едят [електронний ресурс] - https://habr.com/company/funcorp/blog/351638/

2. Lars Hulstaert. Going deep into object detection [електронний ресурс] - https://towardsdatascience.com/going-deep-into-object-detection-bed442d92b34

3. Pedro O. Pinheiro, Tsung-Yi Lin, Ronan Collobert, Piotr Doll´ar. (Jul 2016) Learning to Refine Object Segments (p.15)

4. OpenCV Documentation [електронний ресурс] - https://docs.opencv.org/

5. Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, B., Su, Z., Du, D., Huang, C., Torr, P. (2015). Conditional random fields as recurrent neural nets. In: ICCV.

ДОВІДКА ПРО АВТОРІВ

Смішний Денис Миколайович - студент IV курсу, група ІО-41, комп'ютерна інженерія, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Smishnyi Denys - student of IV course, group IO-41, computer engineering, Department of Computer Engineering, National Technical University of Ukraine "Kyiv Polytechnic Institute named after Igor Sikorsky".

E-mail: denshiksmile@gmail.com

Алещенко Олексій Вадимович – асистент, кафедра обчислювальної техніки, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського".

Aleshchenko Oleksii – assistant professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: alexey.aleshchenko@gmail.com

Denys Smishnyi, Oleksii Aleshchenko

VISUAL DATA PROCESSING BY COMPUTER VISION TECHNOLOGY

Relevance of research. In connection with the increasing rate of implementation of information technology in the daily life of users, technologies that require the attraction of computer vision are gaining popularity.

Formulation of the problem. Creation of a review of existing implementations of machine vision. Analysis of popular algorithms for the operation of neural networks for finding images and objects. Guide examples of commercial implementation of computer vision technology.

Analysis of recent research and publications. At present, you can find a huge amount of materials on the subject of computer vision in broad access. For beginners, the Learning to Refine Object Segments article is from the research team at Facebook's Artificial Intelligence Research Center. Among the existing solutions useful in understanding the existing algorithms, you can select the library documentation with OpenCV open source code.

Uninvestigated parts of general matters defining. To date, there is a small number of solid scientific articles with analysis and comparison of machine vision algorithms, as well as a review of libraries and their implementations.

Setting objectives. The task is to examine the main algorithms for the implementation of computer vision technology, give an example of studying the neural network to identify objects in the picture and review the library OpenCV.

Presentation of the main material. At the first stage we overview the VGG16 algorithm. Research shows, that this algorithm is good for contrast objects, but rather better to use DeepMask algorithm or SharpMask in partial tasks. Also, we overview in a short OpenCV library. It contains 2500 algorithms for computer vision and bridge code for popular languages such as Java and Python for comfortable coding.

Conclusions. During the work the prospects of using computer vision were considered. The main algorithms for the detection of images and objects in images are reviewed. It is determined that the VGG16 algorithm works well for objects contrasting with the background but is very ineffective in typical practical tasks, where the background often merges with the target object. To do this, it's best to use an upgrade algorithm such as DeepMask and SharpMask. The OpenCV library is under review. On its basis, a test application for the detection of QR-codes was created.

Keywords: computer vision, VGG16, OpenCV, object and image recognition.

УДК 004.93.12

Пантелєєв Олександр,
Пилипенко Павло,
Олійник Володимир

СИСТЕМА ВІЗУАЛЬНОГО МУЛЬТІТРЕКІНГА В РЕАЛЬНОМУ ЧАСІ

REALTIMEVISUALMULTITRACKINGSYSTEMY

В статті розглянута проблема візуального мультітрекінга. Запропонована реалізація системи мультітрекінга на базі фреймворка ROS. В системі використовується модифікація підходу трекінга на основі детектування, у якій для побудови траєкторії цілей на проміжних кадрах використовується незалежні трекери на основі кореляційних фільтрів.

Ключові слова: Комп’ютерний зір, візуальний трекінг, кореляційний фільтр, мультітрекінг, трекінг на основі детектування.

Рис.: 2. Бібл.: 6.

The article proposes implementation of the multitracking system based on the ROS framework. It is suggested to use a modification of method of tracking by detection, in which trackers based on correlation filters are used to construct the target trajectory between the key frames.

Keywords: Computer vision, visual object tracking, correlation filter, multitracking, tracking by detection.

Fig.: 2. Bibl.: 6.

Актуальність теми дослідження.Задача відстеження множини об'єктів на відео в реальному часі є невід'ємною частиною багатьох прикладних областей комп'ютерного зору, таких як системи відеоспостереження, аналіз дорожнього трафіку. Особливий інтерес дана задача представляє для автономних роботизованих агентів, здатних активно взаємодіяти з навколишнім світом, зокрема для аналізу оточення безпілотним літальним апаратом, автопілотом автомобіля, промисловим або домашнім роботом.

Постановка проблеми. Задачу мультітрекінга можна сформулювати наступним чином: Необхідно виявити на відео об’єкти, що належать до певних класів, та побудувати для кожного виявленого об’єкта його траєкторію на кадрі. Основною проблемою існуючих систем, які працюють в режимі реального часу, є низька точність та надійність відстеження об’єктів, що рухаються нелінійно.

Аналіз останніх досліджень і публікацій. Найбільш поширеним методом вирішення задачі мультітрекінга на даний момент є трекінг на основі детектування, в якому для побудови траєкторії об'єкта використовується динамічна модель та фільтр Калмана[1, 2]. Такий підхід використовується через високу швидкість розрахунків. Проте через обмеження фільтра Калмана та фактичну відмову від використання візуальних ознак такий трекер погано справляється з об'єктами, які рухаються нелінійно.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячення дослідженню можливостей використання кореляційних трекерів для вирішення задачі мультітрекінгу. Раніше використання кореляційних

трекерів у системах реального часу було неможливим, через велику кількість обчислень, проте нещодавно було запропоновано метод [3], який дозволяє значно збільшити швидкість роботи кореляційного трекера (300FPS для однієї цілі). Така швидкість дозволяє використовувати цей підхід для відстеження множини об’єктів.

Постановка завдання. Задачею даної роботи є розробка гібридної системи візуального мультітрекінга, яка поєднує в собі метод відстеження на основі детектування, та метод відстеження із допомогою кореляційних фільтрів. Розроблена система повинна в режимі реального часу виявляти цілі, та відстежувати траєкторію об’єктів, що рухаються нелінійно з більшою точністю, ніж у існуючих систем.

Викладення основного матеріалу. Для вирішення задачі мультітрекінга пропонується використовувати модифікацію підходу трекінгу на основі детектування, в якій для побудови траєкторії кожного об'єкта між ключовими кадрами використовуються незалежні трекери на основі кореляційних фільтрів. Після виявлення об’єктів детектором, для кожної цілі створюється кореляційний трекер, якій відстежує її до наступного повідомлення від детектора. Після отримання наступної множини об’єктів, список цілей оновлюється.

Для реалізації системи використовується фреймворк ROS [4], який дозволяє розділити систему на окремі вузли (Node). Кожен вузел являє собою незалежний програмний процес. Модулі обмінюються інформацією шляхом публікації повідомлень у відповідні топіки. Кожен вузел, який підписався на топік отримує копію повідомлення, яке відправив інший вузол у цей топік. Обмін даними здійснюється через TCPсокети, таким чином вузли можуть бути фізично розміщені на різних комп’ютерах. Запропонована система має три вузла: камера, детектор, трекер (рис. 1).

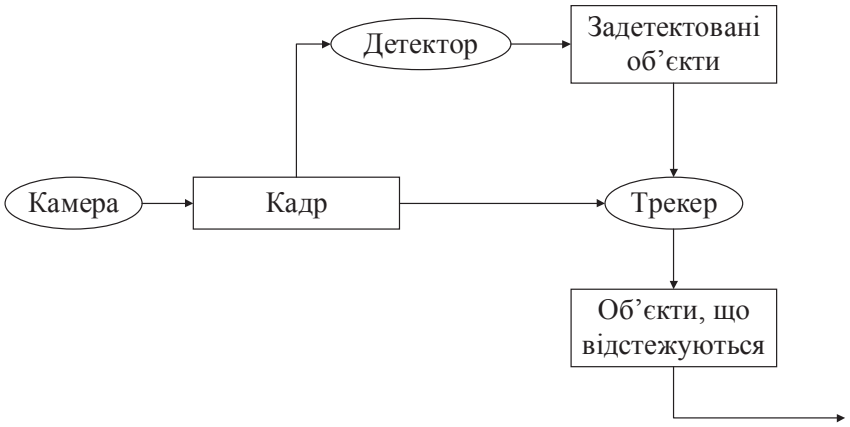


Рис. 1. Схема системи мультітрекінга.

Камера. Цей вузол відповідає за публікацію відеоряду. Як джерело даних може використовуватися веб-камера, відеофайл або директорія з набором кадрів. Кожному кадру присвоюється мітка часу, яка використовується для синхронізації.

Детектор відповідає за виявлення об'єктів на відео. При отриманні нового кадру, на ньому виявляються об'єкти за допомоги нейронної мережі SSD (Single Shot Detector) [5]. Для підвищення швидкості роботи, усі кадри, що надходять на детектор під час обробки поточного зображення, відкидаються.

Трекер. Трекер відповідає за відстеження об’єктів на проміжних кадрах, для цього для кожного об’єкта створюється незалежний кореляційний трекер, який визначає позицію цілі на кадрі. Цей вузол отримує як кадри відео, так і масиви об'єктів, що виявив детектор.

При отриманні нового кадру, оновлюються позиції усіх цілей на цьому кадрі із допомогою їх кореляційних трекерів.

При отриманні набору виявлених об'єктів від детектора, відбувається оновлення списку цілей. Для того щоб правильно розрахувати візуальні ознаки для кожного об’єкту повідомлення зі списком виявлених об'єктів синхронізується з кадром з камери, а потім поточної набір об'єктів зіставляється з об'єктами, які виявив детектор. Зіставлення відбувається за допомогою угорського алгоритму, критерієм подібності об’єктів виступає відношення площин обмежувальних рамок до їх об’єднання (IntersectionOverUnion). За результатами зіставлення можливо три випадки:

- Якщо два об'єкти зіставлені, трекер коригує свою модель, використовуючи позицію і візуальні ознаки відповідного об'єкта, виявленого детектором.
- Якщо об'єкт що відслідковується не був зіставлений з жодним об'єктом з детектора, він вважається втраченим і його відслідковування припиняється.
- Якщо виявлений детектором об'єкт не був зіставлений, вважається, що в кадрі з'явився, новий об'єкт. Для нього створюється новий трекер з новим унікальним ідентифікатором.

Синхронізація. Оскільки вузли системи є незалежними, і повідомлення надходять у довільний момент часу, у вузлі трекера треба здійснювати синхронізацію кадрів та масивів об’єктів які виявив детектор.

Для цього створюються три буфера, в перший додаються повідомлення які містять кадр, у другий – повідомлення, які містять масив об’єктів, що були знайдені на кадрі. Після цього повідомлення, які мають однакову мітку часу, об’єднуються у кортеж і додаються до третього буфера, який у подальшому використовує трекер (рис. 2).

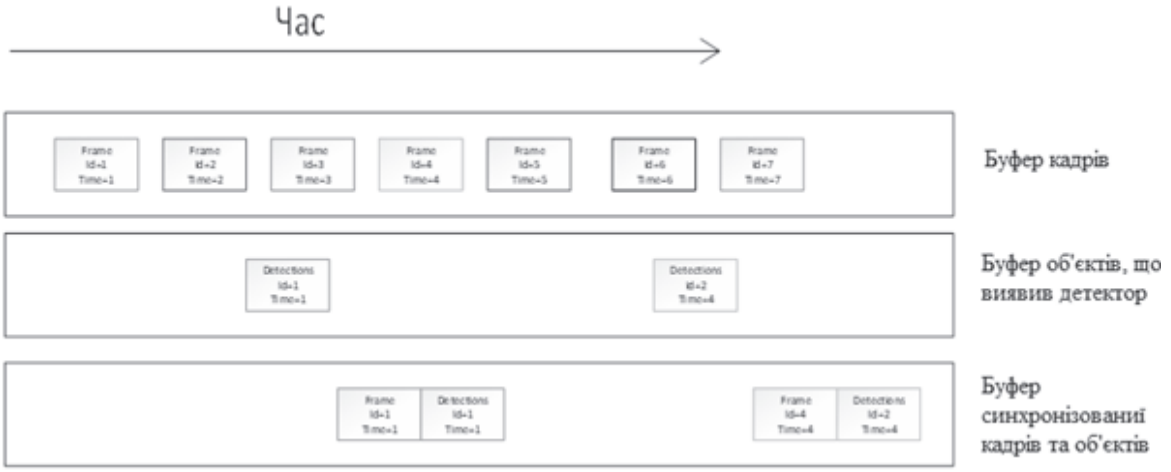


Рис. 2. Схема буферів для синхронізації повідомлень

Експерименти. Для оцінки результатів роботи системи використовується бенчмарк MOT[6]. Датасет містить 11 відео. Відео зняті в різноманітних умовах

(різна освітленість, кількість об’єктів, швидкість та характер руху камери тощо), для того щоб протестувати роботу системи у різних умовах та уникнути пристосування до якогось певного сценарію. Середня швидкість роботи трекера склала 73 FPS, а середня точність – 83%.

Висновки. Запропонована в статті система відстеження множини об’єктів в реальному часі, дозволяє підвищити точність і надійність трекінга об'єктів, що рухаються нелінійно. Це досягається за рахунок використання кореляційних фільтрів візуальних ознак для відстеження цілі на проміжних кадрах. Система мультітрекінга була реалізована на базі ROS, що дозволило зробити компоненти системи незалежними, та здійснювати розподілену обробку даних, завдяки чому підвищується швидкість та надійність системи.

Список використаних джерел

1. Bochinski E., Eiselein V., Sikora T. High-Speed tracking-by-detection without using image information // IEEE International Conference on Advanced Video and Signal Based Surveillance. 2017. C. 1-6.

2. Luo W. Multiple Object Tracking: A Literature Review. [Електронний ресурс] // arxiv.org URL: https://arxiv.org/pdf/1409.7618.pdfДата звернення: 03.04.2018

3. Henriques J., Caseiro R. High-speed tracking with kernelized correlation filters // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015. C.583–596.

4. Morgan Q. ROS: an open-source Robot Operating System // ICRA Workshop on Open Source Software. 2009.

5. Liu W.SSD: Single shot multibox detector // European Conference on Computer Vision. Springer. 2016. C.21–37.

6. Milan A. MOT16: A benchmark for multi-object tracking [Електронний ресурс] // arxiv.org URL: https://arxiv.org/pdf/1603.00831.pdf Дата звернення: 29.04.2018

ДОВІДКА ПРО АВТОРІВ

Пантелєєв Олександр Сергійович – студент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Pantelieiev Oleksandr –student, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: spn220v@gmail.com

Пилипенко Павло Геннадійович – студент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Pylypenko Pavlo –student, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: pylypenko.pg@gmail.com

Олійник Володимир Валентинович – к.т.н., старший викладач, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Oliinyk Volodymyr – Ph.D., associate professor, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: oliinyk.volodymyr@gmail.com

Pantelieiev Oleksandr,
Pylypenko Pavlo,
Oliinyk Volodymyr

REALTIMEVISUALMULTITRACKINGSYSTEMY

Target setting. The problem of visual multiple object tracking is an integral part of many areas of computer vision, such as video surveillance systems, traffic analysis. Thi task is important for autonomous robotic agents, capable of active interaction with the outside world, in particular for the analysis of the environment by an unmanned aerial drone, autopilot of an autonomous car, an industrial or home robot.

Formulation of the problem. The problem of multitracking can be formulated as follows: On the video sequence, system must detect objects belonging to certain classes, and construct for each detected object its trajectory on a frame.

The main problem of existing systems operating in real time is the low accuracy and reliability of tracking objects that are move nonlinearly.

Actual scientific researches and issues analysis. Currently the most common method of solving the multitracking problem is tracking based on detection, with a dynamical model and a Kalman filter. This approach is used because of the high speed of calculations. However, due to the limitation of the Kalman filter, such tracker does not handle objects that move nonlinearly.

Definition of uninvestigated parts of the problem. This article studies possibilities of using correlation trackers for solving the multitracking problem.

The research objective. The objective is to create the system that can automatically detect and track objects of particular classes on video.

Main results. Improvedmethod of multitracker that uses correlational tracker was presented. Distributed system for visual multitracking was developed using ROS framework. Received results were accurate enough and the speed was sufficientfor real-time applications.

Conclusions. The proposed multitracking method allows to increase the accuracy and reliability of tracking of the objects moving nonlinearly. This is achieved by using the correlation filter to track the target between keyframes. The multitracking system based on this method system was implemented using ROS framework, which allows independent components of the system and implements distributed data processing that increase speed and reliability of system.

Key words: Computer vision, visual object tracking, correlation filter, multitracking, tracking by detection.

УДК 004.75

Богдан Бугай,
Юрій Гордієнко

МЕТОД ОПТИМІЗОВАНОГО ЗНАХОДЖЕННЯ ПОДІБНИХ БІНАРНИХ РЯДКІВ У ПЕВНІЙ МНОЖИНІ

METHOD OF OPTIMIZED SEARCH SIMILAR BINARY STRINGS IN A CERTAIN SET

У статті розглядається питання швидкого знаходження подібних рядків у певній множині бінарних рядків з однаковою довжиною, для яких відстань Геммінга не перевищує задане значення.

Ключові слова: відстань Геммінга, бінарний рядок, хеш таблиця.

Табл.: 2. Бібл.: 2.

The article deals with the problem of the rapid finding of similar strings in a certain set of binary strings with the same length, for which the Hamming distance does not exceed a given value.

Key words: Hamming distance, binary string, hash table.

Tabl.: 2. Bibl.: 2.

Актуальність теми дослідження. Порівняння бінарних рядків за допомогою відстані Геммінга ключовий момент в задачах, де вони приймаються в якості характеристики певного текстового або мультимедійного контенту, особливо в задачах де використовується перцептивний хеш в якості характеристики досліджуваного елемента.

Постановка проблеми. Відсутність швидкого методу для знаходження подібних бінарних рядків в певній множині для яких відстань Геммінга не буде перевищувати задану.

Аналіз останніх досліджень і публікацій. Протягом останніх років не з’являлося статей присвячених цій темі.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена вивченню та аналізу запропонованого підходу для оптимізації пошуку подібних рядків в певній множині бінарних рядків з однаковою довжиною, для яких відстань Геммінга не перевищує задане значення.

Постановка задачі. Для певного бінарного рядка потрібно знайти всі рядки в певній множині для яких відстань Геммінга не буде перевищувати задане значення.

Викладення основного матеріалу. Бінарний рядок складається з елементів множини $\{0, 1\}$. Для нього можна виділити три основні характеристики: кількість нулів, кількість одиниць та загальна довжина рядка.

Якщо взяти певну множину бінарних рядків з однаковою довжиною, то їх можна розділити на групи, в яких кількість одиниць в рядку дорівнює номеру групи.

Таблиця 1

Ідентифікатор рядка	Бінарний рядок
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

В Таблиці 1 представлена деяка вибірка бінарних рядків, для простоти бінарний рядок в десятковій системі числення відповідає ідентифікатору. Далі розбиваємо рядки на групи по кількості одиниць.

Таблиця 2

Кількість одиниць	Номер ідентифікатора рядка
0	0
1	1, 2, 4
2	3, 5, 6
3	7

У програмній реалізації Таблиця 2 може бути хеш-таблицею або деякою базою даних, яка надає ефективний доступ за ключем. Даний процес групування показаний для примітивної вибірки, але суть залишається незмінною для вибірки будь-якої величини.

Відстань Геммінга обчислюється як число позицій, в яких відповідні значення двох рядків різні [1]. Для двох бінарних рядків S_1 та S_2 можна підрахувати за формулою [2]:

$$d_h(S_1, S_2) = \sum_{k=1}^{p} |S_1(k) - S_2(k)| \quad (1)$$

де p - це довжина рядку.

Приклад підрахунку:

$$d_h(001{,}0110) = 1 \quad (2)$$

Відстань Геммінга має наступні властивості:

- $d(S_1, S_2) \geq 0$
- $d(S_1, S_1) = 0$
- $d(S_1, S_2) = d(S_2, S_1)$

Так як відстань для двох однакових рядків повинна бути нуль, то відповідно в них кількість одиниць та нулів однакова. Для рядків, між якими відстань повинна бути 1, на одну одиницю чи нуль повинно бути більше ніж в цільового рядка. З цього випливає, що для знаходження всіх подібних рядків між якими є задана відстань Геммінга d потрібно виконати наступний алгоритм:

1) Підраховуємо кількість одиниць в рядку і позначаємо її як C_s;
2) Вибираємо категорії в яких ідентифікатори належать множині S:

$$S = \{C_s - d, \dots, C_s, \dots, C_s + d\} \quad (3)$$

де $C_s - d \geq 0$ та $C_s + d \leq$ length(S)

3) Послідовно проходимо по кожній з категорій та підраховуємо відстань Геммінга для кожного рядка категорії з цільовим рядком;

4) Рядки, для яких відстань буде менша або рівна d, включаємо в результуючу множину;

Даний метод дозволяє проходити лише частину всієї множини для пошуку рядків подібних до цільового. В залежності від відстані Геммінга та характеру запиту ця частина може змінювати свої розміри.

Для тесту згенеруємо рядки, значення яких в десятковій системі числення лежатимуть в межах між 0 та 65535. Код для генерації на мові програмування Python:

```
table = dict()
size = 16
for i in range(2 ** size):
    s = '{0:{fill}{size}b}'.format(i, fill=0, size=size)
    count = sum(int(char) for char in s)
    if count in table:
        table[count].append(s)
    else:
        table[count] = [s]
```

Припустимо, в даній множині нам потрібно знайти подібні рядки для рядка 1111111100000000 з допустимою похибкою 1, то за алгоритмом подібні рядки потрібно шукати в 7, 8 та 9 категорії. Сумарна довжина цих категорій дорівнює $11440 + 12870 + 11440 = 35750$. Це означає, що 29785 рядків не потрібно сканувати. Для прикладу були взяті категорії з найбільшою кількістю рядків, якщо кількість одиниць в рядку прямує до 0 або до довжини рядку, то операцій сканування буде ще менше. Наприклад для рядка 1111111111110000 та з допустимою похибкою 1 потрібно провести $4368 + 1820 + 560 = 6748$ операцій, так як були обрані 11, 12 та 13 категорії. Зменшення кількості операцій відчутне.

Висновки. У роботі продемонстровано, як за допомогою розбиття бінарних рядків на категорії по кількості одиниць в рядку можна досягти приросту в продуктивності пошуку в задачах, де потрібно знаходити подібні рядки із заданою відстанню Геммінга. Даний алгоритм підходить для задач, де використовується пошук по перцептивних хешам.

Представлений алгоритм дуже просто можна покласти в основу багатопоточної програми, так як категорії, по яких потрібно провести сканування незалежні, тобто кожен потік може опрацьовувати свою категорію, потрібно лише використати будь-який примітив синхронізації для результуючої множини для уникнення станів гонки даних.

Є декілька напрямків для майбутньої роботи. Перш за все - покращення пошуку подібних рядків у межах певної категорії.

Список використаних джерел

1. Ionescu M., Ralescu A., (2004). Fuzzy Hamming Distance in a Content-Based Image Retrieval System. International Journal of Information Technology and Knowledge Management, IEEE (p. 2).

2. Xu Z., (2007). Fuzzy Optimization and Decision Making. Springer (p. 3).

ДОВІДКА ПРО АВТОРІВ

Бугай Богдан Андрійович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Buhai Bohdan - student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: bohdan.buhai@yahoo.com

ГордієнкоЮрійГригорович – професор, кафедраобчислювальноїтехніки, НаціональнийтехнічнийуніверситетУкраїни «КиївськийполітехнічнийінститутіменіІгоряСікорського».

Gordienko Yuri– professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: yuri.gordienko@gmail.com

Buhai Bohdan,
Gordienko Yuri

METHOD OF OPTIMIZED SEARCH SIMILAR BINARY STRINGS IN A CERTAIN SET

Target setting. The comparison of binary strings with Hamming's distance is a key point in tasks where they are taken as a characteristic of a particular text or multimedia content, especially in tasks where the perceptual hash is used as a characteristic of the investigated element.

Problem formulation. The absence of a fast method for finding similar binary strings in a certain set for which the Hamming distance does not exceed the given one.

Actual scientific researches and issues analysis. During recent years, there were no articles devoted to this topic.

Uninvestigated parts of general matters defining. This article is devoted to the study and analysis of the proposed approach to optimize the search for similar strings in a certain set of binary strings with the same length, for which the Hamming distance does not exceed a given value.

The research objective. For a certain binary string need to find all the rows in a certain set for which Hamming distance does not exceed the given value.

The statement of basic materials. An analysis of the division method into categories in the task where similar binary strings are to be found is carried out. An algorithm for search optimization is described. The results were well-interpreted and informative.

Conclusions. The work of the algorithm and the obtained results are analyzed. The presented approach proved its effectiveness. The results of the experiments are summarized and following steps described to improve the performance of the algorithm.

Key words: Hamming distance, binary string, hash table.

УДК 004.75

Богдан Бугай,
Юрій Гордієнко

МЕТОД ВИМІРЮВАННЯ ПОДІБНОСТІ ВІДЕОМАТЕРІАЛІВ

METHOD OF MEASUREMENT VIDEO SIMILARITY OF VIDEO CONTENT

У статті розглядається питання знаходження міри подбіності між двома відеофайлами. В якості харакетристик відеокадру було вибрано перцептивний хеш та гістограму.

Рис.: 4. Бібл.: 3.

The article considers the question of finding a measure of similarity between two video files. As video frame characteristics have been selected perceptual hash and histogram.

Key words: comparison of video content, perceptual hash, histogram, Hamming distance.

Fig.: 4. Bibl.: 3.

Актуальність теми дослідження. В багатьох сферах діяльності людини, починаючи від комп'ютерного зору і пошукових систем, систем дедублікації даних та закінчуючи різними системами контролю та модерації відеоматеріалів, є потреба в швидкому аналізі відеоматеріалів та знаходження схожих чи однакових відеофайлів. Дана робота присвячена проблемі знаходження міри подібності між відеоматеріалами.

Постановка проблеми. Відсутність швидкого методу для визначення подібності між відеоматеріалами.

Аналіз останніх досліджень і публікацій. Протягом останніх років з'явилося декілька методів для знаходження подібності між відеоматеріалами, але метод, який використовує перцептивний хеш на основі серднього значення кольору та гістограми в якості характеристики кадру відеофайлу ще не розглядався.

Виділення недосліджених частин загальної проблеми.Дана стаття присвячена вивченню та аналізу запропонованого підходу для виміру подібності відеоматеріалів. Дослідження сфокусовано на перевірці чи достатньо перцептивного хешу на основі середнього значення кольору та гістограми для знаходження подібних відеоматеріалів.

Постановка завдання. Для певного відео X потрібно знайти значення подібності з відео Y.

Викладення основного матеріалу. Відеофайл - це список зображень, які з певною швидкістю, відображаються користувачу:

$$\boldsymbol{V_m = [I_1, \dots, I_j, \dots I_n]} \quad \textbf{(1)}$$

де I_j - це певне зображення у відеоряді.

Кожне зображення несе в собі певний структурний зміст та кольорову гамму. Тому можна припустити, що для порівняння двох відеофайлів, потрібно знайти структурну та кольорову характеристику кожного зображення в даних відеофайлах і порівняти їх між собою.

Кольорова складова зображення визначається гістограмою. У математичному сенсі гістограма - це функція m_i, яка підраховує кількість спостережень, що потрапляють у кожну з незалежних категорій:

$$n = \sum_{i=1}^{k} m_i \tag{2}$$

де n - загальна кількість спостережень, k - загальна кількість категорій.

Розбиття на категорії дозволяє визначити різницю в певному кольорі між гістограмами. Для цього можна використати коефіцієнт кореляції між гістограмами H_1та H_1:

$$d_c(H_1, H_2) = \frac{\sum_{i=1}^{n}(H_1(i) - \overline{H_1}) * (H_2(i) - \overline{H_2})}{\sqrt{\sum_{i=1}^{n}(H_1(i) - \overline{H_1})^2 * \sum_{i=1}^{n}(H_2(i) - \overline{H_2})^2}} \tag{3}$$

де $\overline{H_k} = \frac{1}{n}\sum_{j=1}^{n} H_k(j)$

Дану характеристику потрібно підрахувати для трьох каналів кольору зображення. Коефіцієнт кореляції набуває значень від -1 до $+1$. Значення $+1$ означає, що залежність між H_1та H_1 є лінійною [1].

У зображеннях високі частоти забезпечують деталізацію, а низькі частоти показують структуру. Зображення з великим розширенням містить багато високих частот. Для зменшення високих частот найкраще підходить зменшення розширення зображення. Якщо зменшити зображення, то на зображенні буде присутній лише силует того, що було на оригіналі, це дає змогу виокремити структуру зображення. Для знаходження унікальних ознак зображення буде використовуватися перцептивний хеш.

Перцептивні хеші відрізняються від криптографічних хеш-функцій. У криптографії кожен хеш є випадковим. Дані, які використовуються для генерації хешу, виконують роль джерела випадкових чисел так, що однакові дані дадуть однаковий результат, а різні дані - різний. Після порівняння двох хешів можна зробити тільки два висновки. Якщо хеші відрізняються, - дані різні. Якщо хеші збігаються, то і дані, швидше за все, - однакові. На відміну від криптографічних, перцептивні хеші можна порівнювати між собою і робити висновок про ступінь відмінності двох наборів даних [2].

В процесі хешування зображення можна виділити два головні етапи:

1) Вивчення вмісту зображення;
2) Конструювання хеш-значення, яке однозначно ідентифікує вхідне зображення на основі вмісту зображення;

Існує багато методів для знаходження перцептивного хешу від зображення, кожен з яких має свої переваги та недоліки. В даній статті для наочності знаходження структурної характеристики зображення буде використаний перцептивний хеш на основі середнього значення кольору.

Для знаходження хешу зображення потрібно виконати наступні кроки:

1) Перевести зображення в градації сірого;
2) Змінити розмір зображення до $8\ x\ 8$;
3) Підрахувати середнє значення кольору для зображення;
4) Порівняти кожен піксель з середнім значенням кольору, якщо він більший або рівний, то в результуючу вибірку записати 1, якщо ні 0;
5) Підрахувати хеш;

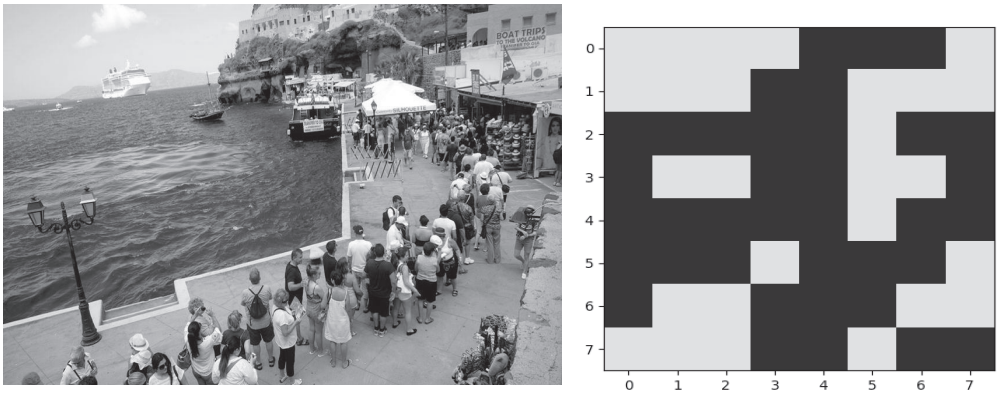


Рис. 1. Результат роботи алгоритму

На основі характеристик зображення відбувається генерація індивідуального, але не унікального відбитка, і ці відбитки можна порівняти з іншими за допомого відстані Геммінга d_h [3]:

$$d_h(M,N) = \sum_{k=1}^{p} |M(k) - N(k)| \quad (4)$$

де p - це довжина рядку, M та N - це перцептивні хеші.

Отже, на основі значень d_c та d_h можна встановити ступінь подібності двох відеокадрів.

Порівняння двох відеофайлів. Кожна секунда відео складається як мінімум із 24 кадрів, котрі в загальному випадку не сильно відрізняються одне від одного, тому для оптимізації даного процесу потрібно з кожної секунди відео вибирати лише декілька кадрів. В результаті буде отримано дві множини відеокадрів S_1 та S_2.

Коефіцієнт подібності між двома відеокадрами I_1, з множини S_1, та I_1, з множини S_2, складається з міри подібності по гістограмі та по перцептивному хешу. Якщо їх не зводити до одного значення то це дає можливість встановлення певних умов на пошук, наприклад коефіцієнт подібності по перцептивним хешам не повинен бути нижчим ніж заданий, також схожий принцип можна застосувати до гістограми де коефіцієнт кореляції не повинен бути нижче ніж задане значення. Значення коефіцієнта подібності по перецептивному хешу можна вирахувати за наступною формулою:

$$S_{ph}(I_1, I_2) = \frac{N - d_h(PH(I_1), PH(I_2))}{N} \quad (5)$$

де N – це довжина перцептивного хешу, PH– це функція обчислення перцептивного хешу. S_{ph} буде належати інтервалу [0, 1].

Для знаходження коефіцієнту подібності по гістограмі потрібно знайти коефіцієнт кореляції для кожного каналу кольору:

$$S_c(I_1, I_2) = (d_c\big(H_r(I_1), H_r(I_2)\big),\ d_c\left(H_g(I_1), H_g(I_2)\right), d_c\big(H_b(I_1), H_b(I_2)\big)) \quad (6)$$

де S_c – це кортеж, значення кожного елементу кортежу належатимуть інтервалу [-1, 1]

У загальному випадку потрібно визначити, яка з двох множин має більший розмір, і з більшої виокремити всі підмножини, які мають розмір меншої, та не змінювати в них порядок елементів відносно оригінальної.

Наприклад, візьмемо дві множини одна розміром 5 , назвемо її A, іншу з розміром 3 , назвемо її B. Тоді більшу множину можна розкласти на підмножини за допомогою функції J та отримати список L з підмножин з A .

Після розбиття на підмножини потрібно підрахувати коефіцієнти подібності S_{ph}між всіма парами (L_i, B). Дане значення можна визначити за формулою:

$$VSPH(L,B) = \sum_{i=1}^{N} \frac{S_{ph}(L_i,B)}{N} \quad (7)$$

де N – довжина підмножини.

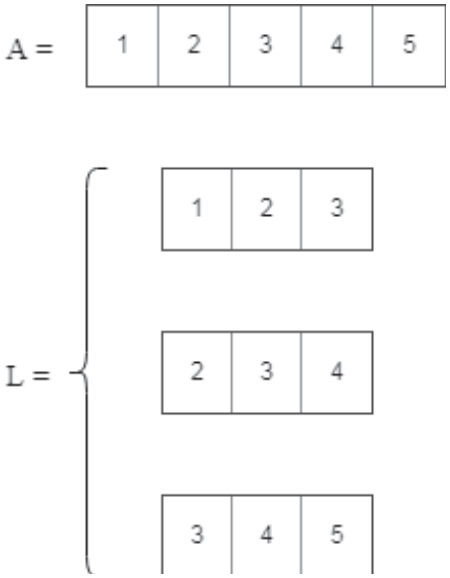


Рис. 2. Приклад розбиття на підмножини

Для підрахунку коефіцієнта подібності кореляцій для двох підмножин можна використати формулу:

$$VSC(L,B) = \sum_{i=1}^{N} \frac{S_c(L_i,B)}{N} \quad (8)$$

У цілому подібність двох відеофайлів описується кортежем з двох значень. Так як дані характеристики підраховуються для всіх (L_i, B), то подібність між A та B буде дорівнювати:

$$VS(A,B) = (\max(VSPH(J(A),B)), VSC(J(A),B)) \quad (9)$$

Отже, знаходимо максимальне значення за структурною характеристикою та значення гістограм на цьому інтервалі, і отриманий кортеж буде значенням міри подібності між двома відеофайлами.

Якщо список L складається з одного елементу, то це частковий випадок, коли два відеофайли мають однакову довжину.

Модифікація методу порівняння двох відеофайлів для подальшої реалізації в паралельних системах із спільною пам'яттю. Представлений метод знаходження міри подібності між двома відеофайлами дуже добре піддається розпаралелюванню, так як процес підрахунку значень для двох відео файлів можна розділити на незалежні частини. Підрахунок ступення подібності

для (L_i, B) для кожної пари можна обраховувати в окремому потоці. А потім, коли всі обрахунки закінчаться, зібрати всі результати в певний вектор і знайти в ньому найбільший коефіцієнт подібності.

Потрібно створити пул потоків, розмір якого не повинен перевищувати кількості ядер в системі, інакше ефективність алгоритму втратиться. Тоді потрібно створити чергу, яка б мала критичну секцію на операції отримання елементів з черги для уникнення дублювання обчислень та гонок даних. Також потрібно створити результуючий вектор, який би забезпечував синхронізацію на етапі вставки елементів для уникнення ситуацій гонки даних.

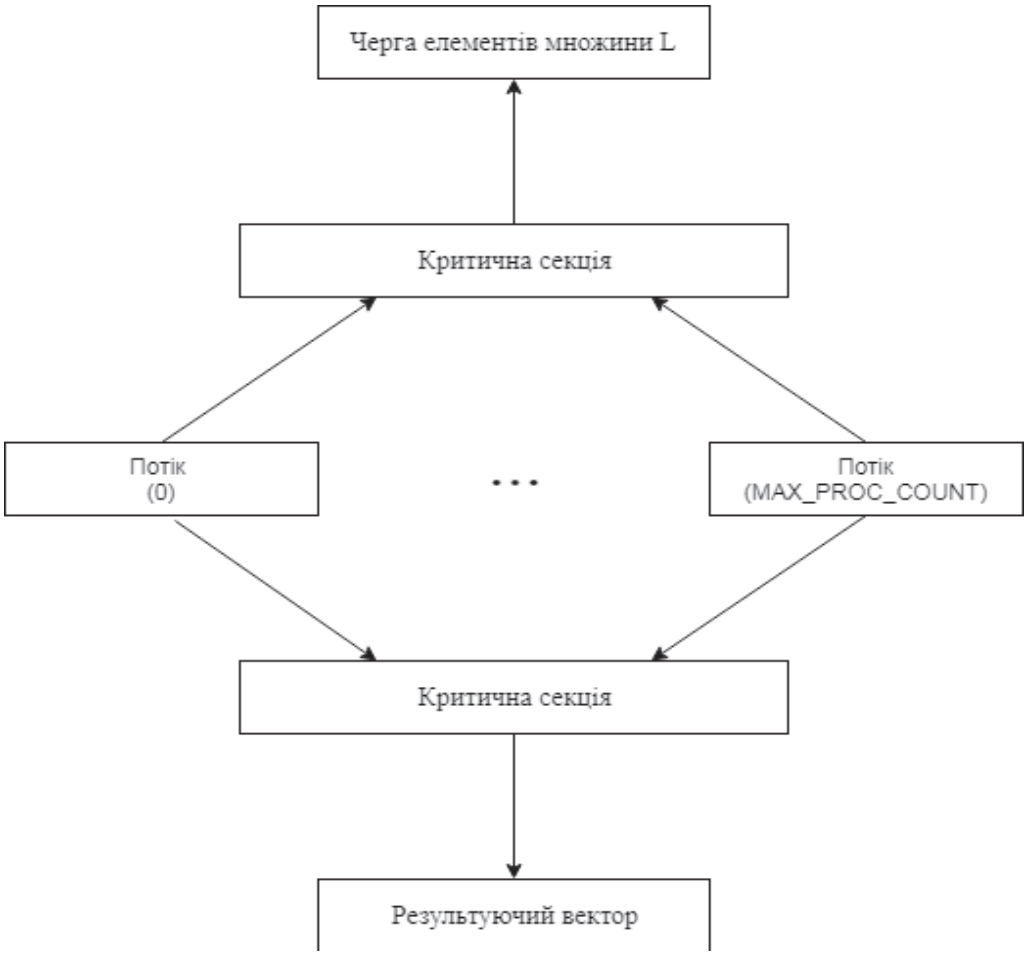


Рис. 3. Алгоритм паралельної обробки множини *L*

Коли потоки обчислення значень подібності між вхідним відеофайлом та елементами множини L закінчать свою роботу, потрібно запустити паралельний алгоритм знаходження максимуму у векторі для оптимізації пошуку максимального значення подібності.

Знаходження максимуму можна розбити на етапи. На першому етапі розділимо результуючий вектор V на декілька векторів і для кожного знайдемо максимум:

$$a_i = \max(V_H) \quad (10)$$

де $i = \overline{1, MAX_PROC_COUNT}$, V_H— це частина вектору яку оброблятиме кожен потік окремо.

Тоді загальний максимум буде дорівнювати:

$$a = \max(a, a_i) \quad (11)$$

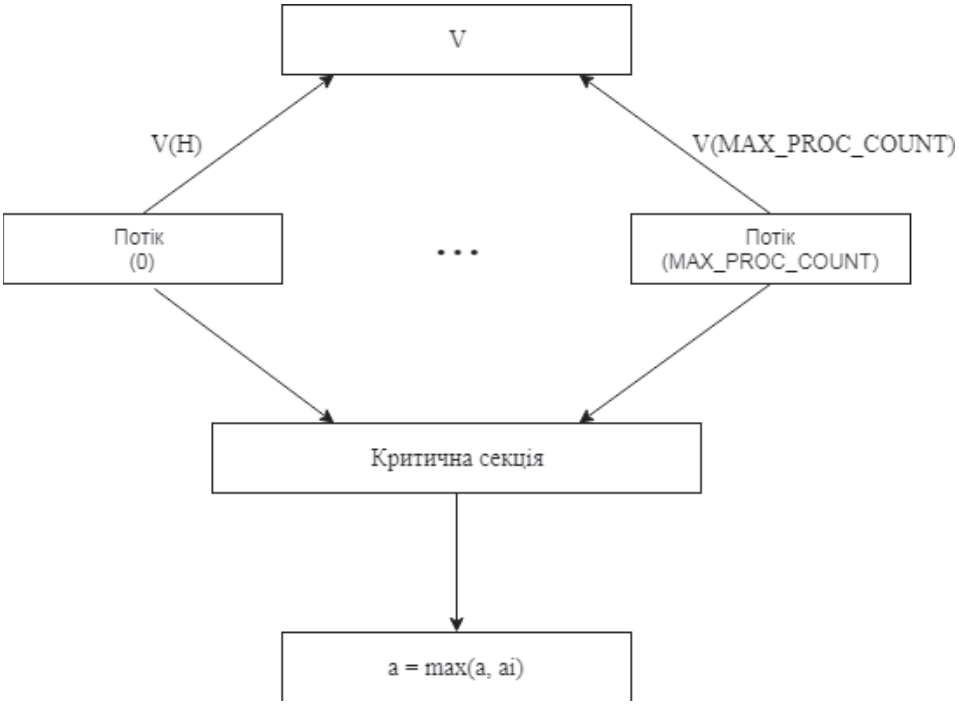


Рис. 4. Алгоритм паралельного знаходження максимального значення подібності

Даний алгоритм можна застосувати в системах із спільною пам'яттю, але при деяких змінах його можна реалізувати для систем з локальною пам'яттю також.

Висновки. У роботі продемонстровано, як за допомогою перцептивних хешів та гістограми знаходити міру подібності між відеофайлами. Даний метод має гарну швидкодію та невразливий до невеликих змін кадрів відеофайлу таких, як незначні шуми, невеликий градус зміни осі відображення та незначні зміни форм фактору. Також представлений алгоритм для реалізації описаного методу в паралельних системах із спільною пам'яттю.

Список використаних джерел

1. Rodgers L., Nicewander W. A. (February 1988). Thirteen ways to look at the correlation coefficient. The American Statistician (p. 5).

2. Vishal M., Brian L. (2006). Perceptual image hashing via feature points: performance evaluation and trade-offs. IEEE Transactions on Image Processing (p.1).

3. Xu Z., (2007). Fuzzy Optimization and Decision Making. Springer (p. 3).

ДОВІДКА ПРО АВТОРІВ

Бугай Богдан Андрійович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Buhai Bohdan - student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: bohdan.buhai@yahoo.com

Гордієнко Юрій Григорович – професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Gordienko Yuri – professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: yuri.gordienko@gmail.com

Buhai Bohdan, Gordienko Yuri

METHOD OF MEASUREMENT VIDEO SIMILARITY OF VIDEO CONTENT

Target setting. In many areas of human activity, from computer vision and search engines, data deduplication systems to various video moderation control systems, there is a need for quick analysis of video material and finding similar video files. This work is devoted to the problem of finding a measure of similarity between video materials.

Problem formulation. The lack of a quick method to determine the similarity between the footage.

Actual scientific researches and issues analysis. During the past few years, there have been several methods for finding the difference between the footage, but the method that uses the perceptual hash based on the average color and the histogram, as a frame file characteristic of the video file, has not yet been compared.

Uninvestigated parts of general matters defining. This article is devoted to the study and analysis of the proposed approach to measure the similarity of video materials. The study focuses on verifying whether or not enough perceptual hash is based on an average color and a histogram to find similar footage.

The research objective. Find the measure of the similarity of two video files.

The statement of basic materials. The analysis of joint use of perceptual hash and histogram as a video frame characteristic. An approach is described for the processing of input video materials for finding a measure of similarity.

Conclusions. The presented method has good performance. Also, it is not reliable at significant changes in video footage, for example large noises or change the video frame area.

Key words: comparison of video content, perceptual hash, histogram, Hamming distance.

УДК 004.75

Болдак А. О., Лисенко Д. В., Ханін М. Є.

СПОСІБ ОПРАЦЮВАННЯ ПРОПУСКІВ В РЕЗУЛЬТАТАХ ЕКСПЕРТНИХ ОПИТУВАНЬ

METHOD OF MISSED ANSWERS PROCESSING IN RESULTS OF EXPERT EVALUATION

У статті запропонований новий спосіб опрацювання пропусків в результатах експертних опитувань, який мінімізує втрати об'єму вибірки при збереженні точності експертних оцінок. Програмна реалізація запропонованого способу використовується в системі екпертних опитувань, розробленій на розподіленій платформі опрацювання даних.

Ключові слова: експертне опитування, заповнення пропусків, об'єм вибірки, програмне забезпечення, розподілене опрацювання даних.

Актуальність дослідження. Експертне оцінювання [1] як процедура оцінювання деякого об'єкта на основі сукупності суб'єктивних оцінок експертів сьогодні широко використовується для вирішення завдань підтримки прийняття рішень в області маркетингу, економічного аналізу [2] та таких нових областях, якими є, наприклад, прогнозування та передбачення розвитку соціо-економічних систем [3].

Постановка проблеми. Не зважаючи на те, що існує достатній спектр способів організації процедури експертного опитування (метод асоціацій, метод парних порівнянь, індивідуальне опитування, метод «Дельфі» та ін.) та відповідних методів об'єктивізації експертних оцінок, об'єм вибірки, тобто потужність множини відповідей експертів на запитання, є одним з найважливіших факторів, що впливають на точність оцінки. Тому на підготовчій стадії експертного опитування використовують один з підходів до вибору кількості експертів, наприклад таких, як [4]:

- довільний – об'єм вибірки обирається довільно без розрахунків точності такого оцінювання;
- довірчий інтервал – вказується необхідна точність оцінювання, після чого розраховується об'єм вибірки, за якої досягається задана точність;
- статистичний аналіз – вказується необхідна точність для окремих груп, які формуються на основі віку, статі, рівню освіти та ін., після чого розраховується необхідний загальний об'єм вибірки.

Але реальний об'єм вибірки зменшується в залежності від кількості пропусків у відповідях експертів та вибраного способу опрацювання цих пропусків, наприклад таких як [5]: методи зважування та заміщення, логічний метод, заповнення середніми або випадковими значеннями, заміна експертів. Використання вказаних методів опрацювання пропусків в реальних умовах викликає певні труднощі, пов'язані з необхідністю залучення занадто великої кількості експертів, організації повторних опитувань та недостатньої точності результатів експертизи.

Постановка завдання. Метою роботи є зменшення втрат об'єму вибірки при збереженні точності оцінок експертів за рахунок розробки та використання нового способу опрацювання пропусків в результатах експертних опитувань.

Викладення основного матеріалу. Пропонований спосіб заснований на гіпотезі, що якщо респонденти в рамках певної тематики дають близькі відповіді на

питання однієї множини, то їхні відповіді на питання тієї ж тематики з іншої множини теж будуть близькими. Для оцінки «близькості» між респондентами введемо формальну модель: респондент – точка багатовимірного простору відповідей.

Для прикладу візьмемо наступну матрицю відповідей:

1, n, n, 1, n, n, 1, 0, 1, 1, n, n, 0, 1, n, n, 0, n, n, n, 0, n, n, n, 0, 1, 0
0, 1, n, 0, n, 0, n, 1, 0, 1, n, 0, n, 1, 1, n, n, 1, n, 1, 1, n, 1, 0, 0, 0, 0
0, n, 0, n, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, n, 0, 0, 0, n, 0, 0, 0, 0, n, 0, 1, 0
1, 0, 0, n, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0
0, 0, 1, 0, 0, 1, 0, 0, 0, n, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1
n, 0, 1, 0, n, 1, 0, 0, n, n, 1, 0, 0, 0, 0, 0, 0, n, 0, 0, 1, 1, 0, 1, 1, 1, n,

де кожний рядок – це відповіді одного експерта, а кожний стовпець – одне питання.

1 та 0 – представлення відповідей у просторі, n – означає відсутність відповіді.

Для розбиття на групи зі схожими відповідями використаємо проекцію на питання, на які всі експерти дали відповідь:

0, 1, 0, 0, 1
1, 1, 1, 0, 0
1, 1, 0, 0, 1
0, 1, 0, 0, 1
0, 0, 1, 1, 0
0, 0, 1, 1, 1

Тоді можна ввести деяку функцію відстані $l(x_i, x_j)$ між точкамиx_i та x_j. В разі ортогоналізації цього простору за допомогою метода головних компонент [6] в якості $l(x_i, x_j)$ можна використати евклідову відстань:

$$l(x_i, x_j) = \sqrt{\sum_{k=1}^{n}(p_{ik} - p_{jk})^2} \quad (1)$$

де x_i, x_j– точки простору, n – його розмірність, p_{ik} – k-та координата відповідної точки.

Тепер необхідно розбити експертів на групи зі схожими відповідями. Для цього використаємо метод k-середніх [7] (англ. k-means) та функцію відстані (1).

Для визначення кількості груп використовується *elbowmethod*[8]. Ідея методу полягає в тому, що ми обраховуємо *k-means* [7] для набору даних з різними k в якомусь діапазоні, та для кожного k обчислюємо *sumofsquarederrors* (*SSE*). Тоді обирається мінімальне k таке, при якому *SSE* насадовольняє.

Отримані кластери (останній стовпець – номер кластеру):

1, n, n, 1, n, n, 1, 0, 1, 1, n, n, 0, 1, n, n, 0, n, n, n, 0, n, n, n, 0, 1, 0, **0**
0, 1, n, 0, n, 0, n, 1, 0, 1, n, 0, n, 1, 1, n, n, 1, n, 1, 1, n, 1, 0, 0, 0, 0, **1**
0, n, 0, n, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, n, 0, 0, 0, n, 0, 0, 0, 0, n, 0, 1, 0, **0**
1, 0, 0, n, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, **0**
0, 0, 1, 0, 0, 1, 0, 0, 0, n, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, **2**
n, 0, 1, 0, n, 1, 0, 0, n, n, 1, 0, 0, 0, 0, 0, 0, n, 0, 0, 1, 1, 0, 1, 1, 1, n, **2**

Розглянемо кластер **0**:

1, n, n, 1, n, n, 1, 0, 1, 1, n, n, 0, 1, n, n, 0, n, n, n, 0, n, n, n, 0, 1, 0, **0**
0, n, 0, n, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, n, 0, 0, 0, n, 0, 0, 0, 0, n, 0, 1, 0, **0**
1, 0, 0, n, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, **0**

Для кожного питання, у якому є пропущенні відповіді, розбиваємо експертів на 2 групи: ті, які відповіли, і ті, які пропустили дане питання.

Візьмемо друге питання. На нього відповів тільки третій респондент цього кластера, у матриці ця відповідь представлена нулем, який і є середнім значенням.

Розглянемо третє питання. На нього відповіли 2 та 3 респонденти, тоді середнє значення $\frac{0+0}{2} = 0$. При необхідності це значення округлюється.

Отримане середнє значення присвоюється іншим експертам цього кластера для цього питання.

Може бути випадок, коли жоден респондент даного кластеру не відповів на питання. Тоді ці поля залишаються незаповненими. Такий випадок є, наприклад, у другому кластері:

0, 0, 1, 0, 0, 1, 0, 0, 0, ***n***, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1, **2**
n, 0, 1, 0, n, 1, 0, 0, n, ***n***, 1, 0, 0, 0, 0, 0, 0, n, 0, 0, 1, 1, 0, 1, 1, 1, n, **2**

Зробивши аналогічні дії з кожним питанням кожного кластеру, отримуємо наступну матрицю:

1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0
0, 1, n, 0, n, 0, n, 1, 0, 1, n, 0, n, 1, 1, n, n, 1, n, 1, 1, n, 1, 0, 0, 0, 0
0, 0, 0, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0
1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0
0, 0, 1, 0, 0, 1, 0, 0, 0, n, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1
0, 0, 1, 0, 0, 1, 0, 0, 0, n, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 1

Якщо відкидати експертів, які відповіли не на всі питання, то об’єм вибірки буде становити $\frac{3}{6} * 100\% = 50\%$ від початкового об’єму. Якщо відкидати питання, на які відповіли не всі експерти, то об’єм вибірки буде становити $\frac{17}{27} * 100\% \approx 63\%$.

Тепер розглянемо початкову матрицю:

1, n, n, 1, n, n, 1, 0, 1, 1, n, n, 0, 1, n, n, 0, n, n, n, 0, n, n, n, 0, 1, 0
0, 1, n, 0, n, 0, n, 1, 0, 1, n, 0, n, 1, 1, n, n, 1, n, 1, 1, n, 1, 0, 0, 0, 0
0, n, 0, n, 0, 1, 1, 1, 1, 1, 0, 0, 0, 1, n, 0, 0, 0, n, 0, 0, 0, 0, n, 0, 1, 0
1, 0, 0, n, 0, 1, 0, 0, 1, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0
0, 0, 1, 0, 0, 1, 0, 0, 0, n, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 1
n, 0, 1, 0, n, 1, 0, 0, n, n, 1, 0, 0, 0, 0, 0, 0, n, 0, 0, 1, 1, 0, 1, 1, 1, n

При відкиданні експертів у даному прикладі об’єм вибірки буде становити 0%, а при відкиданні питань $\frac{5}{27} * 100\% \approx 18{,}5\%$.

Висновки. Таким чином, запропонований спосіб опрацювання пропусків у відповідях експертів дозволяє зменшити втрати об’єму вибірки та забезпечити точність більшу у порівнянні з заповненням пропусків середніми значеннями. Обмеженням запропонованого методу є недостатня кількість запитань, на які відповіли всі респонденти (кількість таких питань повинна перевищувати кількість респондентів).

Програмна реалізація запропонованого способу використовується в системі експертних опитувань на розподіленій платформі опрацювання даних DJ [9].

В подальшому планується розширення запропонованого способу на експертні оцінювання у номінальних та ординальних шкалах.

Список використаних джерел

1. А.И.Орлов (2002) Экспертные оценки. Учебное пособие
2. Damianos P. Sakas, Nikolaos Konstantopoulos (2010) Marketing and Management Sciences: Proceedings of the International Conference on ICMMS 2008
3. M. Zgurovsky (2016) Foresight and construction of the strategies of socio-economic development of Ukraine
4. Н.Г.Каменева, В.А.Поляков (2011) Маркетинговые исследования ст. 78-87.
5. Л.Р. Мухаметова (2009) Методы выборочных обследований ст. 55-57.
6. Jonathon Shlens (2014) A Tutorial on Principal Component Analysis
7. Steinhaus H. (1956). Sur la division des corps materiels en parties. Bull. Acad. Polon. Sci., C1. III vol IV: 801-804.
8. Usingtheelbowmethodtodeterminetheoptimalnumberofclustersfork-meansclustering [Електронний ресурс]. - Режим доступу: bl.ocks.org (дата звернення 24.04.2018) – Назва з екрана.
9. ЕФЕКТИВНІСТЬ СУСПІЛЬНИХ ПЕРЕТВОРЕНЬ. II ТУР[Електронний ресурс]. - Режим доступу: https://dj-next.herokuapp.com/app/Social%20Transformations%20II/ (дата звернення 24.04.2018) – Назва з екрана.

ДОВІДКА ПРО АВТОРІВ

Болдак Андрій Олександрович – доцент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Boldak Andriy – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: boldak.andrey@gmail.com

Лисенко Дмитро Вадимович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Lysenko Dmytro – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: dimalysenko1206@gmail.com

Ханін Максим Євгенійович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Khanin Maxim – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: Maksimkhanin.kpi@gmail.com

A.O. Boldak, D.V. Lysenko, M.E. Khanin

METHOD OF MISSED ANSWERS PROCESSING IN RESULTS OF EXPERT EVALUATION

Relevance of the research topic. Expert evaluation today is widely used to solve decision-making tasks in the field of marketing, economic analysis and such new areas as prediction of the development of socio-economic systems.

Target setting. The actual sample size is reduced, depending on the number of passes in the expert's responses and the chosen method for processing these passes, such as: weighing and substitution methods, logical method, filling in average or random values, replacing experts. The purpose of the work is to reduce the sample size loss while maintaining the accuracy of expert assessments by developing and using a new way of processing passes in the results of expert evaluation.

The statement of basic materials. This method suggests filling passed answers in expert evaluation by the average of groups. For a breakdown into groups, a formal model is introduced: the respondent is the point of a multidimensional response space. In this space, the k-means algorithm is used. To determine the value of k, use the elbow method.

Conclusions. The proposed method for processing passes in the expert's answers allows to reduce the loss of sample size and to provide greater accuracy compared with the filling of the passes with average values.

Key words: expert evaluation, filling of passes, sample volume, software, distributed data processing.

УДК 004.8

Андрій Логвинчук, Олексій Алещенко, Олександр Зальотнов

ЗАСТОСУВАННЯ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ ДО МОДЕЛЮВАННЯ ЕЛЕКТРИЧНИХ КІЛ ЗМІННОГО СТРУМУ

APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS TO ALTERNATING CURRENT CIRCUITS MODELLING

У даній статті розглядається моделювання електричних кіл змінного струму. Метою моделювання є передбачення властивостей даного кола. Для моделювання застосовується нейронна мережа структури "багаторівневий персептрон".

Ключові слова: штучні нейронні мережі, електричні кола, моделювання.

Рис.: 1. Табл.: 2. Бібл.: 5.

This paper deals with the problem of modelling AC electrical circuits. The objective of modelling is prediction of different properties of a given electrical circuit. The model used is a multilayer perceptron artificial neural network.

Key words: artificial neural networks, electrical circuits, modelling.

Fig.: 1. Tabl.: 2. Bibl.: 5.

Актуальність теми дослідження. Вчені та інженери у галузі мікроелектроніки зацікавлені у можливості використання швидких та точних моделей для дослідження і тестування своїх розробок.

Постановка проблеми. Існуючі підходи до моделювання електричних кіл, такі як використання фізичних моделей із застосуванням теорії електромагнетизму або еквівалентні перетворення електричних кіл [1], зазвичай супроводжуються складними аналітичним розрахунками. Реалізація таких моделей за допомогою сучасної обчислювальної техніки не є ефективною.

Аналіз останніх досліджень та публікацій. В останніх працях присвячених штучним нейронним мережам [2, 3] розглядається можливість застосування їх до моделювання фізичних процесів в електричних колах. Зокрема, у праці [2] проводиться огляд, аналіз та виділяються перспективи розвитку штучних нейронних мереж. У роботі [3] досліджується вплив структури штучної нейронної мережі та параметрів її навчання на точність моделювання електричних елементів кола на прикладі джерела струму для контактного зварювання. Робиться висновок про те, що застосування нейронних мереж для подібного роду моделювання є виправданим через низьку похибку, якою можна знехтувати, та універсальність, що дозволяє працювати як з різними значеннями діючої напруги, так і з різними за номіналом електричними навантаженнями.

Виділення недосліджених частин загальної проблеми. Відсутність стандартизованих інструментів моделювання та готових бібліотек електричних компонентів перешкоджають широкому використанню даної методики.

Постановка завдання. Метою цієї статті є дослідження можливості розробки моделей електричних кіл або компонентів, які ефективно використовують обчислювальні ресурси та мають достатню точність моделювання. Довільна ділянка електричного кола може бути представлена у вигляді чотириполюсника. Таким чином, ця стаття присвячена розробці моделі на прикладі польового транзистора.

Викладення основного матеріалу. На сьогоднішній день існує два основні підходи до моделювання електромагнітних процесів. Перший тип – це детальне моделювання, що спирається на фундаментальні фізичні закони, визначені системою рівнянь Максвелла і виконується аналітично. Сучасна обчислювальна техніка накладає певні обмеження на такий тип моделювання, оскільки для них доводиться застосовувати чисельні методи, що є надзвичайно неефективним використанням обчислювальних і часових ресурсів, хоча вони і демонструють високу точність і не покладаються на дані, отримані експериментальним шляхом [1].

Другий тип моделювання базується на еквівалентних перетвореннях електричних кіл, він поєднує різного роду евристики і спрощені теорії з накладанням певних обмежень. Такі моделі потребують значної кількості експериментальних даних, мають гіршу точність, однак обчислюються значно швидше.

Застосування штучних нейронних мереж як окремий підхід до задачі моделювання електричних кіл дозволяє отримувати результати з точністю близькою до аналітичних методів, але значно швидше [4]. Порівняльна характеристика усіх трьох підходів наведена у таблиці 1.

Задачу моделювання електричного кола штучною нейронною мережею можна сформулювати наступним чином: знаючи фізичні параметри польового транзистора, що представлені вектором (1), необхідно визначити амплітуди і фази S-параметрів чотириполюсника, які представлені вектором (2).

$$x = [W, L, a,\ N_d, V_{GS}, V_{DS}]^T \quad (1)$$

$$y = [AS_{11}, PS_{11}, AS_{12}, PS_{12}, AS_{21}, PS_{21}, AS_{22}, PS_{22}]^T \quad (2)$$

Таблиця 1

Порівняння підходів моделювання

Критерій порівняння	Фізичні моделі	Емпіричні та еквівалентні моделі	Штучні нейронні мережі
Швидкість	Повільні	Швидкі	Швидкі
Точність	Висока	Обмежена	Наближається до фізичних моделей
Експериментальні дані	Не потребують	Невелика кількість	Значна кількість для навчання
Теоретичні засади	Рівняння Максвелла	Спрощена теорія, частинні випадки	Не задіяні

На рис. 1 [4] зображено схематично польовий транзистор, який має наступні фізичні параметри: W – ширина затвору, L – довжина затвору, a – товщина каналу, що проводить струм, N_d– густина напівпровідникових домішок, V_{GS} – напруга затвор-виток, V_{DS} – напруга стік-витік.

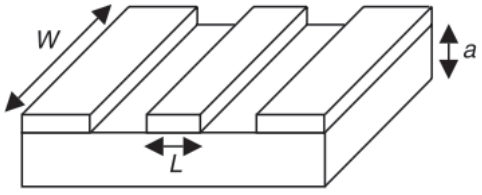


Рис. 1. Схематичне зображення польового транзистора.

Тоді задачу можна виразити формулою:

$$y = f(x), \quad (3)$$

а нейронну мережу для її вирішення – формулою:

$$y = y(x, w). \quad (4)$$

В даному прикладі нейронна мережа зможе представляти поведінку справжнього транзистора тільки після вивчення реального відношення $\boldsymbol{x}$ до $\boldsymbol{y}$ через процес тренування [5]. Пари ($\boldsymbol{x}$, $\boldsymbol{y}$) є зразками даних для навчання нейронної мережі і мають бути виміряні на реальному приладі або згенеровані відповідно до теоретичних засад електромагнетизму.

В даній роботі фізичні параметри $\boldsymbol{x}$ генерувалися випадковим чином, а відповідні їм значення $\boldsymbol{y}$ розраховувались аналітично. Для навчання моделі було використано 5000 зразків даних, значення функції втрат становило 0.18. У таблиці 1 наведено порівняння шуканих параметрів, розрахованих аналітично, та передбачених нейронною мережею.

Таблиця 2

Результати моделювання

Параметр	Передбачене значення	Аналітичне значення
AS_{11}	0.917	0.928
PS_{11}	-63.851	-64.00
AS_{12}	0.026	0.023
PS_{12}	71.200	70.000
AS_{21}	10.813	10.840
PS_{21}	149.008	150.000
AS_{22}	0.570	0.529
PS_{22}	-27.023	-27.000

Висновки. Результати, отримані в даній роботі, демонструють можливість застосування штучних нейронних мереж для симуляції електричних кіл в цілому або їх компонентів окремо за методологією чорної скриньки: модель може прогнозувати різноманітні властивості електричного кола, в залежності від поданих на вхід статичних та динамічних параметрів. Подальше широке застосування такого підходу можливе за умови наявності моделей для стандартних компонентів електричних кіл та наявності зручного інструментарію для роботи з ними. Також варто зазначити перспективи використання інших структур нейронних мереж, таких як мережі радіальних базисних функцій та інші.

Список використаних джерел

1. Harkouss, Y., etal. (May 1998), Modelling Microwave Devices and Circuits for Telecommunications System Design, Proc. IEEEIntl. Conf. NeuralNetworks (pp. 128-133).

2. Е. И. Некрасова (2015), Искусственные нейронные сети, III Научно-техническая конференция с международным участием «Наука настоящего и будущего», сборник материалов конференции (с. 36-37).

3. Карчевский А. П. (2016), Исследование влияния структуры искусственной нейронной сети и параметров ее обучения на точность

моделирования источника питания для контактной сварки, магистерская диссертация (с. 6-8).

4. Zhang, Q. J., and Devabhaktuni, V. K. (April 2003), Neural Networks for Microwave Modeling and Design, IEEE Transactions On Microwave Theory And Techniques, Vol. 51 (pp. 1339-1352).

5. Wang, F. et al. (1999), Neural Network Structures and Training Algorithms for RF and Microwave Applications, Int. Journal of RF and Microwave CAE, Special Issue on Applications of ANN to RF and Microwave Design, Vol. 9 (pp. 216-240).

ДОВІДКА ПРО АВТОРІВ

Логвинчук Андрій Ігорович – студент, кафедра обчислювальної техніки, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського".

Lohvynchuk Andrii – student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: a.logvinchuk-2018@kpi.ua

Алещенко Олексій Вадимович – асистент, кафедра обчислювальної техніки, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського".

Aleshchenko Oleksii – assistant professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: alexey.aleshchenko@gmail.com

Зальотнов Олександр Дмитрович - студент, кафедра обчислювальної техніки, Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського".

Zalotnov Oleksandr– student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: zalotnov.alexander@gmail.com

A. Lohvynchuk, O. Aleshchenko, O. Zalotnov

APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS TO ALTERNATING CURRENT CIRCUITS MODELLING

Target setting. Scientists and engineers in the field of microelectronics are interested in the possibility of using fast and precise models to test their developments.

Actual scientific researches and issues analysis. Existing approaches to modelling electrical circuits such as use of physical models or use of equivalent circuits generally require complex analytical calculations. These are very computationally expensive and cannot be implemented efficiently on modern equipment.

Uninvestigated parts of general matters defining. Lack of standardized modelling tools and components libraries prevent this approach from being mass-used.

Objective setting. The goal of this paper is research on possibility of development of electrical circuit or particular component model, which would be computationally effective and demonstrate desirable precision. This work focuses on development of FET model.

The statement of basic materials. The comparison between existing approaches and neural network modelling is done. The approach to decomposing of original task into abstract problem to be solved by neural network is described. The modelling results display satisfactory precision.

Conclusions. The results obtained in this work demonstrate possibility of neural network usage for modelling electrical circuits or electrical components in a "black box" methodology: given a set of physical parameters, the model can predict various properties of a device that is being modelled. Further use of this approach is possible under condition of existence of standardized models for various standard electrical components and convenient tools to combine them. Also it is worth mentioning that other neural network structure such as radial basis neural networks also can be useful in this field.

Key words: artificialneural networks, electrical circuits, modelling.

UDC 004.82

Данило Сергеєв

ІНТЕГРАЦІЙНА МОДЕЛЬ ПРЕДСТАВЛЕННЯ ЗНАНЬ ДЛЯ СЕМАНТИЧНОГО WEB

INTEGRATION MODEL FOR KNOWLEDGE REPRESENTATION FOR SEMANTIC WEB

Ця стаття представляє інтеграційну модель представлення знань у природно-мовних базах знань, яка дозволяє досягти гарної обчислювальної складності операцій додавання та пошуку даних у базі знань. Така модель може бути використана як проміжна структура даних між спеціалізованими системами обробки тексту та знання-орієнтованими системами в екосистемі Semantic Web.

Ключові слова: квант знань, база знань, обробка природної мови, інтеграційний підхід.

Рис.: 2.Табл.: -. Бібл.: 11.

This paper presents an integration model for knowledge representation in natural language knowledge bases that allows to achieve good computational complexity of operations of adding and searching data in the knowledge base. The model can be used as an intermediate data structure between specialized text-based and knowledge-based systems in Semantic Web ecosystem.

Key words: quantum of knowledge, knowledge base, natural language processing, integration approach.

Fig.: 2. Tabl.: -. Bibl.: 11.

Relevance of research topic. Creating of Semantic Web and instruments based on it has been a well-established area of research for at least last two decades. It remains an active topic, worked on as part of the “Web 3.0” by the W3C committee [1] as well as by independent researchers around the world. Despite this fact, there are still unresolved problems with fully implementing it [2], many of which are linked closely to the data model used in the underlying knowledge base.

Formulation of the problem. Recent research in the field shows promising success in solving some of the underlying problems, namely mining existing Semantic Web for domain-specific tasks [3], fine-tuning user tools in order to retrieve more precise results [4] and creating knowledge bases both for raw natural language data [5] and structured knowledge [6]. However, there still remains the problem of coordinating different parts of the system to achieve cohesive structure of knowledge in it that would allow to coordinate processing of natural language data and structured knowledge. This paper proposes a new model for knowledge representation for natural language knowledge base to be used as a part of Semantic Web ecosystem.

Analysis of recent research and publications. The Semantic Web technologies rely on various data models to store semantically interlinked knowledge, which together can be generalized as knowledge bases, most prominently ontologies, frame networks and semantic networks [7]. On the other hand, data structures for storing text information are mostly based on language-specific syntactical structure, such as parse trees, grammars and n-grams [8]. Both groups of models work with

fundamentally similar types of data, from words to texts in natural language models and from concepts to rules systems in knowledge bases, but they typically are not well-suited to handle the other type.

Both knowledge bases and natural language data structures are used for various high-level tasks, but their functions can be typically brought down to the basic I/O operations of "read-write". Taking into account specifics of aforementioned systems, these operations represent two major functions of a knowledge base:

- reading, or searching and reading, knowledge according to a specific query;
- writing, or adding new knowledge into an existing knowledge base.

Both these functions require additional operations to be executed inside the system, and both natural language and structured knowledge bases introduce additional complexity to these operations.

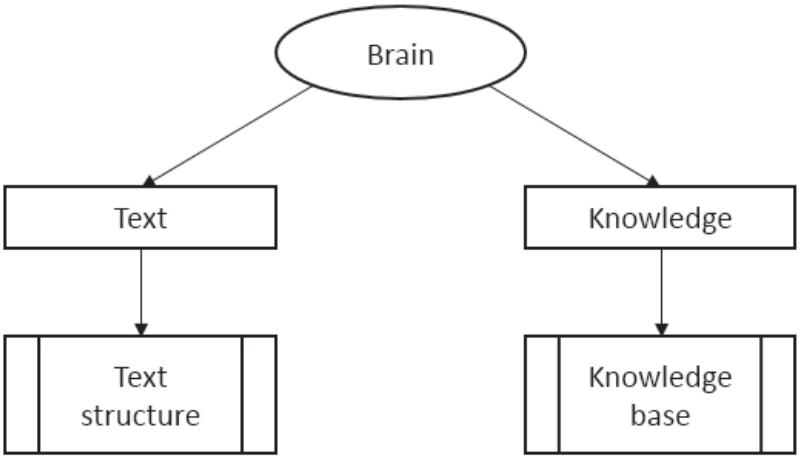


Fig. 1. Data flow in natural language and structured knowledge bases.

Selection of unexplored parts of the general problem. Natural language knowledge bases rely heavily on parsing syntactic structure of the incoming text. This allows for easier addition of new information, as the knowledge structure of incoming text is by definition same as that of the knowledge base, but searching data over such base requires traversing all text-based data structures, which makes applying advanced searching algorithms very hard or outright impossible.

Structured knowledge bases, on the other hand, provide rigid data structure that allows for effective searching, but require any incoming data to be adapted and possibly trimmed in accordance with this structure, making expanding the knowledge base an algorithmically and computationally complex task.

This allows existing knowledge models to be efficient in certain tasks that demand good performance from only certain part of their functions, but effectively prevents any of the models to be used as an average ubiquitous solution.

Presentation of the main material. In this work, a new model of combined knowledge representation is presented that allows to combine best parts of existing models while avoiding their downsides. The model is based on the ideas of integrated approach to modelling of human speech activity, described in [9–11], most notably the proposed basic semantico-syntactic structure, or BSSS, that represents a single "situation", or quantum of knowledge in human brain.

Given that a single BSSS represents a single unit of knowledge, closely tied to corresponding sensory memories, and any fragment of text can be decomposed into several BSSS and relations between them, it can be used as an intermediate structure between knowledge bases of natural language and of structured knowledge.

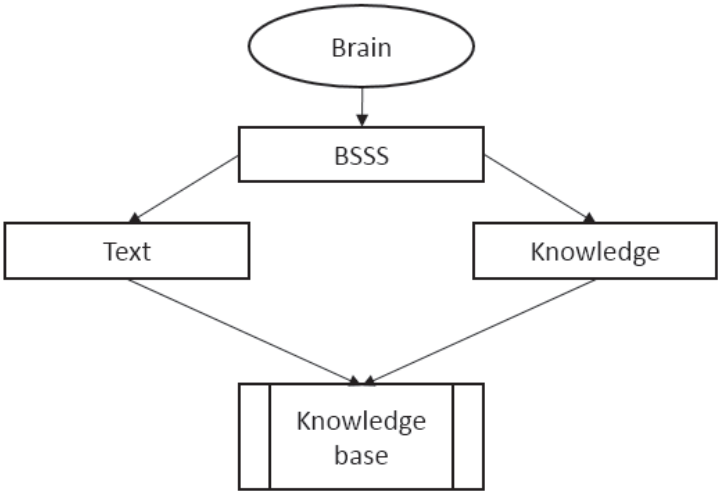


Fig. 2. Data flow in integration model knowledge base.

This model allows to consolidate the knowledge itself in a single, universal structure, filled with BSSS's and relations between them, that provides certain benefits compared to other approaches.

First, like in structured knowledge bases, data is stored in structured, formally well-defined form, which allows for easier processing and automatic manipulation of this data.

Second, like natural language knowledge bases, a model like this provides structural links between the contents of the knowledge base and structure of the text linked to these contents, which preserves easiness of adding new knowledge to the knowledge base from natural language text.

Third, unlike the other approaches, in integrated knowledge base new knowledge is added to an existing structure without altering it – that is, complexity of expanding the base remains relatively constant regardless of its size and doesn't require mandatory rebuilding of previously added data.

It is also worth noting that an integrated knowledge base can be linked to separate natural language and structured knowledge bases, thus allowing to use language-specific language processing tools and domain-specific logic and resolution engine to enhance capabilities of the whole system.

Conclusions. Future work will involve improving building links between natural language constructs, especially multi-word and artificial ones like abbreviations, and complex semantic concepts. More research is also required to create a linguistic processor capable of working with the data model presented in this paper.

References

1. W3.org Semantic web - w3c / W3.org. — 2012.
2. Benjamins V. R. Six challenges for the semantic web / V. R. Benjamins, J. Contreras, O. Corcho, A. Gómez-Pérez // Challenges. — 2004. — Vol. 1, No. i. — P. 1–15.

3. Yong-gui W. Research on semantic web mining / W. Yong-gui, J. Zhen. — 2010. — V1-67-V1-70 p. — ISBN 978-1-4244-7164-5.

4. Ceolin D. Subjective logic extensions for the semantic web / D. Ceolin, A. Nottamkandath, W. Fokkink. — 2012. — 27–38 p.

5. Kaufmann E. Evaluating the usability of natural language query languages and interfaces to semantic web knowledge bases / E. Kaufmann, A. Bernstein // Web Semantics: Science, Services and Agents on the World Wide Web. — 2010. — Vol. 8, No. 4. — P. 377–393.

6. Auer S. Semantic wiki representations for building an enterprise knowledge base / S. Auer, B. Jungmann, F. Schönefeld // Reasoning Web. — 2007. — P. 330–333.

7. Brachman R. Knowledge representation and reasoning / R. Brachman, H. Levesque. — 2004. — ISBN 1558609326.

8. Stavrianou A. Overview and semantic issues of text mining / A. Stavrianou, P. Andritsos, N. Nicoloyannis // ACM SIGMOD Record. — 2007. — Vol. 36, No. 3. — P. 23.

9. Kyslenko Y. I. Back to basics of speech activity / Y. I. Kyslenko // Biologically Inspired Cognitive Architectures. — 2014. — Vol. 8. — P. 46–68.

10. Kyslenko Y. Cognitive architecture of speech activity and modelling thereof / Y. Kyslenko, D. Sergeiev // Biologically Inspired Cognitive Architectures. — 2015. — Vol. 12.

11. Кисленко Ю. І. Архітектура мови / Ю. І. Кисленко. — Київ : ІЗМН, 1998.

ДОВІДКА ПРО АВТОРІВ

Сергеєв Данило Сергійович – асистент, кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

DanyloSergeiev – assistant professor, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: d.sergeiev@gmail.com

Д. С. Сергеєв

ІНТЕГРАЦІЙНА МОЕДЕЛЬ ПРЕДСТАВЛЕННЯ ЗНАНЬ ДЛЯ СЕМАНТИЧНОГО WEB

Актуальність теми дослідження. Створення Semantic Web та інструментів для нього, є добре розвиненою сферою досліджень, як в рамках розробок W3C , так і серед також незалежних дослідників по всьому світу. Незважаючи на це, все ще залишаються невирішеними проблеми з повним його впровадженням, значна частина яких тісно пов'язана зі структурою моделі даних, яка при цьому використовується.

Постановка проблеми. Відсутність моделі представлення знань, яка може однаково якісно працювати з природно-мовними даними та структурованими знаннями.

Аналіз останніх досліджень і публікацій. Недавні дослідження демонструють успіх у вирішенні багатьох задач SemanticWeb, зокрема уточнення запитів для певної галузі пошуку, розробка окремих баз знань для природно-мовних та структурованих знань, взаємодія з зовнішніми системами тощо. При цьому залишається невирішеною проблема координації між різними базами знань в рамках системи.

Постановка завдання. Завданням є розробкацілісної моделі знань, що дозволить координувати обробку природно-мовних даних та структурованих знань.

Викладення основного матеріалу. Представлено модель знань для природно-мовних баз знань, що дозволяє створити базу знань, яка поєднує переваги систем представлення текстових знань та структурованих баз знань і при цьому не наслідує їх недоліки.

Висновки. Розроблена модель дозволяє спростити автоматичну обробку даних у природно-мовних базах знань, зберігаючи повноту знань, зберігаючи постійну складність операції додавання нових знань.

Ключові слова: квант знань, база знань, обробка природної мови, інтеграційний підхід.

УДК 004.8

Ярослав Корнага,
Микола Шиманський,
Андрій Барабаш

ЛОКАЛІЗАЦІЯ ТЕКСТУ НА ЗОБРАЖЕННЯХ ЗА ДОПОМОГОЮ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

TEXT LOCALIZATION ON IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

У статті розглядається питання локалізації тексту на зображеннях складних графічних сцен. Для локалізації застосовуються згорткова нейронна мережа із завчасно натренованою мережею для виділення ознак на основі моделі VGG-16. Для тренування використано набори зображень із ICDAR 2011 та ICDAR 2013.

Ключові слова:локалізація тексту, обробка зображень, машинний зір, нейронні мережі.

Рис.: 1. Табл.: 1. Бібл.: 3.

This paper deals with the problem of scene text localization. Convolutional neural network was used for this task. Pretrained model based on VGG-16 was used for feature extraction. Datasets from ICDAR and ICDAR 2013 were used for training.

Key words: text localization, image processing, computer vision, neural networks.

Fig.: 1. Table.: 1. Bibl.: 3.

Актуальність теми дослідження. Вирішення задачі локалізації тексту на зображеннях складних графічних сцен є необхідною умовою розпізнавання тексту на таких зображеннях, що може мати застосування в системах машинного зору.

Постановка проблеми. Текст на зображенні у повсякденному житті може мати різний колір та розмір шрифту, він може бути рукописним, що само собою означає, що написання може бути дуже варіативним по своїй формі. Колір фону та освітлення також будуть відрізнятись від зображення до зображення. Багатошарові нейронні мережі дозволяють встановлювати закономірності у множині різних можливих вхідних даних, що робить доцільним їх використання для такого роду задач.

Аналіз останніх досліджень і публікацій. За останні роки в області машинного зору було проведено ряд досліджень, які розглядають проблему локалізації тексту. У [1] автори пропонують метод виділення екстремальних регіонів, на основі класифікації яких проводиться пошук окремих символів. Після цього знайдені літери групуються повним перебором.

У [2] пропонується застосування неспеціалізованого генератора для формування регіонів-кандидатів, що можуть містити слова. Після цього застосовується конволюційна нейромережа для регресії текстових прямокутників.

Також у [3] описується метод виділення цілих текстових рядків з використанням їх симетричних характеристик.

Недосліджена частина проблеми. Дана робота присвячена вивчення та аналізу можливостей застосування багатошарових згорткових нейронних мереж для виділення як друкованого, так і рукописного тексту на зображеннях складних графічних сцен.

Постановка завдання. Завданням є тренування повністю згорткової нейронної мережі для виділення текстових областей на зображеннях складних графічних сцен.

Тренування нейромережі. В області машинного зору багатошарові згорткові нейромережі широко застосовуються для задач класифікації об'єктів на зображеннях. Перший шар згортки обчислює на основі оригінального зображення матрицю ознак – набір числових значень, що зазвичай характеризують відповідні їм області на зображенні. Кожен наступний згортковий шар обчислює додатковий набір ознак, на основі виходу попереднього шару. На даний момент існує багато різноманітних класифікаторів зображень на основі згорткових нейромереж, для яких доступні завчасно натреновані моделі. Успішність цих класифікаторів вказує на ефективність шарів нейромережі, що відповідають за виділення ознак. Для прискорення процесу навчання натреновані ваги з цих нейромереж можуть бути використані для ініціалізації ваг в моделі, що буде спеціалізована на класифікації областей зображення на текстові та нетекстові області. На рис. 1 зображена архітектура запропонованої нейронної мережі. Процес обробки зображення починається із виділення ознак на завчасно натренованих шарах, які коректуються на згортковому шарі Конв.1. На наступних шарах з Конв. 2 по Конв. 6 відбувається поступове зменшення розмірності матриці ознак за допомогою максимізаційного агрегування для можливості виділення текстових областей з різними розмірами шрифту: при обробці матриці ознак з великими розмірами більша вірогідність знайти дрібний текст і навпаки. У шарах текстових блоків проводиться обчислення координат кандидатів текстових областей. На останньому етапі обробки проводиться фільтрація масиву кандидатів для відсіювання занадто маловірогідних гіпотез та дублікатів.

Проведення експерименту. Для тренування і тестування системи було використано зображення із баз даних ICDAR 2011 та ICDAR 2013. Дані з цих вибірок використовуються у змаганнях з локалізації, сегментації та розпізнавання тексту на зображеннях. Нейронна мережа була натренована на зображеннях розміром 300×300 з використанням градієнтного спуску. Кількість зображень у тренувальній вибірці: 100000.

Таблиця 1.

Показники якості роботи системи на тестових вибірках

	ICDAR 2011	**ICDAR 2013**
Точність	0,85	0,86
Відгук	0,8	0,81

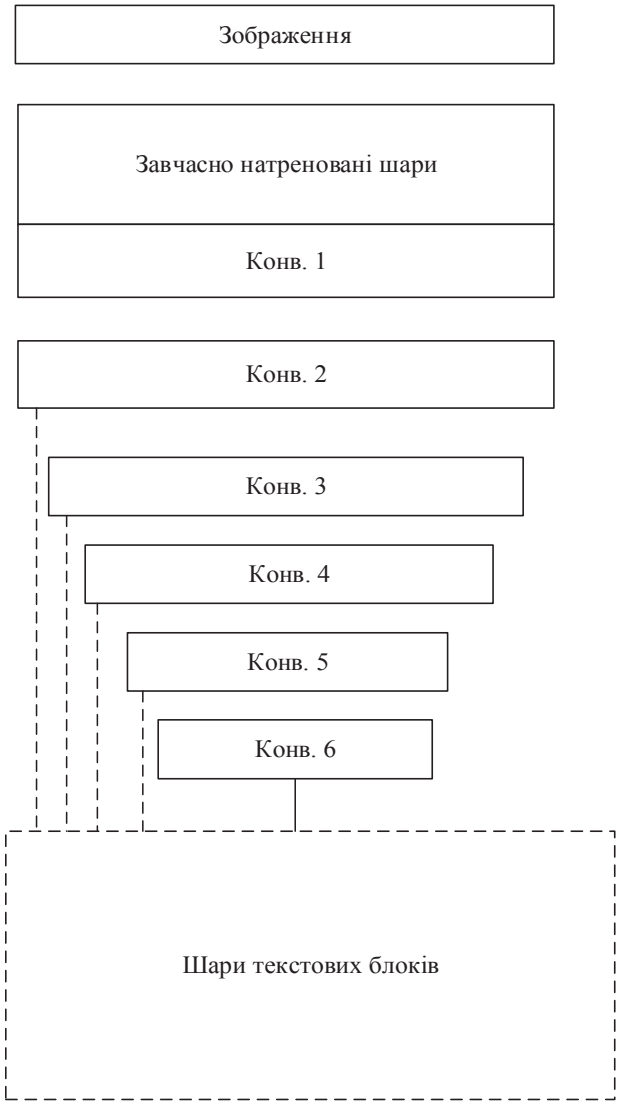


Рис. 1. Архітектура нейромережі.
Штрихованими лініями позначено використання
у шарах текстових блоків ядра згортки розміром 1×5, а суцільною – 1×1

Висновки. Запропонована модель дозволяє забезпечити локалізацію тексту на зображеннях складних графічних сцен. В поєднанні з існуючими методами розпізнавання тексту можна отримати повну систему розпізнавання тексту, яка дозволяє виділяти із зображення текст і перетворювати його у формат, придатний для машинної обробки. Модель може бути використана для локалізації як друкованого, так і рукописного тексту.

Список використаних джерел

1. Neumann, L., and Matas, J. 2012. Real-time scene text localization and recognition. In Proc. CVPR, 3538–3545.

2. Jaderberg, M.; Simonyan, K.; Vedaldi, A.; and Zisserman, A. 2016. Reading text in the wild with convolutional neural networks. IJCV 116(1):1–20.

3. Zhang, Z.; Shen, W.; Yao, C.; and Bai, X. 2015. Symmetrybased text line detection in natural scenes. In Proc. CVPR, 2558–2567.

ДОВІДКА ПРО АВТОРІВ

Корнага Ярослав Ігорович - кандидат технічних наук, доцент кафедри технічної кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Kornaga Yaroslav - associate Professor of the Department of Technical Cybernetics, PhD of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».

E-mail:slovyan_k@ukr.net

Шиманський Микола - студент кафедри Технічної Кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

ShymanskyiMykola - student of the Department of Technical Cybernetics of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».

Барабаш Андрій - студент кафедри Технічної Кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Barabash Andrii - student of the Department of Technical Cybernetics of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute».

YaroslavKornaga, MykolaShymanskyi, Andrii Barabash

TEXT LOCALIZATION ON IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

Target setting. Scene text recognition can have many use cases in computer vision systems and text localization is a major part of this problem

Problem statement. Absence of effective method for text localization in natural scene images.

Actual scientific researches and issues analysis. There are different studies that cover various approaches to text localization task such as usage of maximally stable extremal regions, symmetric features and others.

Uninvestigated parts of general matters defining. This paper is dedicated to study and analysis of possibilities of deep convolutional neural networks for printed and handwritten text localization alike.

The research objective. The task at hand is to create and train a neural network model that can be used for effective text localization on natural scene images.

The statement of basic materials. This paper describes an architecture of a fully convolutional neural network that is proposed for the text localization task. The benefits of using pretrained models as feature extractors were described. Experimental results were presented.

Conclusions. Proposed model provides an effective text localization for natural scene images. It can be used with text recognition systems for text extraction from images and its further processing. This model can be trained for both printed and handwritten text localization.

Key words: text localization, image processing, computer vision, neural networks.

УДК 004.048

Ярослав Корнага, Юрій Тільняк

ДІАГНОСТИКА ТЕХНІЧНОГО СТАНУ АВТОМОБІЛІВ З ВИКОРИСТАННЯМ ШТУЧНИХ НЕЙРОННОЇ МЕРЕЖІ

DIAGNOSTICS OF THE TECHNICAL STATE OF AUTOMOBILES USING THE ARTIFICIAL NEURAL NETWORK

У статті розглянуто основні механізми і принципи роботи сучасної технічної діагностики автомобілів. Показано необхідність впровадження нових технологій для планування і підвищення ефективності технічної експлуатації з описом побудови експертної системи.Обґрунтовано важливість підвищення якості аналізу діагностичної інформації впровадженням штучних нейронних мереж в програму блоку управління.

Ключовіслова:самохідний колісний транспортний засіб, динамічні навантаження,систематехнічноїдіагностики, штучні нейронні мережі, електронний блок керування.

Рис.: 3. Табл.: 1. Бібл.: 5.

In paper, the basic mechanics and principle of robotics of the automobile technical diagnostics of cars are rooted. It shows the need for new technologies for planning and implementation of technology performance in the field of forensic expert systems. The importance of writing an anchor analysis of diagnostic information in the form of neural units in the program of management is fixed.

Keywords: self-propelled wheeled vehicle, dynamic loads, system of technical diagnostics, artificial neural networks, electronic control unit.

Fig.: 3. Tabl.: 1. Bibl.: 5.

Постановка завдання. У сучасних автомобілях без комп'ютерної діагностики, визначення стану автомобіля стає неможливим. Електронні схеми, розроблені для управління, обладнані системою самодіагностики, яка створена для того, що б інформувати водія про несправності. У процесі роботи автомобіля, на різних режимах (запуск двигуна, прогрів, розгін і гальмування, холостий хід), безперервно йде зчитування показників декількох десятків датчиків.

Актуальність теми дослідження.На показники датчиків система реагує в залежності від типу сигналу, один сигналізує про пошкодження датчика і розрив ланцюга, другий про вихід рівня сигналу за межі закладені в пам'яті блоку керування, що свідчить про вихід самого датчика з ладу, а система діагностики зберігає код помилки в довготривалій пам'яті для можливості аналізу фахівцем. Вихід за рамки параметрів середовища датчика свідчить про позаштатну роботу, що потребує рішення про подальші дії керування автомобілем [3].

Постановка проблеми. Аналіз роботи систем діагностики технічного стану автомобіля виявив ряд недоліків: по-перше, вони не враховують динамічне навантаження, які діють на робочі органи. По-друге, не виконується одна з головних задач технічного діагностування – прогнозування технічного стану автомобіля. Система, яка існує, не спиралась на машинне навчання і спроможна лише сповіщати про наявність помилки в блоках керування чи механічних елементах [4]. Таким чином, актуальним є прогнозування технічного стану автомобіля в процесі виконання ним технологічних операцій, тобто в динаміці роботи [2].

Мета дослідження. Ці недоліки можуть бути усунені, якщо в режимі реального часу, за даними поточних вимірювань критичних параметрів, ідентифікувати динамічні навантаження, відповідні їм динамічні режими роботи і технічний стан автомобіля. За допомогою експертної бази даних в якій набрана статистика про діючі навантаження, а також напрацювання і залишковий ресурс автомобіля і його систем, можна здійснювати прогнозування технічного стану автомобіля. Ці завдання можуть бути вирішені за рахунок впровадження штучних нейронних мереж [1].

Нечіткі нейронні мережі в технічній діагностиці транспортного засобу. Підготовка відповідної бази даних, по якій буде проводитися навчання, а в подальшому експлуатація нечіткої нейронної мережі (власне діагностування несправностей), вимагають проведення таких вимірювань, які будуть однозначно свідчити про фактичний стан автомобіля. При цьому слід виділити ті фрагменти характеристик, які відрізняються одна від одної. Для досягнення цієї мети можуть виконуватися будь-які операції (як лінійні, так і нелінійні) на всій базі даних.

Схема застосування штучної нейронної мережі для виявлення несправностей представлена на Рис. 1

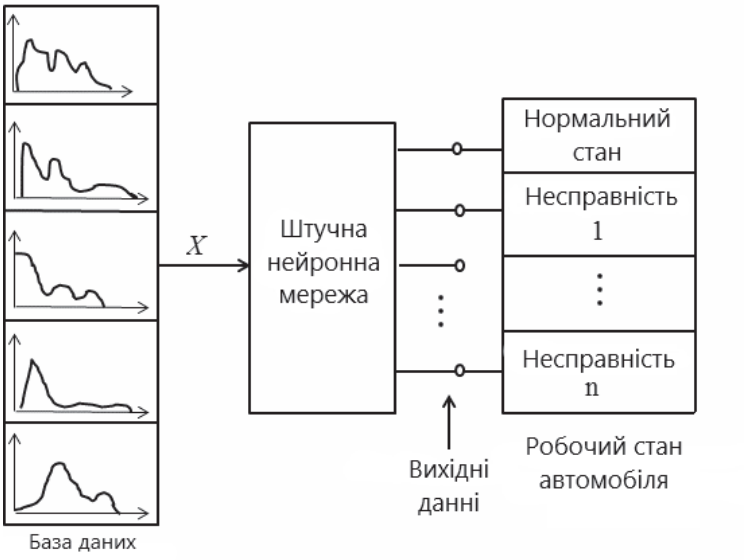


Рис. 1. Схема застосування штучної нейронної мережі для виявлення несправностей в автомобіля

Найважливіша властивість штучних нейронних мереж свідчить про їх великий потенціал в галузі прогнозування відмов і в можливості паралельної обробки інформації усіма нейронами [4]. Завдяки цій здатності, при великій кількості між нейронних зв'язків, досягається одночасно обробка значного обсягу вимірювальної інформації в реальному масштабі часу.

За рахунок використання штучної нейронної мережі сервіс здатний визначити, коли саме несправність стане критичною. У той же час система зможе вказати часовий інтервал, протягом якого це, швидше за все, відбудеться.

Структура нейронної мережі. На рис. 2 показана структура нейромережевої моделі ЕБК, побудованої на основі багатошарового рекурентного персептрона. Кількість нейронів у прихованому шарі вибирається виходячи з мінімальної середньоквадратичної помилки навчання нейронної мережі.

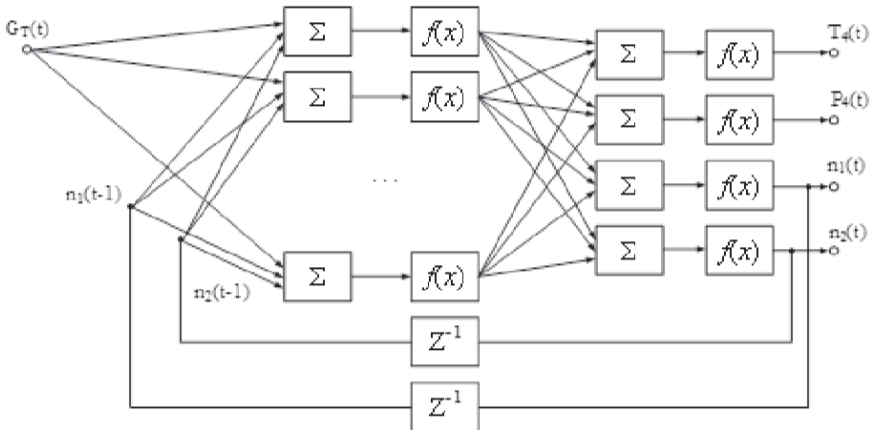


Рис. 2. Структура нейромережевої моделі

Кількість нейронів у прихованому шарі вибирається виходячи з мінімальної середньоквадратичної помилки навчання нейронної мережі (Рис. 3).

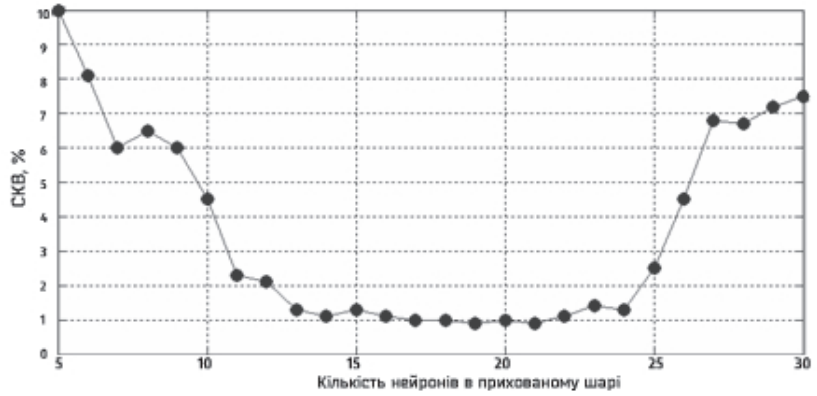


Рис. 3. Помилка нейронної мережі в залежності від кількості нейронів прихованого шару

Помилки 1 і 2 роду. Порівняльний аналіз точності класичного і нейронечітких методів класифікації відмов наведено в табл. 1. У ній зображені ймовірності помилок 1 і 2 роду при класифікації відмови вимірювального каналу температури газів за турбіною, дефектів компресора і камери згоряння. Дані, наведені в табл. 1, доводять, що інтелектуальні методи більш ефективно і якісно здійснюють процес ідентифікації відмов у вузлах і агрегатах двигуна внутрішнього згорання.

Таблиця 1.

Порівняльний аналіз методів

Метод класифікації відмов	**Імовірність помилки визначення відмови, %**					
	Відмова вимірювального каналу t4		**Дефект компресора низького тиску**		**Дефект камери Згоряння**	
	Помилка 1 роду	**Помилка 2 роду**	**Помилка 1 роду**	**Помилка 2 роду**	**Помилка 1 роду**	**Помилка 2 роду**
До пусковий контроль	1,21	0,75	1,72	1,17	2,41	1,93
Інтелектуальний метод	0,47	0,27	0,56	0,41	0,77	0,55

Висновки. На основі аналізу особливостей експертної системи діагностики несправностей СКТЗ і методів аналізу нейронної мережі розроблена експертна система діагностики несправностей автомобілів. Проведений аналіз алгоритмів роботи нечітких нейронних мереж свідчить про те, що ці мережі, за певних обмежень, з успіхом можуть використовуватися у прогнозуванні технічного стану в бортових діагностичних системах.

Результати показують властивість штучних нейронних мереж та їх великий потенціал в галузі прогнозування відмов, можливості паралельної обробки інформації усіма нейронами. Завдяки цій здатності, при великій кількості навчальної вибірки, досягається збільшення точності в прогнозуванні несправностей СКТЗ.

Генеруючи сигнали ЕБК двигуна, в якості прикладу, в ході роботи було встановлено наступні переваги використання нейромережевих алгоритмів для вирішення вищезазначених завдань: підвищення ефективності діагностики, простота навчання і до навчання. Обґрунтовано можливість використання штучних нейронних мереж в системах технічної діагностики СКТЗ.

Список використаних джерел

1. Вікторова О.В. Використання «м'яких» обчислень в інтелектуальних інформаційно-вимірювальних системах дорожніх машин / О. В. Вікторова, А.О. Коваль // Вестник ХНАДУ: сб. науч. тр. – 2011. – Вып. 53. – С. 111–117.

2. Вікторова О.В. Методика побудови функцій приналежності інформативних параметрів динамічних режимів роботи дорожньої машини / Восточно-европейский журнал передовых технологий. – 2011. – № 5/3(53). – С. 11–15.

3. Жернаков С.В. К вопросу о построении гибридных нейро-нечетких экспертных систем кдиагностики и контроля ГТД. Управление в сложных системах. – Уфа: УГАТУ, 1999. – С. 119-126.

4. Круглова Т. Н. Нечеткий экспертный метод диагностирования технического состояния очистного комбайна // Наукові праці Дон НТУ. – 2010. – № 18(172). – С. 179–185.

ДОВІДКА ПРО АВТОРІВ

Корнага Ярослав Ігорович –кандидат технічних наук, доцент,доцент кафедри технічної кібернетикиНаціонального технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Kornaga Yaroslav Igorovich –Candidate of Technical Sciences, Associate of professor Technical Cybernetics Department of the National Technical University of Ukraine " Igor Sikorsky Kiev Polytechnic Institute "

E-mail:slovyan_k@ukr.net

Тільняк Юрій Ярославович –магістр кафедри технічної кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»

Tylnyak Yuriy Yaroslavovych – Master of the Department of Technical Cybernetics of the National Technical University of Ukraine "Igor Sikorsky Kiev Polytechnic Institute"

E-mail: georgetilnuak@gmail.com

Yaroslav Kornaga, Yuri Tilnyak

DIAGNOSTICS OF THE TECHNICAL STATE OF AUTOMOBILES USING THE ARTIFICIAL NEURAL NETWORK

Relevance of research topic. In modern cars without computer diagnostics, it is impossible to determine the state of the car. Electronic circuits designed for control are equipped with a self-diagnostic system that is designed to inform the driver of a malfunction. In the process of operation of the car, in different modes (engine start, warm up, acceleration and inhibition, idle run), continuous reading of indicators of several dozens of sensors.

Formulation of the problem. Analysis of the work of diagnostic systems of the technical condition of the car revealed a number of shortcomings: firstly, they do not take into account the dynamic load acting on the working bodies. Secondly, one of the main tasks of technical diagnostics is not fulfilled – the prediction of the technical condition of the car. Thus, it is relevant to predict the technical condition of the car in the process of performing its technological operations, that is, in the dynamics of work.

Analysis of recent research and publications. Over recent years, there are more articles devoted to the technical diagnosis of self-propelled wheeled vehicles, in particular, due to the emergence of new methods for diagnosing the use of neural networks.

Selection of unexplored parts of the general problem. This article is devoted to the study and analysis of the proposed approach for diagnosing the technical condition of cars. The research focuses on the study of the use of artificial neural networks and electronic control units.

Setting objectives. Existing systems do not rely on machine learning and are only able to report the presence of errors in control units or mechanical elements. By using an artificial neural network, the service is able to determine when the fault will become critical. At the same time, the system will be able to specify the time interval during which it will most likely happen.

Presentation of the main material. The necessity of introducing new technologies for planning and improving the efficiency of technical operation with the description of constructing an expert system based on technology using the capabilities of neural networks is shown. As an instrument for the practical solution of applied problems in the field of diagnosing and forecasting the performance of a self-propelled vehicle.

Conclusions. The importance of improving the quality of analysis of diagnostic information by introducing artificial neural networks into the control unit program is substantiated.

Keywords: self-propelled wheeled vehicle, dynamic load, system of technical diagnostics, artificial neural networks, electronic control unit.

УДК 004.8

Вадим Мухін,
Ярослав Корнага,
Андрій Любівий

МОДЕЛЮВАННЯ Й РОЗПІЗНАВАННЯ ЖЕСТІВ СКЕЛЕТОНУ РУКИ

MODELING AND RECOGNITION OF GESTURE OF HAND SKELETON

В даній статті розглянуті проблеми моделювання й розпізнавання скелетону руки відзнятої з однієї камери без додаткових датчиків за рахунок детектування ключових точок руки та побудови моделі руки з цих точок для класифікації її серед описаних моделей жестів. В результаті було розроблено систему, систему для знаходження ключових точок руки за допомогою нейронної мережі, яка використовуючи їх, створює прототип образу руки й класифікує модель жесту.

Ключові слова: моделювання об’єктів, згорткові нейронні мережі.

Рис.: 5. Бібл.: 4.

In this article, problems of modeling and recognition of a skeleton of a hand from image which is taken with one camera without additional sensors are considered. Detected key points from the image with hand are using for the construction of a model of a hand from these points to classify it among the described gesture models. As a result, a system was developed, a system for finding key points of the hand using a neural network, which uses them, creates a prototype of the image of the hand and classifies the gesture model.

Keywords: modeling of objects, convolutional neural networks.

Fig.: 5. Bibl.: 4.

Актуальність теми дослідження. Проблема моделювання й розпізнавання жестів стає більш актуальною в останні дні у зв'язку зі зростаючою кількістю систем з якими людина може взаємодіяти за допомогою жестів. На сьогоднішній день величезна кількість техніки знаходиться у нас в домі: комп’ютери, телевізори, телефони, планшети й інші. Ми з ними постійно взаємодіємо, і ефективність цієї взаємодії впливає на наше життя. Частіше всього ми з ними взаємодіємо тактильним способом: відчуття доторку й тиску. Ми натискаємо на кнопки на пристрої чи на екран телефона. Але не завжди це можливо, бо в нас можуть бути зайняті руки; або зручно, бо прилад з яким ми взаємодіємо знаходиться на відстані від нас. І тому було б добре взаємодіяти з приладами за допомогою рук, а саме жесті.

Таким чином, проблемою є створення системи, що дасть можливість «спілкуватись» з іншими за допомогою жестів, що підвищить гнучкість й зручність використання. Дана робота присвячена проблемі розпізнавання жестів руки з моделей, які були побудовані за допомогою нейронної мережі.

Постановка проблеми. Через появу нових методів і підходів у галузі штучного інтелекту, тема розпізнавання зображень стала більш актуальною в останні роки.

Аналіз останніх досліджень і публікацій. Існують два основних способи розпізнавання моделей рук. Перший спосіб заснований на одяганні спеціальних електронних пристроїв (спеціальні рукавиці з датчиками), який в основному використовують в кіноіндустрії та добре працює, але є дорогим і не завжди придатний для використання в інших галузях[1-3].

Другий в основному використовує комп'ютерне бачення, яке передбачає обробку зображень за допомогою камер або(й) інших датчиків.

Протягом останніх років з'являється все більше статей присвячених обробці зображень з камери за допомогою спеціальних алгоритмів або за допомогою нейронних мереж. Якість роботи нейронних мереж значно підвищилась через використання нових архітектур нейронних мереж та великої кількості нових досліджень щодо підвищення якості та швидкості їх навчання. А також, через розвиток технологій та збільшення кількості даних[2-4].

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена вивченню та аналізу запропонованого підходу для розпізнавання моделей рук. Дослідження сфокусовано на вивченні застосування згорткових нейронних мереж для детектування ключових точок руки й побудови моделі руки й розпізнавання цієї моделі.

Постановка завдання. Завдання є створити систему, яка дасть можливість, використовуючи згорткову нейронну мережу, детектувати ключові точки руки з одного зображення й після цього будувати модель руки, щоб порівняти її з описаними моделями для детектування жестів.

Викладення основного матеріалу. Проведено аналіз використання згорткових мереж для детектування ключових точок, створення моделі руки з них й подальшого розпізнавання цих моделей.

Для детектування ключових точок використовуємо архітектуру Convolutional Pose Machines (CPM). CPM повертає heatmaps кожної ключової точки.

CPM має неявне навчання залежностей між зображенням і виведенням зразу декількох ключових точок, тісна інтеграція між навчання й виведенням та модульний послідовний дизайн. На кожному етапі CPM, отримується зображення та heatmaps, виготовлені на попередньому етапі, як вхідне й намагається виправити та уточнити положення ключових точок на руці.

Так як, створення згорткової мережі з багатьма модулями призводить до загальної мережі з багатьма шарами, що призводить до проблеми затухання градієнтів під час навчання, то створюються додаткові виходи нейронної мережі після кожного модуля, що використовуються лише під час навчання, які намагаються так само визначити положення ключових точок за допомогою heatmaps. Створення додаткових виходів під час навчання дає можливість поповнювати градієнти та спрямовувати мережу для створення все більш точних heatmaps.

Функція витрат, яку ми прагнемо мінімізувати на виході кожного етапу на кожному рівні, виглядатиме:

$$f_t = \sum_{p=1}^{P+1} \sum_{z \in Z} ||b_t^p(z) - b_{t`}^p(z)||_2^2 \quad (1)$$

Ідеальна heatmap для частини p, як $b_t^p(Y_p = z)$, що створює гаусові піки на місцях істинного розташування кожної ключової точки p, де z – це положення цієї ключової точки, а t – номер класифікатора(модуль нейронної мережі).

Загальна ціль для усієї архітектури виглядатиме, як додавання втрат на кожному етапі, тобто:

$$F = \sum_{t=1}^{T} f_t \quad (2)$$

Було зроблено знімок лівої руки для проведення обробки по ідентифікації жесту. Після подання знімку з рукою на нейронну мережу ми отримуємо heatmaps, з яких знаходимо ключові точки руки для моделювання жесту (рис. 1).

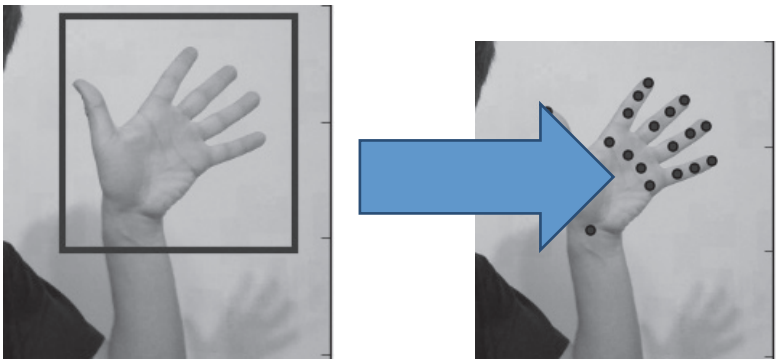


Рис. 1. Детектування ключових точок руки

Після цього нормалізуємо модель до кожної описаної моделі за допомогою rotate, shift, scale і flip (рис. 2, рис. 3, рис 4).



Рис. 2. Rotate моделі руки

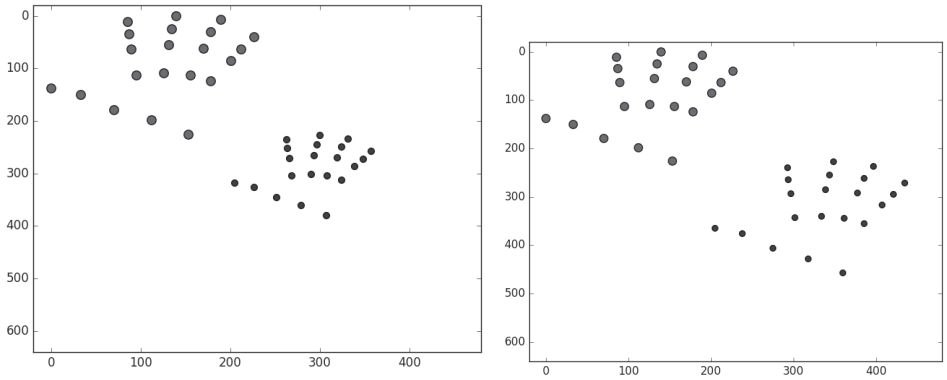


Рис. 3. Scale моделі руки

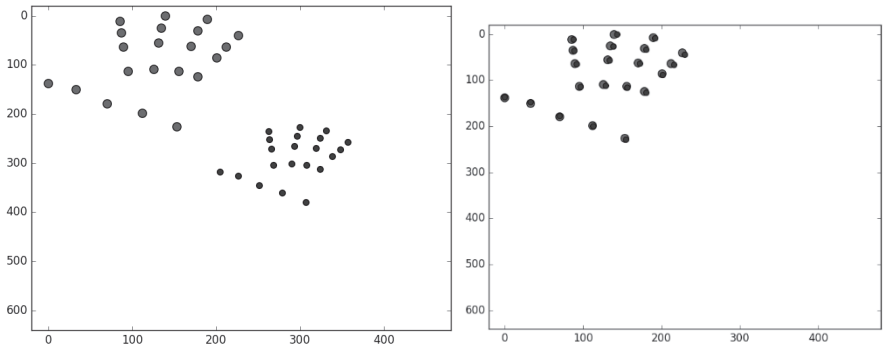


Рис. 4. Shift моделі руки

Після нормалізації побудованої моделі до даного жесту, вона порівнюється з ним та визначається чи задовольняє дана модель описаний жест. Як видно з рис. 4 наша модель руки задовольняє в межах допустимого відхилення заданий жест.

Загальний алгоритм. На рис. 5 показаний загальний алгоритм побудови та розпізнавання моделіруки.

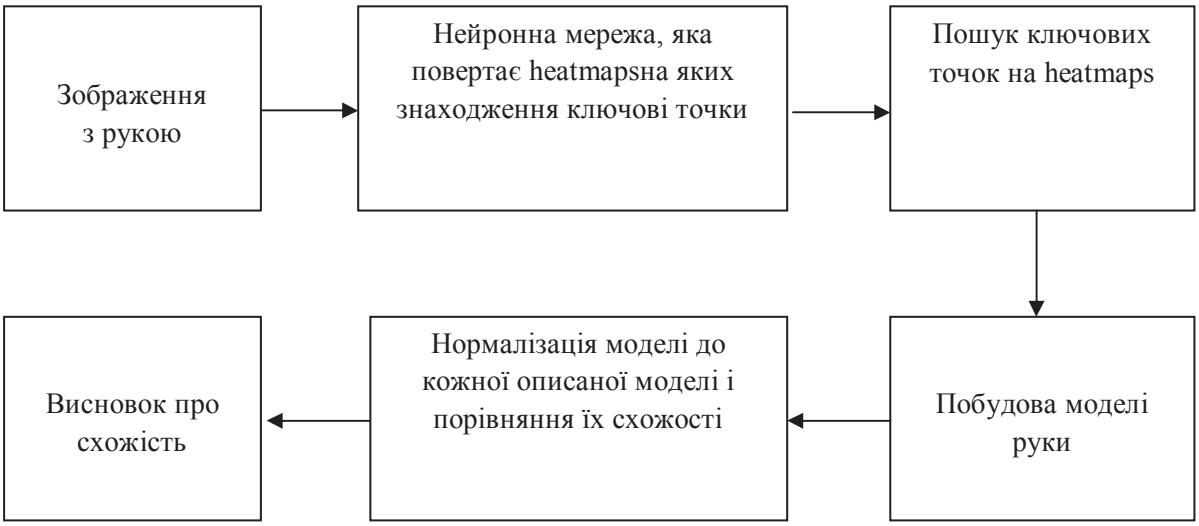


Рис. 5. Алгоритм побудови та розпізнавання моделі руки

Висновки. В даній роботі було проаналізовано структуру та параметри моделі Convolutional Pose Machine(CPM), яка використовується для визначення heatmaps знаходження точок руки з однієї камери. Також було вирішено проблеми аналізу та знаходження ключових точок руки, й подальшого використання для моделювання й розпізнавання жестів скелетону руки. Було порівняно отриманий жест з еталонним та визначено, наша модель руки задовольняє в межах допустимого відхилення заданий жест.

References

1. Shih-En Wei, Varun Ramakrishna, Takeo Kanade, Yaser Sheikhn(2016). Convolutional Pose Machines. arXiv preprint arXiv: 1602.00134,

2. J. Tompson, R. Goroshin, A. Jain, Y. LeCun, and C. Bregler. Efficient object localization using convolutional networks. In CVPR, 2015.

3. Alfredo Canziani, Adam Paszke, Eugenio Culurciello(2017). An Analysis of Deep Neural Network Models for Practical Applications. arXiv preprint arXiv:1605.07678

4. Xingyi Zhou, Qingfu Wan, Wei Zhang, Xiangyang Xue, Yichen Wei(2016). Model-based Deep Hand Pose Estimation. arXiv preprint arXiv:1606.06854

ДОВІДКА ПРО АВТОРІВ

Мухін Вадим Євгенійович - доктор технічних наук, професор кафедри математичних методів системного аналізу Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Mukhin Vadim Yevgenievich - Doctor ofTechnical Sciences, Professor of the Department of Mathematical Methods of System Analysis, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

E-mail: v_mukhin@i.ua

Корнага Ярослав Ігорович - кандидат технічних наук, доцент кафедри технічної кібернетики Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського».

Kornaga Yaroslav Igorovich - Candidate of Technical Sciences, Associate Professor of the Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

E-mail: slovyan_k@ukr.net

Любівий Андрій Ігорович – студент, кафедри Технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Liubivyi Andrii Igorovich – student, Department of Technical Cybernetics, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: 120494al@gmail.com

Vadim Mukhin,
Yaroslav Kornaga,
Andrii Liubivyi

MODELING AND RECOGNITION OF GESTURE OF HAND SKELETON

Target setting. The problem of modeling and recognizing gestures becomes more relevant in recent days due to the growing number of systems that a person can interact with using gestures.

Actual scientific researches and issues analysis. Due to the emergence of new methods and approaches in the field of artificial intelligence, the topic of image recognition has become more relevant in recent years.

Uninvestigated parts of general matters defining. This article is devoted to the study and analysis of the proposed approach for the recognition of models of hands. The research focuses on the study of the use of convolutional neural networks for detecting key points of the hand and constructing a model of the hand and recognizing this model.

The research objective. The purpose of this paper is to investigate a system that will enable, using a convolutional neural network, to detect key points of a hand from one image, and then build a hand model to compare it with the described model gestures.

The statement of basic materials. To detect key points, we use the Convolutional Pose Machines (CPM) architecture. CPM returns the heatmaps of each key point.

CPM has implicit learning of the relationship between the image and the output of several key points at once, the close integration between learning and output and modular sequential design. At each stage of the CPM, images and heatmaps made in the previous step are received as input and are trying to correct and clarify the position of the key points on the hand. After detection of keypoints we construct and compare hand build with described models.

Conclusions. The content, structure and parameters of the Convolutional Pose Machine (CPM) model used to determine heatmaps for finding points of a hand from one camera to solve the problem of finding the key points of the hand, and their further using for modeling and recognition of hand skeleton gestures are analyzed. The approach proved to be good for creating and recognizing hand models.

Keywords: modeling of objects, hand, convolutional neural networks, classification.

UDC 004.8

Ярослав Подзірей, Павло Регіда

КЛАСТЕРИЗАЦІЯ. ЗМЕНШЕННЯ РОЗМІРНОСТІ ОЗНАК ОБ'ЄКТІВ В МАШИННОМУ НАВЧАННІ ТА ВІЗУАЛІЗАЦІЯ ДАНИХ. МЕТОД t-SNE.

CLUSTERING. DECREASE OF DIMENSION OBJECT DESCRIPTION IN MASTER EDUCATION AND VISUALIZATION OF DATA. T-SNE METHOD

У даній статті розглядається питання кластеризації, зменшення розмірності ознак вибірки вхідних об'єктів аналізу і візуалізації даних за допомогою алгоритму t-SNE. Пропонується варіант можливого покращення даного алгоритму. Для тестування використовується мова Python та популярна бібліотека scikit-learn.

Ключові слова: машинне навчання, кластеризація, зменшення розмірності простору вибірки ознак, візуалізація даних, алгоритм t-SNE.

Рис.: 1.. Бібл.: 5.

In this article, the question of clustering, reducing the size of the characteristics of the sample of input objects of analysis and data visualization using the t-SNE algorithm. A variant of possible improvement of this algorithm is proposed. For testing, Python language and the popular scikit-learn library are used.

Key words: machine learning, clustering, diminishing of the dimension of the sample of traits, data visualization, t-SNE algorithm.

Fig .: 1 .. Bible: 5.

Актуальність теми дослідження. Машинне навчання розвивається дуже швидкими темпами і кількість даних для аналізу росте з кожним днем. Така тенденція, в свою чергу, призводить до поступового ускладнення аналізу даних і розуміння взаємозв'язків між ними, що в машинному навчанні є ключовою особливістю. Саме тому люди в першу чергу шукають можливості для спрощення представлення даних без втрати деякої важливої інформації. Дана стаття присвячена одному з відносно молодих алгоритмів, який досить добре справляється з задачами подібного типу.

Постановка проблеми. Не дивлячись на те, що алгоритм t-SNEдосить добре себе зарекомендував в сфері машинного навчання та аналізу даних, він все ж має певні недоліки, які намагаються бути вирішеними в даній статті.

Аналіз останніх досліджень і публікацій. Протягом останніх років з'являється все більше статей присвячених різним модифікаціям алгоритму t-SNE, проте більшість цих досліджень направлені на підвищення швидкості даного алгоритму, оскільки він має одну з найбільших складностей серед конкурентів. Більшість інших недоліків даного алгоритму не є досить детально вивченими на даний час.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена дослідженню потенційності покращення алгоритму t-SNE за рахунок можливості використання інформації про розмітку та склад вибірки, а також можливості додавати нові точки в вибірку для покращення візуального результату без перерахунку всіх координат.

Постановка завдання. Провести модифікацію алгоритму t-SNEта аналіз його тестування.

Викладення основного матеріалу. Зазвичай в машинному навчанні ми маємо справу з досить багатовимірними вибірками даних, оскільки дуже часто аналіз даних проводиться одночасно відносно багатьох параметрів і уникнути цього в еру технічного розвитку неможливо. Але при цьому ми хочемо якось дивитися на ці дані, розуміти, як вони влаштовані, які там взаємозв'язки, які ознаки важливі, а які - ні, як співвідносяться класи між собою. Існує досить велика кількість алгоритмів які допомагають провести певну характеристику даних без пониження розмірності, але такі данні досить важко сприймаються людиною, оскільки ми не можемо мислити більш ніж в 3 вимірах.

Хотілося б відобразити всю вибірку в двовимірний або тривимірний простір так, щоб відразу було видно всі закономірності в даних, вся їх структура. Наприклад, якщо класи сильно перемішані між собою, якісь класи виділяються і їх можна добре відокремити від решти, тобто провести процес кластеризації. Власне, саме так ми приходимо до задачі візуалізації даних: це окремий випадок нелінійного зниження розмірності, коли розмірність простору, в яке ми намагаємося спроектувати нашу вибірку, це 2 або 3.

На сьогоднішній день, одним з найбільш популярних підходів до вирішення такого класу задач є використання методу t-SNE(t-distributed stochastic neighbor embedding).

Суть алгоритму t-SNE полягає в розрахунку двох матриць схожості: перша - для вихідних точок в багатовимірному просторі, друга - для кінцевих точок в двовимірному. Друга матриця підбирається таким чином, щоб якомога краще відображати властивості першої.

На відміну від алгоритмів конкурентів, метод t-SNEне потребує максимальної близькості розміщення об’єктів у результуючому просторі відносно заданого. Даний алгоритм потребує лише збереженню пропорцій об’єктів, що дає досить великі переваги, оскільки при досить великому скупченні об’єктів в багатовимірному просторі, досить важно відобразити дані закономірності в маломірному, тобто данні виходять більше чіткими і зрозумілими для аналізу.

Алгоритм створює статистичну модель різних особливостей даних, які формують вірогідну картину будь-якого заданого значення. Він випадковим чином розсіює точки даних і починає переміщати ці точки крок за кроком в процесі, який намагається знайти золоту середину, де точки збалансовані статистично. Тобто, точки тягнуть і натискають одна на одну, ґрунтуючись на тому, як незбалансовані вони з усіма іншими точками, вони самоорганізуються за алгоритмом. Рух кожної характеристики точок в кожному напрямку визначається впливом інших точок, які повинні бути розміщені на відстані за їх особливою ймовірністю.

Загальні формули для обчислення ймовірностей в заданому та результуючому просторах:

$$p(x_j|x_i) = \frac{\exp\left(\| x_i - x_j\|^2/2\sigma^2\right)}{\sum_{k \neq i} \exp\left(\|x_i - x_k\|^2/2\sigma^2\right)} \quad (1)$$

$$q(\tilde{x}_j|\tilde{x}_i) = \frac{\left(1 - \|\tilde{x}_i - \tilde{x}_j\|^2\right)^{-1}}{\sum_{k \neq i} \left(1 + \|\tilde{x}_i - \tilde{x}_k\|^2\right)^{-1}} \quad (2)$$

Ці формули показують на скільки j-та точка близька до і-тої з заданим відхиленням σ. Сигма відрізняється для кожної точки. Воно підбирається так, щоб виконувалась оцінка перплексії:



(3)

де H(Pi) – ентропія Шеннона в бітах



(4)

Для розуміння фізичних властивостей алоритму, перплексію можна вважати пом'якшеною оцінкою ефективної кількості сусідніх точок для заданої. Зазвичаний беруться значення від 5 до 50.

Для порівняння розподілень ймовірностей заданого і результуючого простору з точки зору математики дуже добре підходить дивергенція Кульбака-Лейблера:

$$p(x_j|x_i)\log\frac{p(x_j|x_i)}{q(\tilde{x}_j|\tilde{x}_i)} \rightarrow \min_{\tilde{x}_1,\ldots,\tilde{x}_z} \tag{5}$$

Загальне розподілення ймовірностей:

$$p_{ij} = \frac{p_{i|j} + p_{j|i}}{m}, \tag{6}$$

де m – кількість точок в наборі даних.

Заданий алгоритм мінімізує суму всіх ймовірних відстаней за допомогою алгоритму градієнтного спуску з наступним градієнтом:

$$\frac{\partial C\,cost}{\partial y_i} = 4\sum_j (p_{ij} - y_{ij})(y_i - y_j)(1 - \| y_i - y_j \|^2)^{-1} \tag{7}$$

Для розуміння фізичних властивостей даного процесу варто зазначити, що алгоритм буде певним чином притягувати точки простору відображення для близьких між собою точок заданого простору, та відштовхувати для точок, які розміщені далеко в заданому багатовимірному просторі.

Спроба модифікації заданого алгоритму проводилась з урахуванням вже існуючих рішень, які зазвичай направлені на зменшення швидкості виконання або зменшення складності алгоритму. Ціллю модифікації алгоритму є можливість використовувати інформацію про розмітку вибірки та додавати нові точки без перерахунку всіх координат.

Для початку пропонується взяти до уваги вибірку $X^{m0} = \{\mathbf{x}_1,\ldots,\mathbf{x}_m 0\}$. Для неї оцінимо $P^{m0\times m0}$ і знайдемо його відображення Z^{m0} и $Q^{m0\times m0}$. Додамо нову множину об'єктів $X^{m-m0} = \{\mathbf{x}_m 0_{+1},\ldots,\mathbf{x}_m\}$. Не міняючи положень $\mathbf{z}_i$, $i = m^0+1,\ldots,m$ початкових точок знайдемо відображення Z^{m-m0} и $Q^{(m-m0)\times m}$, зміщуючи точки в напрямку градієнта.(7)

Для підвищення надійності Z^{m0} використаємо інформацію про розмітку y_i об'єктів $\mathbf{x}_i$, $i = 1,\ldots,m^0$. Пропонується використовувати інформацію про розмітку при обчисленні $P^{m0\times m0}$, наприклад:

$$p_{j|i} = \frac{\exp(-||\mathbf{x}_i - \mathbf{x}_j||^2/\sigma_{ij}^2)}{\sum_{k\neq i}\exp(-||\mathbf{x}_i - \mathbf{x}_k||^2/\sigma_{ij}^2)}, \quad \sigma_{ij} = \begin{cases}\sigma^2, y_i = y_j,\\ \varepsilon \rightarrow 0, y_i \neq y_j\end{cases} \tag{8}$$

Для тестування була використана задача:

Для існуючого документу C —послідовність символів $c_1,\ldots,c_{Ld}$. В даному документі присутні зашифровані блоки інформації. Необхідно для кожного символу c_l документу провести відповідну кластеризацію за характеристикою приналежності заданого символу до зашифрованого блоку інформації.

Результати виконання модифікації:

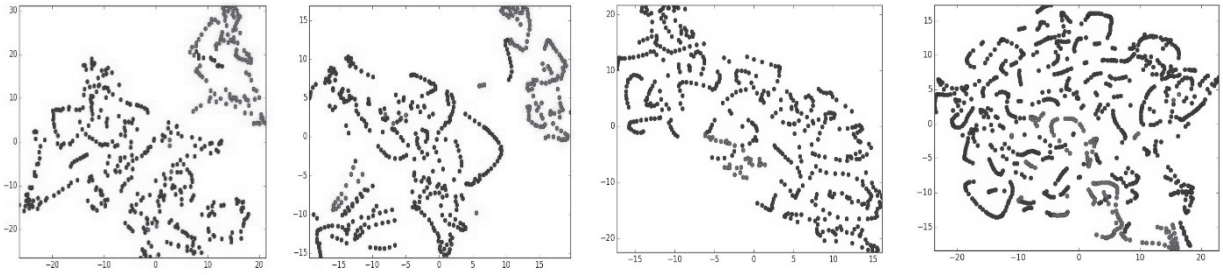


Рис. 1. Результати виконання для,
відповідно, 500, 500, 1000, 1000 речень в файлі

Висновки. Проаналізувавши результати можна зробити висновок, що задана модифікація алгоритму, на даний момент, має досить непоганий результат лише для малого об'єму даних. При вхідних параметрах великого розміру, результат виходить не дуже чітким, не можна відразу побачити чітку кластеризацію об'єктів, а відповідно, не можна провести якісний аналіз даних. Даний метод модифікації потребує більш детального вивчення.

Список використаних джерел

1. L.J.P. van der Maaten and G.E. Hinton. Visualizing High-Dimensional Data Using t-SNE. Journal of Machine Learning Research 9(Nov):2579-2605, 2008. PDF

2. G.E. Hinton and S.T. Roweis. Stochastic Neighbor Embedding. In Advances in Neural Information. Processing Systems, volume 15, pages 833–840, Cambridge, MA, USA, 2002. The MIT Press. PDF

3. L.J.P. van der Maaten. Accelerating t-SNE using Tree-Based Algorithms. Journal of Machine Learning Research 15(Oct):3221-3245, 2014. PDF

4. Hyunsoo Kim, Haesun Park, and Hongyuan Zha. Distance Preserving Dimension Reduction for Manifold Learning Proceedings of the 2007 SIAM International Conference on Data Mining

5. Wes McKinney. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. O'Reilly, 2012.

ДОВІДКА ПРО АВТОРІВ

Подзірей Ярослав Іванович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Yaroslav Podzirei – student, Department of Computer Engineering,
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: ginofa00@gmail.com

Регіда Павло Геннадійович – аспірант, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Rehida Pavlo – PhD student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: pavel.regida@gmail.com

Yaroslav Podzirei, Pavlo Rehida

CLUSTERING. DECREASE OF DIMENSION OBJECT DESCRIPTION IN MASTER EDUCATION AND VISUALIZATION OF DATA. T-SNE METHOD

Target setting. People are primarily looking for opportunities to simplify the presentation of data without losing some important information. This article is devoted to one of the relatively young algorithms, which quite well copes with problems of this type.

Formulation of the problem. Despite the fact that the t-SNE algorithm has proven itself well in the field of machine learning and data analysis, it nevertheless has certain disadvantages that are trying to be resolved in this article.

Actual scientific researches and issues analysis. In recent years, there are more articles devoted to various modifications of the t-SNE algorithm, but most of these studies are aimed at increasing the speed of this algorithm, since it has one of the greatest difficulties among competitors. Most of the other disadvantages of this algorithm are not sufficiently studied at present.

Uninvestigated parts of general matters defining. This article is devoted to the study of the potential of improving the t-SNE algorithm due to the possibility of using information about the markup and composition of the sample, as well as the ability to add new points in the sample to improve the visual result without recalculating all coordinates.

The research objective. Modify the t-SNE algorithm and analyze its testing.

General model structure. Usually in machine learning, we deal with rather high-dimensional data samples, because very often the analysis of data is carried out simultaneously with respect to many parameters and it is impossible to avoid this in the era of technical development. But at the same time, we want to somehow look at these data, understand how they are arranged, what are the interrelationships there, which features are important, and which - no, how the classes relate to each other. There is a fairly large number of algorithms that help to carry out a certain characteristic of the data without diminishing the dimension, but such data are hardly perceived by a person, since we can not think in more than 3 dimensions. I would like to display the entire sample in two-dimensional or three-dimensional space so that all the regularities in the data, all their structure, were immediately visible. For example, if the classes are strongly mixed with each other, some classes are allocated and they can be well separated from the rest, that is to hold the process of clusterization. Actually, this is the way we come to the task of data visualization: this is a special case of a nonlinear decrease in dimensionality, when the dimension of the space in which we try to design our sample is 2 or 3.To date, one of the most popular approaches to solving such a task class is the use of the t-SNE method (t-distributed stochastic neighbor embedding).

Conclusions. After analyzing the results it can be concluded that the given modification of the algorithm, at the moment, has a fairly good result for only a small amount of data. When input parameters are large, the result is not very clear, it is impossible to immediately see a clear clustering of objects, and accordingly, it is impossible to conduct a qualitative analysis of data.

Key words: machine learning, clustering, diminishing of the dimension of the sample of traits, data visualization, t-SNE algorithm.

UDC 004.8

Вадим Щур, Віктор Стешин

ВИКОРИСТАННЯ МЕТОДІВ РОЗПІЗНАВАННЯ СИМВОЛІВ ДЛЯ ВИЗНАЧЕННЯ ДОСТОВІРНОСТІ ОСОБИСТОГО ПІДПИСУ

USING CHARACTER RECOGNITION METHODS TO DETERMINE THE AUTHENTICITY OF A PERSONAL SIGNATURE

У статті розглядається питання аналізу і вдосконалення існуючих методів розпізнавання символів для визначення достовірності особистого підпису. Для цього використовуються нейронні мережі, можливості навчання та збільшення точності яких, дозволяють ефективно вирішувати проблеми визначення за умов різноманітності і неточності рукописних підписів. Система розпізнавання тренуються на базі неодноразового введення особистих підписів та встановлюють відповідні ваги нейронної мережі.

Ключові слова: розпізнавання символів,згорткові нейронні мережі, багатошаровий перцептрон, навчання мережі.

Рис.: 4. Бібл.: 5.

The paper deals with the analysis and improvement of existing methods of character recognition to determine the authenticity of a personal signature. To do this, using neural networks, learning opportunities and increasing the accuracy of which, can effectively solve the problems of definition in terms of diversity and inaccuracy of handwritten signatures. The recognition system trains on the basis of repeated input of personal signatures and establishes the corresponding weights of the neural network.

Key words: character recognition, convolutional neural networks, multilayer perceptron, network learning.

Fig.: 4. Bibl.: 5.

Актуальність теми дослідження. Протягом багатьох років підпис є необхідною умовою визнання істинності особи людини в біометричних системах контролю і управління доступом. Основною перевагою використання розпізнавання підпису як методу верифікації є той факт, що більшість сучасних портативних комп'ютерів і електронних пристроїв вже дозволяють використовувати рукописний ввід даних, тому немає ніякої необхідності створення принципово нових пристроїв біометричного збору інформації. У той же час, існує дуже мало систем для розпізнавання, які можуть забезпечити достатньо високу точність розпізнавання, зберігаючи прийнятний рівень ефективності.

Постановка проблеми. Відсутність методів розпізнавання символів, які б могли забезпечити достатньо високу точність розпізнавання для визначення достовірності особистого підпису.

Аналіз останніх досліджень і публікацій. Протягомостанніх роківз’являється все більше статей присвячених темі верифікації особистого підпису, зокрема, завдяки появі нових методів розпізнавання символів з використанням нейронних мереж. Проте запропоновані методи, не показують достатньої ефективності, щоб задовільнити поставлені задачі.

Виділення недосліджених частин загальної проблеми. Дана статтяприсвячена вивченню, аналізу та вдосконаленню існуючих методів розпізнавання символів. Дослідження сфокусовано на застосуванні результуючих методів використовуючи нейронні мережі.

Постановка завдання. Завданнямє покращення існуючих методів розпізнавання символів, створити модель,неоднократне введення особистих підписів дозволить ідентифікувати підпис, який належить власнику в процесі навчання нейронної мережі.

Аналіз засобів реалізації. Базовим завданням особистого підпису в біометричній системі являється визначення правдивості та завіреності документів людиною. Сучасні інформаційні системи уже давно використовують письмове джерело введення даних. Використання розпізнавання особистого підпису в системі контролю та доступу являється перспективним напрямком та логічним для отримання будь-якого доступу.

Основний алгоритм цього процесу складається із навчання системи тобто введення початкових даних та визначення відповідності еталонному значенню. Після виконання всіх етапів система надає доступ, або особливі можливості (Рис.1, 2).

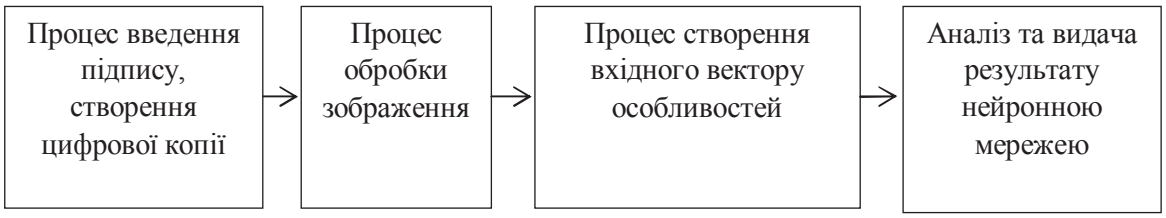


Рис. 1. Процес отримання доступу по особистому підпису.

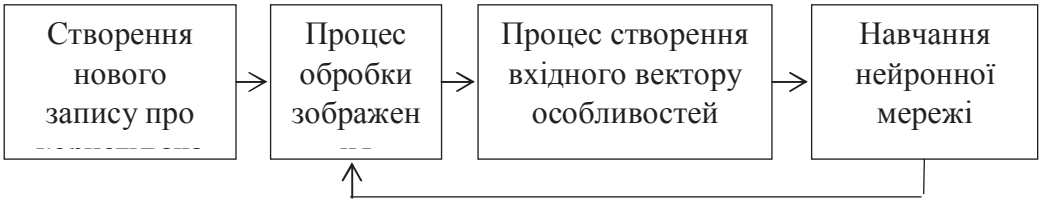


Рис. 2. Процес навчання мережі.

Нейронні мережі, завдяки можливості постійного навчання та узагальнення, дозволяють ефективно виконувати аналіз введеної інформації в умовах мінливості та різноманітності рукописних підписів. Розробка найпростішої нейронної мережі на базі багатошарового перцептрону дозволяє визначити відповідність вхідного малюнку відображенню символу. Такий спосіб використовується для отримання символьного співпадіння у розпізнаванні текстів. В цьому випадку кожний символ виділяється у окрему частину і виконується його аналіз. Розпізнавання особистого підпису має схожий алгоритм, але має ряд особливих характеристик.

Для роботи з визначенням достовірності підписів вибрана Згорткова нейронна мережа. Дана мережа характеризується високою продуктивністю та точністю у роботі із зображеннями і порівнюється із можливостями людського

зору, оскільки дозволяє визначити приналежність, або іншими словами точність відповідності зображення, еталону не лише у визначеній позиції, але і по всій області спостереження, незважаючи на кут повороту, масштабу [1].

Структура Згорткової мережі наведена нижче:

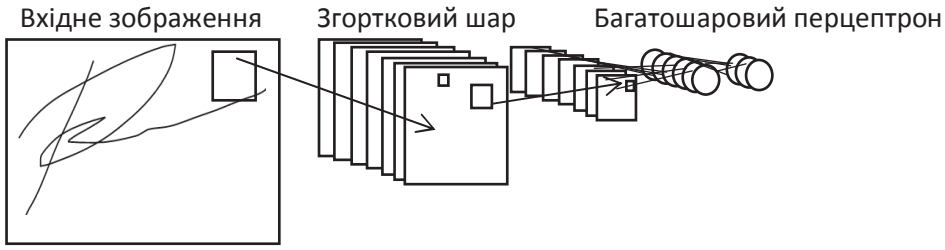


Рис. 3. Структура Згорткової нейронної мережі.

Як показано на рис. 4. Структура даної мережі складається із вхідного слою, що відповідає областям введеного особистого підпису, згортковим шарам, що відповідають багаторазовому згортанню зображення для досягнення можливостей аналізу областей з нечітким розміщенням та шарам звичайної нейронної мережі, що дає відповідь належності особистого підпису до еталонних [2].

Завдяки згортковому шару нейронної мережі досягається мета, отримуються вхідні параметри для багатошарового парсетрона по всій області зображення (1):

$$h_{ij}^{k} = \tanh\left((w^{k} * x)_{ij} + b_{k}\right) \quad (1)$$

Результат кожного згортання h^k потрапляє у функцію активації, яку зазвичай вважають вбудованою в шар згортання, значення якого залежить від зміщення ваг $(w^k * x)_{ij}$ та похибки b_k, далі отримується результуюче значення багатошарового перцептрона з визначення достовірності підпису [3].

Навчання даної мережі відбувається методом зворотного розповсюдження помилки, що дає можливість налаштувати мережу задаючи кожного разу новий приклад навчання. Це дозволить використовувати дану мережу для класифікації підпису та віднесенню її до свого класу. Недоліком вибраної мережі для подальших досліджень є великий час навчання та можливість отримання збою, а також велика кількість варіації даних з виділенням необхідних, можливо надлишкових, обчислювальних потужностей.

Виконавши навчання та введення інформації на основі досліджень сформовані наступні результати (Рис. 4).

По результатах роботи системи по розпізнаванню підпису, доступ до даних в рідкісних випадках отримує також і зловмисник. Із збільшенням кількості зразків навчання, збільшується і ймовірність входу користувачем, зловмисник в даному випадку збільшує даний показник із меншою швидкістю.

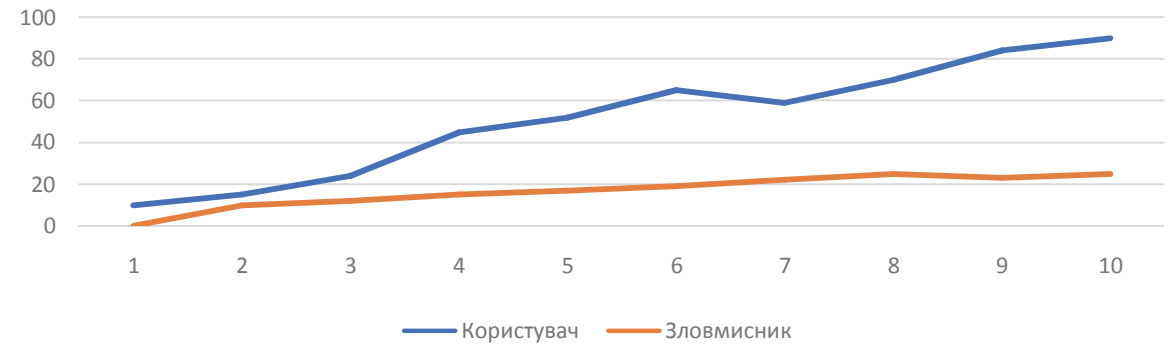


Рис. 4. Результат аналізу підпису та ймовірність отримання доступу.

Висновки. Замість звичайного розпізнавання багатошаровим перцептроном, використано згорткову нейронну мережу, що дозволило підвищити точність аналізу підпису. При збільшені кількості зразків навчання, відповідно збільшується і ймовірність розпізнавання. Тому, при великій кількості образів можна досягти непоганої точності розпізнавання. Але, з іншого боку, це призводить до великих затрат часу на навчання, та можливості отримання збою. Питання збільшення ефективності даного методу потребує більш детального вивчення.

Список використаних джерел

1. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel: Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation, 1(4):541-551, Winter 1989.

2. Matt Mazur. A Step by Step Backpropagation Example [Електронний ресурс] - Режим доступу: https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/

3. DeepLearning 0.1 documentation [Електронний ресурс] – Режим доступу: http://deeplearning.net/tutorial/lenet.html

4. Faundez-Zanuy, Marcos (2007). «On-line signature recognition based on VQ-DTW». Pattern recognition 40 (3): 981-992.

5. M. M. Lange, S.N. Ganebnykh (2005). «Classification of 2D Grayscale Objects in a Space of the Multiresolution Representation».

ДОВІДКА ПРО АВТОРІВ

Щур Вадим Юрійович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Shchur Vadym –student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: robinzon004@gmail.com

Стешин Віктор Васильович – асистент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Steshyn Viktor – assistant, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: v.steshyn@kpi.ua

Vadym Shchur, Viktor Steshyn

USING CHARACTER RECOGNITION METHODS TO DETERMINE THE AUTHENTICITY OF A PERSONAL SIGNATURE

Relevance of the research. For many years, signatures are a prerequisite for recognizing the truth of a person's personality in biometric systems of control and access control. The main advantage of using signature recognition as a verification method is the fact that most modern portable computers and electronic devices already allow the use of handwriting data, so there is no need to create fundamentally new devices for biometric information collection. At the same time, there are very few systems for recognition that can provide a high accuracy of recognition, while maintaining an acceptable level of efficiency.

Target setting. The lack of character recognition methods that could provide sufficiently high recognition accuracy for determining the authenticity of a signature.

Actual scientific researches and issues analysis. In recent years, there are more and more articles devoted to the verification of personal signature, in particular, due to the emergence of new methods for character recognition using neural networks. However, the proposed methods do not show sufficient effectiveness to meet the objectives.

Uninvestigated parts of general matters defining. This article is devoted to the study, analysis and improvement of existing methods of character recognition. The research is focused on applying the resulting methods using neural networks.

The research objective. The task is to improve the existing methods of character recognition, create a model, repeatedly entering personal signatures will identify the signature that belongs to the owner in the process of training the neural network.

The statement of basic materials. To work with the definition of the authenticity of signatures, a convolutional neural network was chosen. This network is characterized by high performance and accuracy in working with images. Due to the convolutional layer of the neural network, the goal is achieved, the input parameters for the multilayer perceptron obtained across the entire image area. The training of this network occurs by the method of back propagation of the error, it makes it possible to configure the network by setting a new learning example each time. This will allow using this network to classify the signature and assign it to its class.

Conclusions. Instead of the usual recognition by a multilayer perceptron, a convolutional neural network was used, which made it possible to improve the accuracy of the signature analysis. As the number of training samples increases, the probability of recognition increases accordingly. Therefore, with a large number of images, you can achieve a good recognition accuracy. But, on the other hand, this leads to a large expenditure of time for training, and the possibility of getting a malfunction. The question of increasing the effectiveness of this method requires more detailed study.

Key words: character recognition, convolutional neural networks, multilayer perceptron, network learning.

УДК 004.93

Коломієць Олег Олексійович

МЕТОДИ ІНТЕГРАЦІЇ ЕЛЕМЕНТІВ ДОПОВНЕНОЇ РЕАЛЬНОСТІ З СИСТЕМОЮ РОЗПІЗНАВАННЯ ОБРАЗІВ

METHODS FOR INTEGRATION OF ELAMENTS OF AGUMENTED REALITY WITH IMAGE RECOGNITIONSYSTEMS

Дана стаття присвячена методам інтеграції елементів доповненої реальності з системою розпізнавання образів. Мета дослідження – опис алгоритмів позиціонування елементів доповненої реальності на зображенні з використанням даних отриманих від системи розпізнавання образів та обробки відео-потоку для створення доповненої реальності з використанням лінійної регресії. Для розпізнавання обличчя використано сторонню бібліотеку Vision.

Ключові слова: доповнена реальність, опорні точки, проекція, лінійна регресія, апроксимація

Бібл.: 6.

This article is focused on methods for integration elements of augmented reality with pattern recognition system. The research aims to describe the algorithms that adds elements of augmented reality to the image based on results from pattern recognition system and process video for creation of augmented reality using linear regression. Third party library Vision was used for face detection.

Keywords: augmented reality, landmarks, projection, linear regression, approximation

Bibl.: 6.

Актуальність теми дослідження. У зв’язку із значним поширенням та зростом обчислювальних потужності мобільних пристроїв для них відкриваються нові області застосування. Однією з них є створення доповненої реальності на основі зображень та відео потоку, отриманих з камери мобільного пристрою. Тому протягом останніх декількох років сфера створення елементів доповненої реальності набула значної актуальності.

Постановка проблеми. Створення елементів доповненої реальності складається із наступних основних етапів: аналіз оригінального кадру, позиціонування об’єкту, який потрібно додати та генерація вихідного кадру, шляхом накладання необхідного об’єкту на оригінальний кадр. Вважатимемо, що процес розпізнавання об’єктів відбувається з використанням сторонньої бібліотеки. Розглянемо проблему позиціонування та генерації вихідного кадру.

Постановка задачі. Мета даного дослідження – розробка алгоритмів додавання елементів доповненої реальності на зображення (на основі даних, отриманих від системи розпізнавання образів) та відео.

Поняття доповненої реальності.Доповнена реальність (розширена реальність) - результат введення у поле сприйняття будь-яких сенсорних даних (зображення, звук тощо) з метою доповнення інформації про навколишнє середовище та покращення сприйняття інформації. Доповнену реальність, іноді, називають змішаною, що створюється поєднанням реальних об’єктів сприйняття із накладеними за допомогою комп’ютера елементами.

Доповнена реальність пов’язана із розпізнавання образів, адже вона, переважно, базується на накладанні ефектів на існуюче зображення. Процес накладання потребує визначення позиції об’єкта на зображенні, для цього потрібне розпізнавання та визначення координат певних елементів на ньому.

Розпізнавання образів - це підвид машинного навчання що відображає завдання ідентифікації та розпізнавання об’єктів на зображенні або потоці відео.

В межах даного дослідження розглядатиметься накладання ефектів на обличчя людини.

Аналіз вхідних даних та розробка алгоритмів. Для розпізнавання обличчя доцільно використати одне з готових рішень OpenCV, Google Play Services Vision тощо. В контексті застосування на мобільних пристроях під управлінням ОС Android кращим варіантом є Google Play Services Vision, адже ця бібліотека є частиною пакету Google Play Services, який розроблявся спеціально для використанні у Android, отже враховує особливості даної платформи. Бібліотека дозволяє розпізнати обличчя, визначити його ключові точки (англ. Landmarks): очі, основа носа, губи тощо, нахил та поворот голови.[1]

Ключовим моментом у розробці системи для додавання елементів доповненої реальності є розміщення ефекту на зображенні. Для двовимірних ефектів можна виділити такі основні параметри для позиціонування:

- Координати опорної точки ефекту на зображенні;
- Масштаб ефекту відносно зображення, що редагується;
- Поворот ефекту відносно вертикалі.

Ефекти потрібно поділити на типи, які мають однакову логіку визначення позиції, на приклад, “окуляри”, “вуса” тощо.

Для реалізації розширюваної системи у моделі збереження даних доцільно використати шаблон проектування композит (англ. Composite), який відображатиме структуру типів ефектів [2]

Позиціонування ефектів на зображенні. Розглянемо приклад визначення позиції ефекту “окуляри”. За опорну точку для даного ефекту доцільно обрати центр зображення(для фронтального вигляду окулярів). Для позиціонування використаємо такі опорні точки обличчя: ліве і праве око та основу носа.

1. Визначимо точку проекції основи носа на пряму, що проходить через точки очей, позначимо її як Projection (таким чином враховуватиметься поправка на нахил обличчя);

2. Абсцису точки розташування ефекту знайдемо за формулою

$$x = \frac{\left(\frac{leftEye.x + rigthEye.x}{2} + Projection.x\right)}{2}, \quad (1)$$

де leftEye.x - абсциса точки лівого ока, rigthEye.x - правого, projection.x - точки проекції;

3. Ординату обчислимо за формулою :

$$y = \frac{(leftEye.y + rigthEye.y + Projection.y)}{3}, \quad (2)$$

де leftEye.y - ордината точки лівого ока, rigthEye.y - правого, projection.y - точки проекції;

4. Масштаб визначимо як відношення відстані між очима до половини ширини зображення(виходячи з того, що око повинне знаходитись посередині скельця, відстанню між скельцями знехтуємо)

5. Значення нахилу голови відносно нормалі отримаємо з результатів розпізнавання обличчя.

За переліченими вище даними можна намалювати ефект поверх вхідного зображення. Також слід врахувати масштаб вхідного зображення, адже розпізнавання обличчя відбувається для його абсолютних розмірів.

Розробка алгоритму для обробки відео-потоку.

Відео-потік - це набір кадрів, які змінюються із певною частотою. Для людського ока комфортна частота зміни кадрів - 24 FPS (framepersecond), мінімальна прийнятна частота кадрів 12-4 [3]. Потреба у забезпеченні плавності відео накладає жорсткі обмеження на алгоритм додавання елементів доповненої реальності.

Основним завданням алгоритму є мінімізація часу, необхідного для обробки фреймів вхідного потоку. Зменшення часових затрат на розпізнавання можна досягнути з використанням комбінованого методу обробки вхідних даних: проводити розпізнавання лише деяких кадрів із заданою періодичністю, а для кадрів, які не обробляються системою розпізнавання образів прогнозувати розташування необхідних опорних точок за допомогою регресійного аналізу.

Поняття і обґрунтування доцільності застосування регресійного аналізу для прогнозування опорних точок об'єктів.

Регресійний аналіз - це статична техніка для дослідження і моделювання взаємозв'язків між змінними. Лінійна регресія є підвидом регресійного аналізу - це лінійний підхід до моделювання зв'язку між скалярним результатом (залежною змінної) і однією або декількома описовими(незалежними) змінними. З загальному випадку розглядається вектор регресорів (незалежних змінних). Випадок коли досліджується взаємозв'язок між результуючою змінною(y) та однією незалежною змінною(x) називається простою лінійною регресією.[4]

Об'єкти на відео рухаються за певною криволінійною траєкторією. Оскільки регресійний аналіз для невідомої криволінійної функції є складною та ресурсоємкою задачею, її не доцільно використовувати в умовах обмежених ресурсів та часу для обчислень.

Криволінійну функцію можна апроксимувати до ламаної лінії(англ. polygonal chain) [5]. Аналіз та моделювання ламаної лінії є значно простішою, з точки зору необхідних ресурсів, задачею. Її можна розбити на лінійні ділянки та досліджувати кожну із них окрему, як двовимірну лінію, задану рівнянням

$$y = \hat{a} + \hat{b}x, \tag{3}$$

Для визначення параметрів рівняння$(\hat{a}, \hat{b})$ застосуємо просту лінійну регресію.[6]

$$\hat{b} = \frac{\sum_{i=1}^{n} x_i * y_i - n * \bar{x} * \bar{y}}{\sum_{i=1}^{n} {x_i}^2 - \bar{x}^2}, \tag{4}$$

$$\hat{a} = \bar{y} + \hat{b}\bar{x}, \tag{5}$$

де $\bar{y}, \bar{x}$ – середні вхідних даних(відомих значень координат опорних точок)

Опис алгоритму

1. Ініціалізувати список попередніх опорних точок recognizedLandmarks, граничний кут зміни вектору руху опорних точок – d = const, період розпізнавання кадрів – n = const, лічильник розпізнаних кадрів – currentIteretion = 0

2. Початкове розпізнавання
   2.1. Провести розпізнавання 1 кадру, записати отримані опорні точки у масив попередніх точок
   2.2. Провести позиціонування ефектів та відобразити оброблений кадр
   2.3. Провести розпізнавання наступного, записати отримані опорні точки у масив попередніх точок
   2.4. Провести позиціонування ефектів та відобразити оброблений кадр
3. Для кожного наступного кадру
   3.1. Провести регресійний аналіз для отриманих з попередніх розпізнавань опорних точок
   3.2. На основі функцій руху опорних точок визначити координати опорних точок для поточного кадру
   3.3. Провести позиціонування ефектів та відобразити оброблений кадр
4. Паралельно із пунктом 3 для кожного n-го кадру (виконувати у паралельному потоці)
   4.1. currentIteration++
   4.2. Провести розпізнавання кадру, додати розпізнані опорні точки у список recognizedLandmarks
   4.3. Перевірити напрямок вектору руху опорних точок: для пар опорних точок (recognizedLandmarks[current], recognizedLandmarks [current-1]) та (recognizedLandmarks[current-1], recognizedLandmarks[current-2]) побудувати прямі та порівняти їх коефіцієнти нахилу – якщо різниця коефіцієнтів нахилу більша ніж d – стався перегин кривої руху об'єкту –> повернутись до пункту 1; інакше продовжити

Для додавання елементів доповненої реальності на кожен фрейм, використовується описаний вище алгоритм для обробки зображень.

Висновки. Дане дослідження описує алгоритми накладання елементів доповненої реальності на зображення та відео-потік. Основне завдання описаних алгоритмів зменшити затримки при обробці відео-потоку та запобігти значному зменшенню показника плавності(кількості кадрів за секунду) вихідного відео-потоку.

Список використаної літератури

1. MobilevisionAPI [Електронний ресурс] – Режим доступу до ресурсу: https://developers.google.com/vision/introduction.

2. Design Patterns: Elements of Reusable Object-Oriented Software / E.Gamma, J. Vlissides, R. Johnson, R. Helm., 1994. – 395 с.

3. LaValle S. Vurtual reality / Steven LaValle. – Cambridge: Cambridge University Press, 2017. – 428 с.

4. Montgomery D. C. Introduction to linear regression analysis / D. C. Montgomery, E. A. Peck, G. G. Vining. – New Jersey: Wiley, 2012. – 642 с. – (1).

5. Hiroshi IMAI. Polygonal Approximations of a Curve — Formulations and Algorithms / Hiroshi IMAI, Masao IRI // Machine Intelligence and Pattern Recognition / Hiroshi IMAI, Masao IRI., 1988.

6. Seber G. A. Linear regression analysis / G. A. Seber, A. J. Lee. – New Jersey: WILEY, 2012. – 549 с. – (2).

ДОВІДКА ПРО АВТОРІВ

Коломієць Олег Олексійович – студент 6-го курсу кафедра обчислювальної техніки, факультету інформатики та обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kolomiiets Oleh – 6th year student of computer technics department of National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Kolomiiets Oleh

METHODS FOR INTEGRATION OF AUGMENTED REALITY ELEMENTS WITH IMAGE RECOGNITION SYSTEMS

The actuality of research. Augmented reality on mobile devices become very popular during the last couple of years. Mostly it related to increasing the performance of such devices. But still, there are issues with the processing of the video which requires image recognition for adding augmented reality elements. This issue is connected with the high complexity of image recognition that impacts on video FPS (frame per second) rate.

Problem setting. The high complexity of image recognition which leads to low FPS of the output video.

Target setting. Make investigation the problem of adding augmented reality element into photo and video. The aim of the research is to describe the algorithms that add elements of augmented reality to the image based on results from pattern recognition system and process video for the creation of augmented reality using linear regression.

Main material description. During this research was conducted an investigation of possible ways for improving video processing and created an algorithm which combines image recognition for getting coordinates of required landmarks on the image with regression analysis for prediction of these landmarks for frames when we are not running recognition algorithms. Linear regression is used for cases when the object is moving on the linear trajectory in another case we run additional recognition and mark all previous landmark as inappropriate for regression analysis (fallback mechanism)

Conclusions. This research describes algorithms for creation augmented reality based on results from image recognition system. The key purpose of video processing algorithms is reducing of the latency of output video and increase its FPS rate.

УДК 378.1

Віктор Порєв

ЕЛЕМЕНТИ ІНТЕЛЕКТУАЛЬНОЇ ПРОГРАМНОЇ СИСТЕМИ ДЛЯ ПІДТРИМКИ НАВЧАЛЬНОГО ПРОЦЕСУ

ELEMENTS OF THE INTELLIGENT SOFTWARE SYSTEM TO SUPPORT EDUCATIONAL PROCESS

The article deals with the questions of building an intellectual information system to support the educational process. Some aspects of data analysis are discussed in planning and supporting educational activities. The algorithmic basis for organizing interdisciplinary connections is considered. Provides suggestions on time analysis of learning processes.

Key words: algorithm, intellectual information system, support of educational process.

Fig.: 9. Bibl.: 6.

У статті розглядаються питання побудови інтелектуальної інформаційної системи для підтримки навчального процесу. Висвітлюється деякі аспекти аналізу даних при плануванні та підтримці навчальної діяльності. Розглядається алгоритмічна основа впорядкування міждисциплінарних зв'язків. Наводяться пропозиції щодо часового аналізу процесів навчання.

Ключові слова: алгоритм, інтелектуальна інформаційна система, підтримка навчального процесу.

Рис.: 9. Бібл.: 6.

Актуальність теми дослідження. Вимоги щодо вдосконалення освітньої діяльності, її гнучкості та здатності швидко адаптуватися до сучасних реалій пропагуються в усьому світі. Оптимізація та підтримка освітньої діяльності є доволі складною множиною різноманітних процесів, у яких потрібно враховувати безліч факторів. Як здається, допомогти людям у цій діяльності може розвинена комп'ютерна система, яка втілює ідеї штучного інтелекту.

Постановка проблеми. Обмеженість людських можливостей і необхідність охоплення величезних обсягів інформації, знань та досвіду, які накопичуються упродовж десятків років призводить до необхідності появи комп'ютерних інтелектуальних систем, які здатні не тільки накопичувати, а й аналізувати дані та процеси. Проблемою є відсутність готових рецептів побудови таких систем, зокрема, у галузі освітньої діяльності.

Аналіз останніх досліджень і публікацій. Темі використання комп'ютерних технологій для підтримки навчального процесу присвячено вже достатньо багато досліджень та публікацій. На Україні створено і відомо декілька діючих автоматизованих систем такого типу, зокрема, АІС "Навчальні плани НТУУ "КПІ" [1], технології фірми ПОЛІТЕК СОФТ [2] та інші [3].

В результаті аналізу літературних та рекламних джерел, присвячених системам підтримки автоматизації навчального процесу, можна зробити висновок, що ці системи будуються на платформах систем керування базами даних; і основна функція, яка реалізується у таких системах – зберігання та накопичення навчальних планів.

Суттєвим недоліком багатьох існуючих систем є те, що вони фактично підтримують тільки зберігання і накопичення вже готових документів, не надаючи можливостей автоматизованого проектування та планування таких навчального процесу.

Останні роки активно ведуться дослідження щодо впровадження ідей, методів та засобів штучного інтелекту, зокрема, створенню інтелектуальних систем, які самі здатні навчатися [4]. Відомі публікації щодо створення інтелектуальних систем підтримки освітньої діяльності [5].

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена висвітленню алгоритмічних та методологічних питань автоматизації планування, проектування та підтримки навчального процесу засобами інтелектуальних інформаційних систем.

Постановка завдання. Потрібно шукати та розроблювати методи і алгоритми, які дозволяють автоматизувати як рутинні так і достатньо творчі аспекти планування та проектування. Одним з шляхів на цьому напрямку є використання та розроблення методів інтелектуального аналізу даних, побудови інтелектуальних систем, які можуть навчатися і розвиватися.

Викладення основного матеріалу. Загальне функціонування інтелектуальної інформаційної системи для підтримки навчального процесу можна відобразити наступним чином(Рис. 1).

Далі варто зосередитися на деяких конкретних аспектах функціонування такої інформаційної системи.

Виявилось, що при плануванні зручною є спеціально розроблена автором статті модифікація діаграм Ганта. Як і у звичайних діаграм Ганта, кожна робота відображається відносно осі часу стрічкою, початок якої означає час початку роботи, а довжина стрічки означає тривалість роботи. Пропонується щоб висота стрічок-робіт була пропорційною обсягу роботи. Тоді роботи, які виконуються одночасно, малюються одна під одною, що дозволяє оцінити сумарний обсяг робіт, які виконуються. Для відображення освітнього процесу у часі кожна дисципліна відображається згідно розташуванню на числовій осі, яка поділена на семестри. Червона пунктирна лінія показує межу обсягів для кожного семестру (рис. 2).



Рис. 1. Основні етапи діяльності за рік

Аналіз міждисциплінарних зв’язків. В принципі, в навчальних програмах від викладачів слід вимагати тільки перелік дисциплін, які необхідно засвоїти для вивчення певної дисципліни Д і не обов’язково фантазувати, де ця дисципліна може бути у нагоді. А вже коли перелік дисциплін-попередниць буде включений у базу даних – модель навчального процесу, то система вже сама зможе знайти і відобразити усі дисципліни-наступниці, які будуть використовувати результати навчання вказаної дисципліни Д.

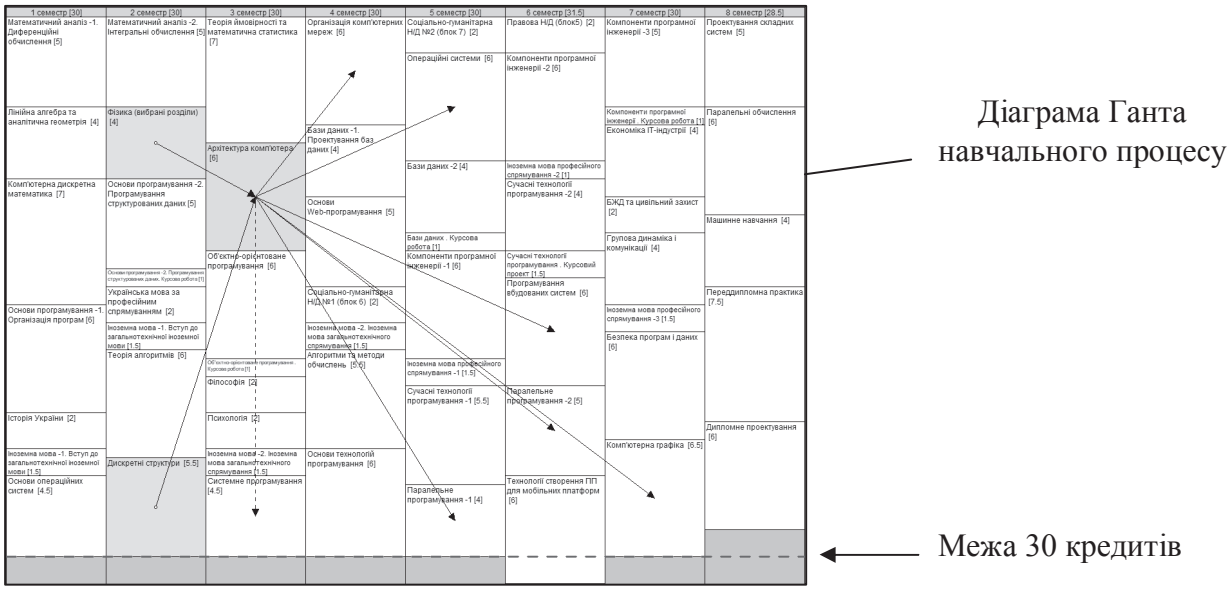


Рис. 2. Відображення міждисциплінарних зв’язків у інтелектуальній системі

Одною з проблем при плануванні є вірогідність циклів у графах міждисциплінарних зв’язків. Наприклад, викладач, який розробляє навчальну програму для дисципліни В, вважає, що для вивчення цієї дисципліни потрібні знання, навички, уміння, які є результатами засвоєння деяких інших дисциплін, зокрема, дисципліни А. Інший фахівець, який викладатиме дисципліну А, може вважати, що для вивчення цієї дисципліни необхідно у попередніх семестрах засвоїти декілька дисциплін, зокрема, дисципліну В. Це можна відобразити у вигляді графа зв’язків (рис. 3).

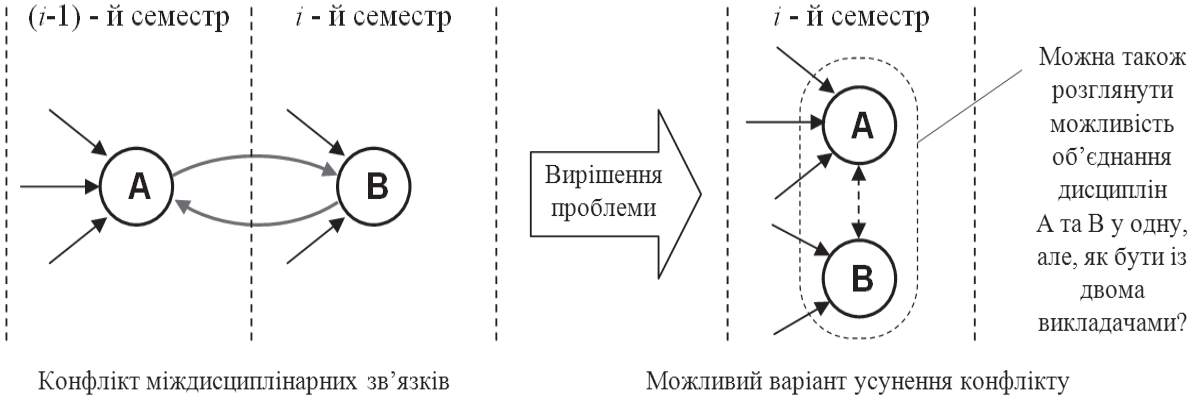


Рис. 3. Приклад простого конфлікту – мінімальний цикл на графі зв’язків

Аналіз коректності міждисциплінарних зв'язків може бути виконаний за допомогою алгоритму впорядкування у часі, запропонованому у [6].

Нехай загальна кількість елементів буде *N*. Алгоритм складається з двох кроків:

Крок 1. Усі елементи розташовуються на першій стадії часу.

Крок 2. Виконується основний цикл впорядкування елементів у часі:

```
for (int i=0; i<=N; i++)
 if (MoveToNextStages() == 0) break;
```

У тілі циклу декілька разів виконується функція

```
int MoveToNextStages()
{
int done = 0;

for (int i=0; i<N; i++)
  {
  mxstprev = MaxStagePrev(i);
  if (стадія i-го елементу <= mxstprev)
    {
    MoveToStage(i, mxstprev+1);
    done = 1;
    }
  }
return done;
}
```

Функція **MaxStagePrev(i)** знаходить для *i*-го елемента максимальну стадію елемента-попередника. Процедура **MoveToStage(i,s)** розташовує *i*-й елемент на стадії *s*.

Основний цикл кроку 2 наведеного вище алгоритму виконується максимум *N*+1 ітерацій. Чому *N*+1 а не *N*? Це зроблено для виявлення циклів у зв'язках елементів. Якщо елементів усього *N*, то їх можна розташувати максимум по *N* стадіям часу (це буде суцільний послідовний ланцюжок по одному елементу на кожну стадію). Якщо якийсь елемент займе *N*+1 стадію, то це вже буде розрив у часі – помилка. Відповідно цей елемент належить до циклу і його зв'язки необхідно проаналізувати для розв'язання конфлікту. При наявності циклів у зв'язках елементи будуть безкінечно пересуватися на наступні стадії. Тому зроблене обмеження кількості ітерацій – остання *N*+1-а ітерація призначена для контролю можливості впорядкування елементів.

Якщо циклів немає, то цей алгоритм може завершитися суттєво раніше до *N*-ї ітерації. Це залежить від порядку розташування елементів у робочих масивах.

Проектування навчального процесу. Розробка моделі навчального процесу та підготовка основних керівних документів, таких як, освітня програма, навчальний план тощо, є складним ітераційним процесом, у якому потрібно враховувати багато обмежень, рекомендацій, а також попередній досвід роботи у цій галузі. Інтелектуальна інформаційна система значною мірою допомагає це зробити за потрібний час з мінімізацією помилок.

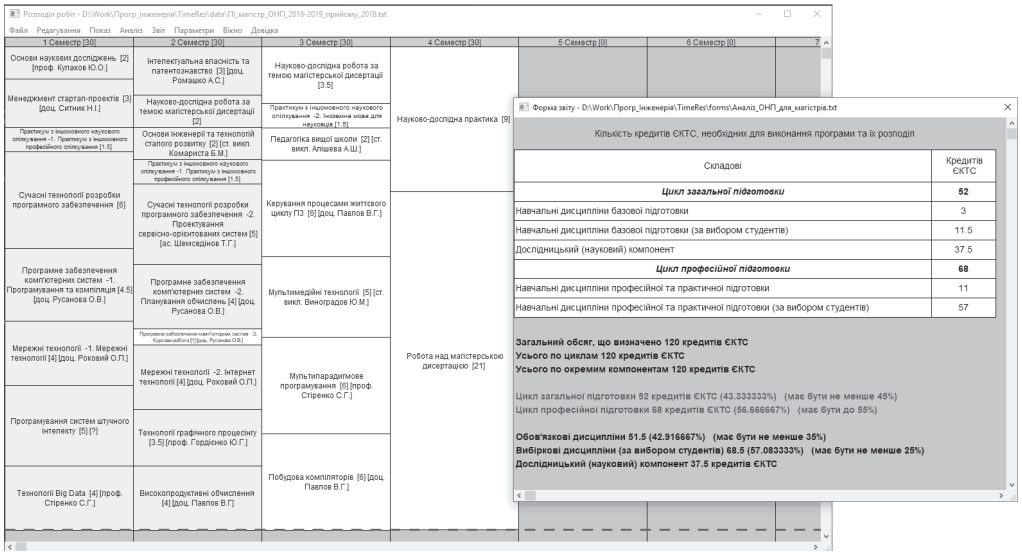


Рис. 4. Контроль помилок на етапі розробки освітньої програми

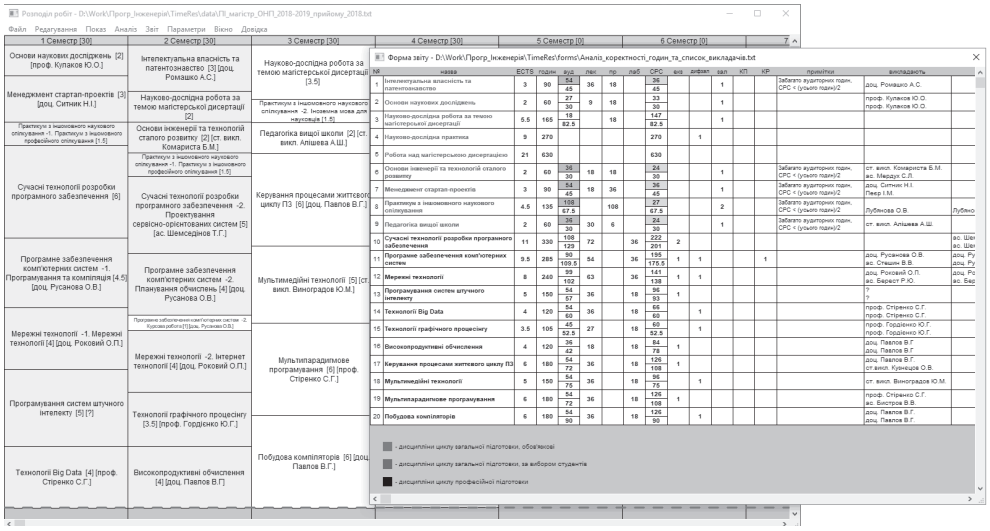


Рис. 5. Контроль помилок на етапі розробки навчального плану

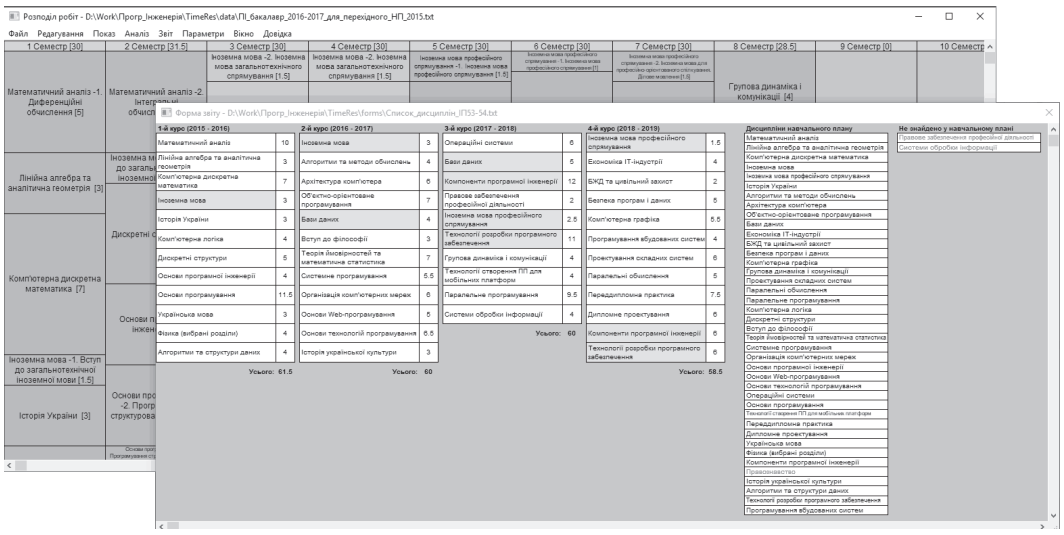
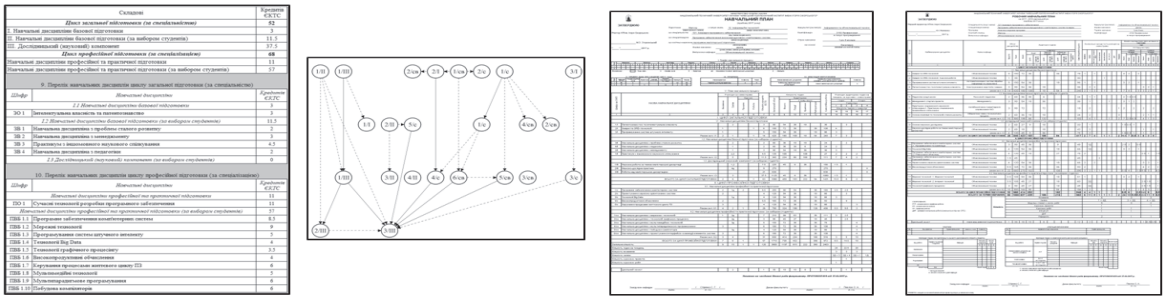


Рис. 6. Аналіз викладання дисциплін за чотири роки

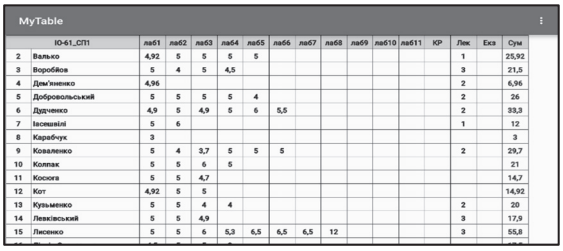
Формування звітів. На основі моделі, яка зберігається у вигляді даних та правил, достатньо легко можна отримувати потрібні документи-звіти



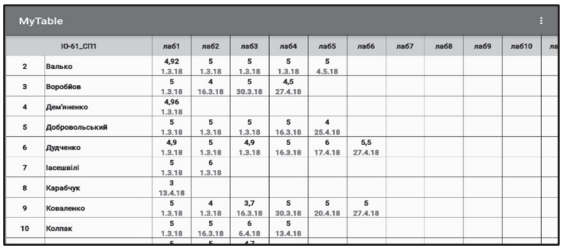
Складові освітньої програми | Навчальний план | Робочий навчальний план

Рис. 7. Основні документи – продукти інформаційної системи

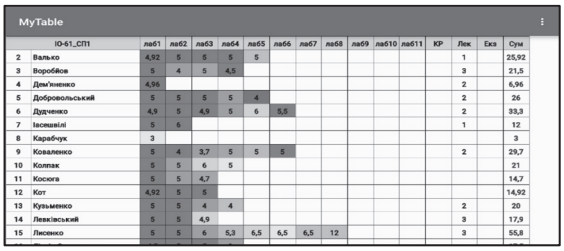
Підтримка поточної роботи викладача. Розвиток інформаційних комп'ютерних технологій засобів надає широкі можливості для підтримки навчального процесу. Приклад – розроблена автором статті програма для обліку та аналізу успішності студентів. Програма зветься MyTable і працює на планшетах та смартфонах у середовищі ОС Android.



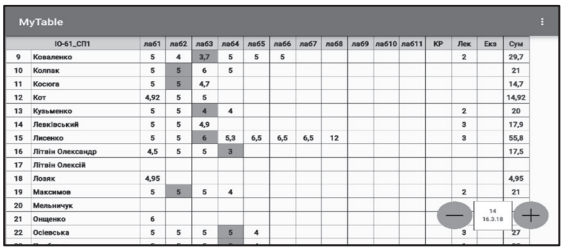
а) звичайний вигляд таблиці групи



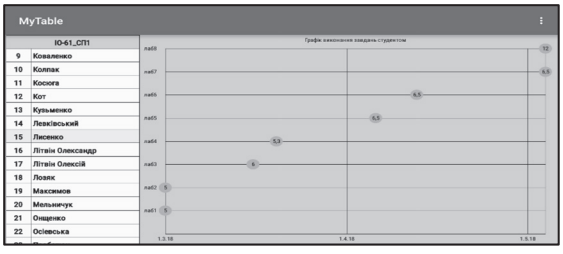
б) додатково відображаються дати виставлення оцінок



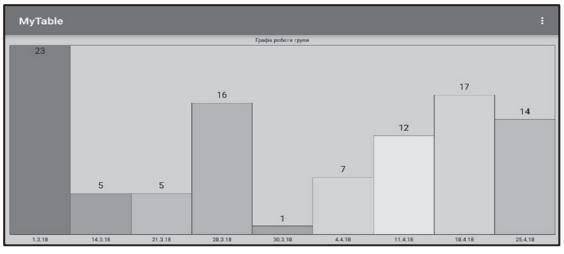
в) кольором позначається час оцінок



г) часовий зріз – виділено оцінки, отримані у вказаний день



д) графік роботи студента – терміни виконання завдань та оцінки



е) графік роботи групи – кількість отриманих оцінок по дням занять

Рис. 8. Вигляд електронної таблиці обліку та аналізу успішності студентів на мобільному пристрої

Одної з особливостей такої мобільної таблиці є автоматичне збереження дати у момент виставлення оцінки. Це відкриває додаткові можливості викладачу аналізувати та оцінювати успішність кожного студента і усього навчального процесу у цілому, робити висновки щодо вдосконалення планування та поточної роботи. Наприклад, якщо розглядати графіки успішності, то можна виділити три основні різновиди

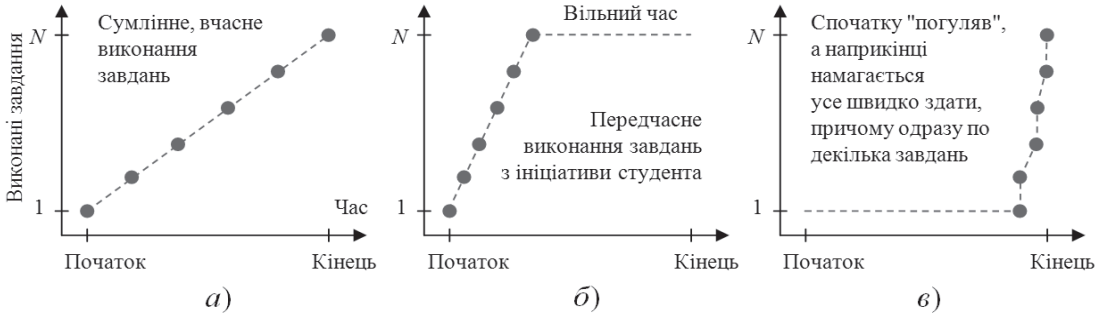


Рис. 9. Три основні різновиди графіків успішності студентів

Досвід показує, що вирішення питання кінцевої оцінки є нетрівіальним завданням – не можна однозначно визначити, наприклад, що студент із графіком (*в*) завжди знає матеріал гірше студентів, у яких графіки типу (*а*) або (*в*). Графік успішності – часовий ряд можна використати як елемент навчально-особистого портрету студента, який складатиметься із різноманітних відомостей про студента. На його основі можна робити аналіз, класифікацію, прогнозувати досягнення і, відповідно, адаптувати навчання. Множина різноманітних відомостей про багатьох студентів та викладачів за декілька років є основою бази даних, яка, в свою чергу, буде доповнюватися правилами логічного виводу, що буде складати базу знань. Так буде відбуватися поглиблення впровадження методів та засобів штучного інтелекту.

Висновки.Розглянуто підхід до створення інтелектуальної інформаційної системи для підтримки навчального процесу. Запропоновано алгоритмічну та методологічну основу для побудови елементів комп'ютерної інтелектуальної інформаційної системи. Наведені приклади реалізації окремих функцій системи підтримки навчального процесу.

Список використаних джерел

1. АІС "Навчальні плани". Конструкторське бюро інформаційних систем НТУУ "КПІ". [Електронний ресурс]. – Режим доступу: http://kbis.kpi.ua/kbis/index.php?option=com_content&task=view&id=84&Itemid=91

2. ПОЛІТЕК СОФТ. Програмне забезпечення для вищих навчальних закладів України. [Електронний ресурс]. – Режим доступу: http://www.politek-soft.kiev.ua/

3. Медовая Е.В. Автоматизация процесса создания учебного плана в ОНАПТ // Автоматизація технологічних і бізнес-процесів. – 2014, Volume 6, Issue 4, С. 151-154. [Електронний ресурс]. – Режим доступу: http://www.journal-atbp.com/journals/atbp20.pdf

4. Интеллектуальная информационная система//Википедия https://ru.wikipedia.org/wiki/Интеллектуальная_информационная_система

5. И. И. Казмина, Е. В. Нужнов. Интеллектуальная поддержка образовательных процессов на уровне специальности (профиля) // Открытое образование. – 2013. – №6. – С.80-84.

6. Порєв В. М. Алгоритмічні аспекти автоматизованого проектування навчальних планів // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. пр. – К.: Век+. – 2016. – № 64. – С.84-89.

ДОВІДКА ПРО АВТОРІВ

Порєв Віктор Миколайович – кандидат наук, старший викладач, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Victor Porev – senior lecturer, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: v_porev@ukr.net

V.M. Porev

ELEMENTS OF THE INTELLIGENT SOFTWARE SYSTEM TO SUPPORT EDUCATIONAL PROCESS

Relevance of research topic. The requirements for improving the educational activity, its flexibility and the ability to quickly adapt to current realities are promoted all over the world.

Formulation of the problem. Limited human capabilities and the need to cover the vast amounts of information, knowledge and experience that accumulate over decades lead to the emergence of computer intelligence systems that can not only accumulate, but also analyze data and processes. The problem is the lack of ready-made recipes for constructing such systems, in particular, in the field of educational activities.

Analysis of recent research and publications. As a result of the analysis of literary sources devoted to the systems supporting the automation of the educational process, we can conclude that these systems are based on the platforms of database management systems; and the main function that is implemented in such systems - the storage and accumulation of curricula.

Selection of unexplored parts of the general problem. This article is devoted to the coverage of algorithmic and methodological issues of automation of planning, designing and supporting the educational process by means of intelligent information systems.

Setting objectives. It is necessary to search and develop methods and algorithms that allow to automate both routine and sufficiently creative aspects of planning and design.

Presentation of the main material. The methodological and algorithmic aspects of building an intellectual information system are considered. The algorithm of control and regulation of interdisciplinary connections is proposed. The modification of Gantt charts is proposed. The main aspects of storage and analysis of time series for educational activities are highlighted.

Conclusions. An approach to creating an intellectual information system for the support of the educational process is considered. The algorithmic and methodological basis for construction of elements of computer intellectual information system is proposed. Examples of implementation of separate functions of the educational process support system are given.

Key words: algorithm, intellectual information system, support of educational process.

Секція 4. GN (Глобальні мережі, grid та cloud системи). Section 4. GN (Global networks, grid and cloud systems).

UDC 004.8

Ярич Андрій

ПИТАННЯ ВИБОРУ CLOUD-ДОДАТКУ НА ОСНОВІ АНАЛІЗУ GOOGLE, AMAZON ТА APPLE РІШЕНЬ

QUESTION OF CLOUD APP CHOOSING ON GOOGLE, AMAZON AND APPLE ANALYSIS

У статті розглядається питання вибору програмного додатку хмарної технології. Для початку використовується набір даних про основні характеристики Google, Apple та Amazonрішення. Вибір базується на основі аналізу ключових параметрів. Розглядаються рішення AmazonAWS (S3), GoogleDrivета AppleiCloud.

Ключові слова: iCloud, Amazon, GoogleDrive, шифрування, AES.

Рис.: 3.Табл.: 0. Бібл.: 4.

The article deals with the choice of a software application for cloud technology. To begin with, it “uses” a set of data on the key features of Google, Apple and Amazon solutions. The choice is based on the analysis of key parameters. Amazon AWS (S3), Google Drive and Apple iCloud samples were considered.

Key words: iCloud, Amazon, Google Drive, encryption, AES.

Fig.: 3. Tabl.: 0. Bibl.: 4.

Target setting. Cloud technologies become more and more demand in the world; people use storage services for documents, photo and another files so it’s important to have a file protection and fast downloading. Due to these facts cloud services are actual topic for discussion.

Actual scientific researches and issues analysis. In connection to the invention of new methods and approaches in the field of cloud technologies, the topic of cloud choosing has become more studied in recent years.

Uninvestigated parts of general matters defining. There are a lot of uninvestigated moments that should be opened. I am mostly going to focus on the 1TB to 2TB range for ease of comparison, and because that should address the needs of most users. I am also sticking with the big names.

The research objective. The purpose of this paper is to investigate the applications of the iCloud, AMAZON AWS (S3) and Google Drive. As a solution, the article will focus on data protection and price comparison, on choosing the best application.

The statement of basic materials. After your data reaches Google, it gets unencrypted then re-encrypted using 128-bit AES. While not the 256-bit algorithm that most other services use, this is still normal. This is done on-the-fly before the data is actually stored, which prevents the possible leakage of unencrypted data on their hard drives. The AES encryption keys that were used to encrypt data are then, themselves,

encrypted with a rotating set of master keys. This adds another layer of security by requiring a *second* set of encryption keys.

Amazon offers what is called ‘Server Side Encryption’ for S3, or if user is looking at S3 bucket in the AWS console, it will be called ‘Default encryption’ under the properties sheet. Consumer can enable encryption simply by toggling a button in the UI and telling AWS how he/she would like the key to be managed. One option is AES-256 keys managed per bucket, and the other option is allow KMS — the AWS Key Management Service — handle key management.

iCloud secures information by encrypting it when it's in transit, storing it in iCloud in an encrypted format, and using secure tokens for authentication. For certain sensitive information, Apple uses end-to-end encryption that provides high level of data security. Information is protected with a key derived from information unique to device, combined with device passcode.

General model structure. The structure of the model was chosen to compare the level of security models. First is AMAZON:

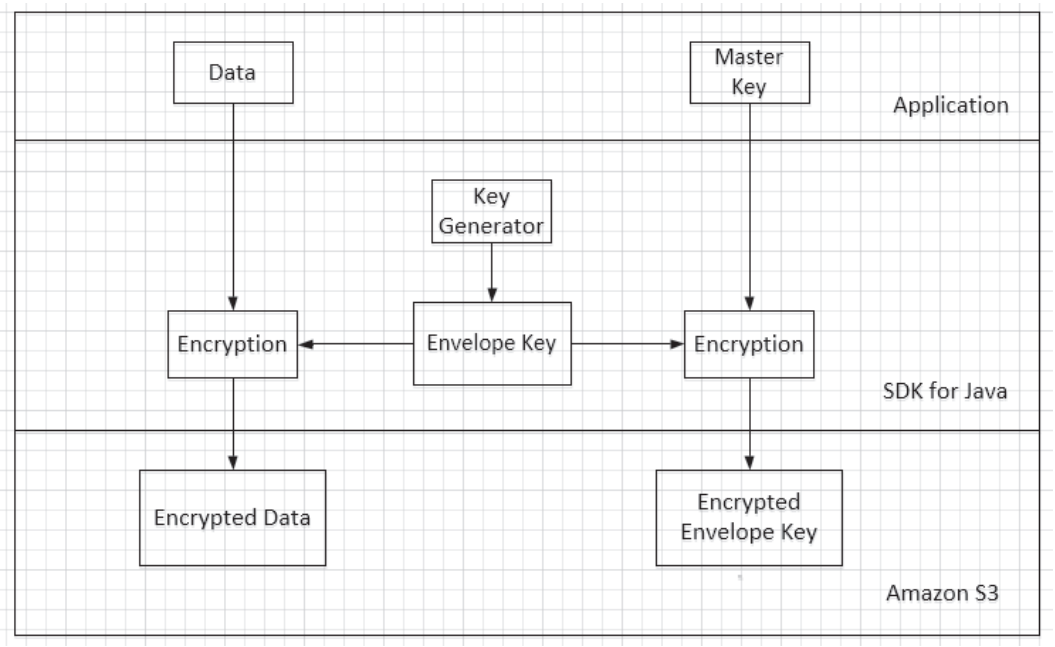


Fig. 1. Client-Side Data Encryption for Amazon S3

Second thing is Google’s variant:

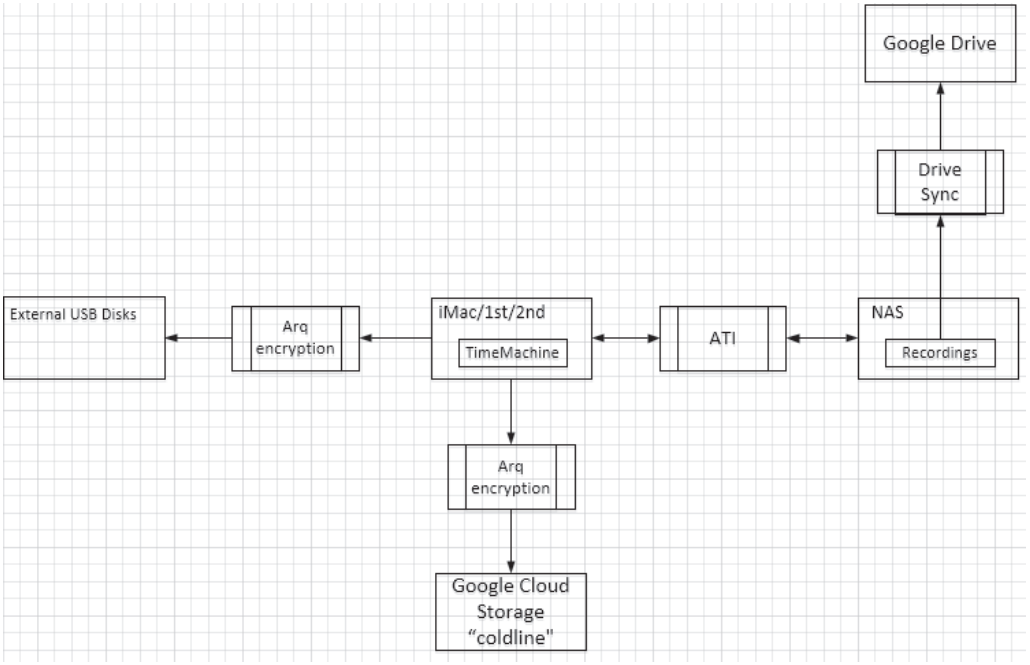


Fig. 2. Google Data Protection strategy

Third thing is Apple's variant:

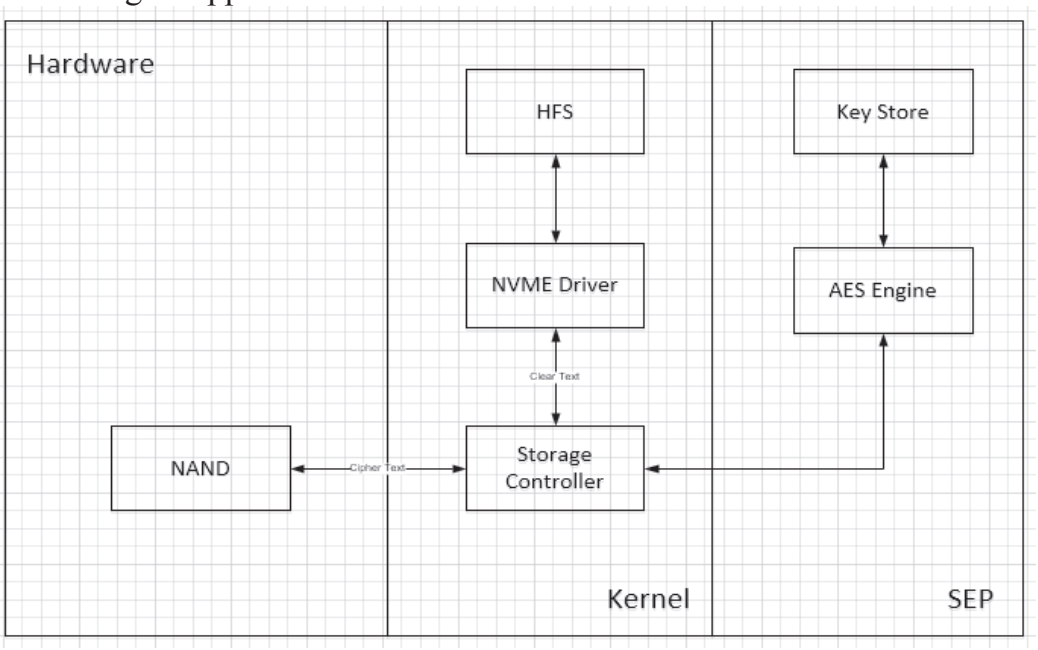


Fig. 3. iCloud Filesystem Data Protection

Conclusions. The paper has demonstrated the ability of iCloud to be the best cloud based storage. It can be seen that the use of such service produces qualitative results. ICloud has enough data protection level and price is cheaper than competitors have.

There are several directions for future work. One is to change the test protection's model for all services, increasing the number of hidden units and adding more productive parts. Another is to create a completely new cloud scheme. These changes will definitely improve the results. It also would be interesting to test the hybrid model that consists of 3 templates.

References

1. Google Cloud Platform Documentation. – Link:https://cloud.google.com/docs/. – Date: 22.04.2018.

2. How Do I Enable Default Encryption for an S3 Bucket? – Link:https://docs.aws.amazon.com/AmazonS3/latest/user-guide/default-bucket-encryption.html. – Date: 22.04.2018.

3. How the Apple iCloud Works. – Link: https://computer.howstuffworks.com/cloud-computing/icloud4.htm. – Date: 25.04.2018.

4. Price comparison. – Link:https://www.wired.com/story/amazon-just-killed-the-best-deal-in-tech/. – Date: 25.04.2018.

ДОВІДКА ПРО АВТОРІВ

Ярич Андрій Миколайович – магістрант, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Yarych Andriy – magistrate, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: andriyarych@zoho.com

Ярич Андрій

ПИТАННЯ ВИБОРУ CLOUD-ДОДАТКУ НА ОСНОВІ АНАЛІЗУ GOOGLE, AMAZON ТА APPLE РІШЕНЬ

Актуальність теми дослідження. Хмарні технології стають все більш й більш поширеними на світі; люди використовують сервіси зберігання для документів, фото та інших файлів, тому так важливо мать захист даних та їх швидке завантаження.

Постановка проблеми. На даний момент існує достатньо широка необізнаність населення щодо потенціалу та репутації різноманітних хмарних додатків, тому слід виділити найкращий, що повністю задовольнятиме потребам користувачів.

Аналіз останніх досліджень і публікацій. У зв'язку з появою нових методів та підходів на ниві хмарних технологій тема вибору хмарних додатків стає більш популярною в останні роки.

Виділення недосліджених частин загальної проблеми. Існує достатньо невивчених моментів, які повинні буть відкритими. Слід звернути увагу на рішення з об'ємом в 1 та 2 ТБ, оскільки це найбільш релевантно для багатьох юзерів; акцент був поставлений на програмах від відомих компаній.

Постановка завдання. Завданням є повністю розкрити особливості усіх рішень; на основі отриманих даних (у більшій мірі – від рівня захисту інформації та ціни послуг) вибирається найбільш актуальне та потрібне програмне рішення.

Викладення основного матеріалу. Перш ніж дані залишать пристрій, він зашифровується за стандартом TLS. Це той самий стандарт, який використовується для шифрування веб-переглядачів на веб-сайтах із безпечним HTTPS. Після того, як дані досягають Google, вони шифруються за допомогою 128-бітної AES. Хоча це не 256-бітний алгоритм, який використовують більшість інших служб, однак непогано.

Amazon пропонує те, що називається «Шифрування за замовчуванням» під панеллю властивостей. Можна ввімкнути шифрування, просто перемикаючи кнопку в інтерфейсі користувача та повідомляючи AWS, як хочеться керувати додатком. Одним з варіантів є керування ключами AES-256, а інший варіант - дозволити KMS - службі керування ключами AWS – працювать із ключами.

Про iCloud. Щоб використовувати повне шифрування, то для Apple ID потрібно мати двофакторну автентифікацію. Щоб отримати доступ до даних на новому пристрої, може знадобитися введення паролю для існуючого або колишнього пристрою Apple.

Висновки. Дана стаття продемонструвала першість "яблука" серед всіх розглядуваних додатків; було чітко видно, що використання даного сервісу дає результати відносно юзера та його сподівань; є декілька напрямів подальшого розкриття питання. Наприклад, можна протестувати гібридну модель хмари або створити якісно нову; також за ідею можна взять покращення системи захисту інформації iCloud.

Ключові слова:iCloud, Amazon, GoogleDrive, шифрування, AES.

УДК 004.027

Максим Бондарчук,
Юрій Зорін

АЛГОРИТМ РОЮ ЧАСТИНОК ПОШУКУ НАЙКОРОТШОГО ШЛЯХУ В ТЕЛЕКОМУНІКАЦІЙНІЙ МЕРЕЖІ

PARTICLE SWARM OPTIMIZATION ALGORITHM FOR SHORTEST-PATH FINDING IN TELECOMMUNICATION NETWORKS

Вроботі представленомодифікований алгоритм рою частинокдля задачі пошуку найкоротшого шляхув телекомунікаційних мережах. Завдяки застосуваннюкількох модифікацій, серед яких метод селективноїрегенерації й нова версія кодування на основі пріоритетів,продуктивність алгоритмубула суттєво підвищена.

Ключові слова: алгоритм рою частинок, селективна регенерація, кодування на основі пріоритетів, скидання швидкості.

Рис.: 1. Табл.: 1. Бібл.: 6.

The paper presents a modified Particle Swarm Optimization algorithm for shortest-path finding in telecommunication networks. Due to introduction of several modifications, among them selective particle regeneration methodand a new version of priority-based encoding, the algorithm performance has been substantially improved.

Key words: particle swarm optimization, selective regeneration, priority-based encoding, velocity reinitialization.

Fig.: 1. Tabl.: 1. Bibl.: 6.

Актуальність теми дослідження. В наш час важко уявити життя без телекомунікаційних мереж. Мережа інтернет, яка є телекомунікаційною мережею, розширюється з кожним днем долучаючи тисячі нових користувачів. Україна є однією з країн, в якій інтернет розвивається швидкими темпами. Незважаючи на те, що майже всі куточки планети покриті мережею, у багатьох місцях все ще наявна незначна пропускна спроможність мережі. В цьому випадку ефективне розв’язання задачі пошуку найкоротших шляхів між заданими вузлами є надзвичайно важливою.

Постановка проблеми. Відсутність добре досліджених методів розв’язання проблеми передчасної збіжності алгоритму рою частинок та способів його застосування до задачі пошуку найкоротшого шляху в мережі робіть їх досить актуальними.

Аналіз останніх досліджень і публікацій. Останніми роками метаевристичні алгоритми викликають все більший інтерес дослідників. Зокрема запропоновано застосування таких алгоритмів як імітація відпалу, оптимізація за допомогою мурашиної колонії та інших до розв’язання задачі пошуку найкоротших шляхів в мережі.

Виділення недосліджених частин загальної проблеми. Робота присвячена дослідженню модифікацій алгоритму рою частинок та його застосуванню до пошуку найкоротшого шляху в мережі. Незважаючи на те, що запропоновано досить велика

кількість модифікацій алгоритму рою частинок, питанню удосконалення когнітивної поведінки рою приділено недостатню увагу.

Постановка завдання. Завданням роботи є розробка модифікованого алгоритму рою частинок, що має більш високийпоказник успішності (доля знаходження правильного шляху на 100 тестів)в порівнянні з оригінальним алгоритмом при застосуванні його до задачі пошуку найкоротшого шляху в мережі.

Викладення основного матеріалу. Оригінальний алгоритм рою частинок [1,2] створений для знаходження екстремуму функції багатьох змінних $f(x_1, x_2, \dots, x_k)$. Всі частинкирою є потенційними розв'язками і на етапі ініціалізації розміщується в довільних точках простору пошуку $X_i(x_1, x_2, \dots, x_k)$. На кожній ітерації частинкипереміщуються за рахунок зміни швидкості $V_i(v_1, v_2, \dots, v_k)$. Для зміни швидкості використовується інформація про поточнерозміщення, а також інформація про попередні розв'язки, які були знайдені даною частинкою P_i, а також всіма іншими частинками рою G.

Швидкість обчислюється за формулою

$$V_{id} = V_{id} + \varphi_p r_p (P_{id} - X_{id}) + \varphi_g r_g (G_d - X_{id}), \quad (1)$$

де φ_pта φ_g– когнітивний і соціальний параметри, якими скеровується напрям руху в бік відповідно локально чи глобально найкращих розв'язків (зазвичай це дійсні числа більше 2), r_pта r_g– випадкові числа в діапазоні [-1, 1], $i = 1, 2, \dots, N$, N– розмір популяції, $d = 1, 2, \dots, D$– кількість вимірів функціїf.

Формула переміщення частинки визначається як

$$X_{id} = X_{id} + V_{id}. \quad (2)$$

Обмін інформацією між частинками здійснюється через вектор G (найкращий отриманий популяцією розв'язок).

З метою покращення можливостей локального пошуку алгоритму в перспективних областях вводиться множник w інерції швидкості, який зменшується по мірі виконання ітерацій. Внаслідок цього частинки набувають властивостей, які забезпечують ретельніше дослідження перспективних областей простору пошуку, про які стало відомо на попередніх ітераціях. Інерцію швидкості обчислюється як

$$w_{iter} = w_{max} + (w_{max} - w_{min}) \frac{iter}{I}, \quad (3)$$

де w_{max} та w_{min} – верхня та нижня границі інерції відповідно та є константами, які визначаються емпірично залежно від функції, що оптимізується, $iter$ – номер поточної ітерації алгоритму, I – максимальна кількість ітерацій.

Ще однією з модифікацій алгоритму рою частинок є застосування методу звуження множника.Цей метод дозволяє більш точно скеровувати напрям руху частинки з використанням коефіцієнтівφ_p таφ_g. Коефіцієнт для формули швидкості обчислюється за формулою

$$\chi = \begin{cases} 2\left(\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|\right)^{-1}, \varphi = \varphi_p + \varphi_g > 4 \\ 1, \varphi = \varphi_p + \varphi_g \le 4 \end{cases}. \quad (4)$$

З урахуванням даних модифікацій формула зміни швидкості набуває вигляду

$$V_{id} = \chi\left[wV_{id} + \varphi_p r_p (P_{id} - X_{id}) + \varphi_g r_{2g} (G_d - X_{id})\right]. \quad (5)$$

Однією з проблем базового алгоритму рою частинокє його передчасна збіжність до локальних оптимумів. Для вирішення даної проблеми застосуємо підхід, що називається скидання швидкості. Метод полягає у тому, що кожна частинка на кожній ітерації з певною ймовірністю p_l,яку має назву ймовірності усунення, переміщується в довільну точку простору [3].

```
Згенерувати початкову конфігурацію для кожної частинки:
Виконати довільну ініціалізацію X_i
Виконати довільну ініціалізацію V_i
    Обчислити path_i
    Обчислити значення функції пристосованості f(X_i)
    P_i ← ∞
    G ← ∞
iteration_count ← 0
// max_iteration – максимальна кількість ітерацій

while (iteration_count<max_iteration)
для кожної частинки i
                Обчислити V_i із застосуванням інерції швидкості та звуження
            множника
        Оновити X_i
        Обчислити path_i
        Обчислитизначення функції пристосованості f(path_i)
        if f(path_i) < f(P_i)
                P_i ← path_i
        if f(P_i) < f(G)
                G ← P_i
        Застосувати скидання швидкості
        Застосувати метод селективної регенерації

    iteration_count ← iteration_count + 1
end while
returnG
```

Рис. 1. Псевдокод модифікованого алгоритму рою частинок.

З метою покращення пошукової здатності алгоритму застосовано дещо видозмінений метод, що має назву селективна регенерація частинок [4]. Метод полягає у визначенні «відстані» в термінах значення функції пристосованості між кожною частинкою рою та глобально найкращою частинкою. Серед тих частинок, «відстань» яких є меншою за аналітично задане значення Ω довільним чином вибирається d відсотків, для яких виконується реініціалізація. Основною метою цього є намагання допомогти якійсь кількості частинок вирватись з зони можливого локального мінімуму, до якого може потрапити найкраща частинка й потягти за собою інші. Оскільки певна кількість частинок щойно вийшла з околиці найкращої частинки, є доцільним на певний час змінити когнітивні

параметри поведінки, щоб дані частинки не повернулисямиттєво до найкращої. Для цього будемо змінювати значення φ_p на $2\varphi_p$ для певної кількості ітерацій z.

Для застосування модифікованого алгоритму рою частинок до розв'язаннязадачі знаходження найкоротшого шляху в графі використаємо кодування шляху з урахуванням пріоритетів[5,6]. Метод полягає в тому, що кожній вершині графу присвоюється пріоритет з діапазону [0, 1] із врахуванням її поточного розміщення. Простір пошуку розбивається на області й кожній області призначається певний пріоритет. В процесі побудови шляху при переході з поточної вершини графа частинка серед ребер, що не належать до вже пройденого шляху, обирає те, яке має мінімальне значення добутку ваги ребра на пріоритет. Зворотній процес називається декодуванням шляху з урахуванням пріоритетів.

Результати комп'ютерного тестування. Для тестування запропонованого алгоритму згенеровано штучні набори даних, які представляють мережі з різною кількістю вузлів та ребер. Виконано порівняльний аналіз оригінального (PSO) та модифікованого (MPSO) алгоритмів. В табл. 1 наведені параметри конфігураціймережі коефіцієнти успішності алгоритмів (successrate – SR), тобто доля знаходження найкоротшого шляху, що взято як показник міри якості алгоритму. Всі результати усереднені на 50 запусках алгоритмів. При тестуванні використано параметри $\varphi_g = 1.5$, $\varphi_p = 2.5\,\phi_p = 2.5$,$\Omega = 0.1$,$c = 0.5$,$z = 15$.

Таблиця 1

Порівняльний аналіз алгоритмів.

Топологія мережі	Кількість вузлів	Кількість ребер	SR PSO	SR MPSO
1	50	500	0.8655	0.9008
2	60	720	0.8207	0.9164
3	70	980	0.8281	0.9161
4	80	1280	0.8260	0.9269
5	90	1620	0.8298	0.9054
6	100	2000	0.8646	0.9390
7	110	2420	0.8571	0.9218
8	120	2880	0.8701	0.9018
9	130	3380	0.8750	0.9224
10	140	3920	0.8361	0.9099

Висновки. Результати тестування продемонстрували, що запропонований алгоритм має кращі показники успішності приблизно на8% порівняно з оригінальним, а усереднений коефіцієнт успішності є більшим за 0.9. Більш детальним аналізом встановлено, що кожна із запропонованих модифікацій дає приблизно однаковий вклад в покращення продуктивності алгоритму (в межах 2-3%).

Предметом подальших досліджень можуть бути модифікації алгоритму, пов'язані з застосуванням шумовихметаевристик, а також розбиттяколонії частинок на групи й застосування в якості Gнайкращої частинки в групі.

Список використаних джерел.

1. J. Kennedy, R. Eberhart. Particle Swarm Optimization.Proceedings of IEEE International Conference on Neural Networks. 1998. Vol. 65, №. 3.P.1942- 1948.
2. Zhang Y. A Comprehensive Survey on Particle Swarm Optimization Algorithm and its Applications. Mathematical Problems in Engineering. 2015. V.56. P.1-26.
3. P. Sharma, N. Khurana. Study of Optimal Path Finding Techniques. Proceedings of IEEE International Conference on Neural Networks. VII. 2001. P.124-130.
4. Chi-Yang Tsai,I-Wei Kao. A Particle Swarm with Selective Particle Regeneration for Multimodal Functions. Wseas Transactions on Information Science abd Applications. 2009. P.242-250.
5. Зорін Ю.М., Бондарчук М.Ю Алгоритм рою частинок для знаходження найкоротшого шляху в графі.Прикладна математика та комп'ютинг-2017: збірник тез доповідей. 19-21 квіт. 2017 р. Київ: Просвіта. 2017. С.22-26.
6. Зорін Ю.М., Бондарчук М.Ю Алгоритм рою частинок для знаходження найкоротшого шляху в графі.Прикладна математика та комп'ютинг-2018: збірник тез доповідей. 21-23 бер.. 2018 р. Київ: Просвіта. 2018. С.49-53.

ДОВІДКА ПРО АВТОРІВ

Зорін Юрій Михайлович – доцент, кафедра системного програмування і спеціалізованих комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Zorin Yury – associate professor, System Programming and Specialized Computer Systems Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: yzorin@gmail.com

Бондарчук Максим Юрійович – магістрант, кафедра системного програмування і спеціалізованих комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Bondarchuk Maksym – student, System Programming and Specialized Computer Systems Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: bondarchuk.m.y@gmail.com

Maksym Bondarchuk,
Yury Zorin

PARTICLE SWARM OPTIMIZATION ALGORITHM FOR SHORTEST-PATH FINDING IN TELECOMMUNICATION NETWORKS

Relevance of research topic. In our time it's hard to imagine life without telecommunication networks. The Internet, which is a telecommunication network, is expanding every day by engaging thousands of new users. Even though almost all the planet corners are covered by the network, in many places there is still a small bandwidth of the network. In this case, an effective solution to the problem of finding the shortest paths between the given nodes is an extremely important task.

Target setting. The lack of well-researched methods for solving the problem of premature convergence of the particle swarm algorithm and its application to the problem of the shortest path finding in networks made it an actual topic.

Actual scientific researches and issues analysis. In recent years, meta-heuristic algorithms are of increasing researchers interest. In particular, the application of such algorithms as Simulated Annealing, Ant Colony Optimization and others to the solution of the shortest paths finding problem in the network have been proposed.

Uninvestigated parts of general matters defining. The paper is devoted to the study of the particle swarm algorithm modifications and its application to finding the shortest path in the network. Despite the fact that a sufficiently large number of modifications of the particle swarm algorithm has been proposed, insufficient attention to the issue of improving cognitive behavior has been paid.

The research objective. The lack of well-studied methods for solving the problem of premature convergence of the particle swarm algorithm and its application to the problem of finding the shortest path in networks make it an actual topic.

The statement of basic materials. After a thoroughanalysis of the original particle swarm optimization algorithm analysis performed an introduction of the following modifications: selective particle regeneration, velocity reinitialization, construction factor method, particle killing method and priority-based encoding technique have been proposed. Computer experiments have shown that all the modifications combined give around 8% in success rate improvement.

Conclusions. The test results have shown that the proposed algorithm has better performance compared to the original one. The objective of further research may be modifications of the algorithm associated with the use of noising metaheuristics.

Key words: particle swarm optimization, selective regeneration, priority-based encoding, velocity reinitialization.

УДК 004.724.4

Олександр Калюжний,
Юрій Кулаков,
Михайло Діброва

СПОСІБ ОРГАНІЗАЦІЇ БАГАТОШЛЯХОВОЇ МАРШРУТИЗАЦІЇ З ВИКОРИСТАННЯМ ЧАСТКОВО-ПЕРЕСІЧНИХ ШЛЯХІВ

A METHOD FOR ORGANIZING MULTIPATH ROUTING USING PARTIALLY-OVERLAPPING PATHS

В роботі запропоновано модифікований хвильовий алгоритм пошуку шляхів, котрий дозволяє покращити пропускну здатність та забезпечити більш оптимальне балансування навантаження мережі. В основі даного алгоритму лежить пошук та використання частко-пересічних маршрутів передачі даних.

Ключові слова: багатошляхова маршрутизація, хвильовий алгоритм, частково-пересічні маршрути

In this work we propose a modified wave algorithm for finding paths, which can improve the bandwidth and provide a more optimal load balancing of network. The basis of this algorithm is the search and use of partially overlapping routes.

Key words: multipath routing, wave algorithm, partially-overlapping path

Актуальність теми дослідження. Розширення сфери використання комп'ютерних мереж та підвищення їх навантаженості вимагає більш високі вимоги до якості обслуговування та використанню доступних ресурсів. Дана робота присвячена способу організації багатошляхової маршрутизації, як одному із способів покращення роботи мережі.

Постановка проблеми. Покращення існуючих методів багатошляхової маршрутизації за допомогою використання частково-пересічних маршрутів передачі даних.

Аналіз останніх досліджень і публікацій. Сучасні способи формування множини шляхів передачі даних основані на алгоритмах пошуку у глибину, пошуку у ширину або комбінаторних методах. Більшість таких алгоритмів маршрутизації [1, 2] розроблені для пошуку непересічних маршрутів, котрі в свою чергу не можуть повністю використовувати всі ресурси мережі. Так у роботі [3] розглянуто хвильовий алгоритм маршрутизації, котрий адаптований під пошук множини непересічних маршрутів. У роботі [4] також розглянуто модифікацію методу гілок та меж адаптованого під багатошляхову маршрутизацію.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена способу організації багатошляхової маршрутизації за допомогою використання частково-пересічних маршрутів. Дослідження сфокусовано на алгоритмі маршрутизації котрий може сформувати саме частково-пересічні маршрути.

Постановка завдання. Завданням є розробити алгоритм маршрутизації, котрий буде знаходити множину частково-пересічних шляхів передачі даних всередині мережі і буде оптимальним по часовій складності.

Викладення основного матеріалу. Знаходження всіх частково-пересічних маршрутів у графі не є оптимальним варіантом, хоча і можливе, так як це

фактично пошук абсолютно усіх можливих шляхів між двома вузлами. В загалом, таке рішення можливо застосувати, лише у спробі максимально розподілити та збалансувати трафік всередині мережі. Однак це спричинить дуже великі затрати на обрахування шляхів, затримку передачі даних та деякі інші критерії QoS.

Щоб обрати більш оптимальні маршрути та зменшити часову складність їх знаходження в даній роботі представлено модифікацію хвильового алгоритму. Особливістю даного рішення є можливість сформувати множину частково-пересічних маршрутів з обмеженням максимальної довжини шляху.

Пошук частково-пересічних маршрутів на основі класичного хвильового алгоритму має декілька фундаментальних проблем. По-перше – в процесі пошуку, велика кількість маршрутів в результаті утворить цикли. Це спричинено тим, що для знаходження шляхів з перетинами необхідне повторне використання певних вузлів та ліній зв'язку.

Для вирішення даної проблеми в розробленому алгоритмі після першого використання лінії зв'язку від вершини A до вершини B, відбувається видалення зворотного зв'язку B → A.

Також відбувається видалення зв'язків між суміжними вузлами всередині одної хвилі. Що також дозволяє оптимізувати роботу алгоритму.

Для зменшення загальної затримки передачі даних в алгоритм закладено механізм автоматичного відсічення дуже довгих шляхів. Максимальна довжина маршруту, котрий може знайти алгоритм вираховується за формулою:

$$S_{max} = S_{min} + K, \quad (1)$$

де: S_{max} – максимальна довжина шляху, S_{min} – довжина найкоротшого шляху, K – коефіцієнт подовження мінімального шляху.

Формула для знаходження K:

$$K =]S_{min}/2 + 1[\quad (2)$$

При використанні формул (1) та (2) найдовший шлях може бути лише у півтора рази довшим за найкоротший.

Реалізація даного механізму обмеження окрім зменшення затримки передачі даних дозволяє алгоритму зменшити час його виконання. Це викликано тим, що на кожному кроці алгоритм запам'ятовує кількість вершин, котрі належать кожному шляху. В результаті після знаходження найкоротшого шляху, дані про його довжину вносяться до глобальної змінної і після досягнення іншими шляхами значення максимальної довжини заданого формулою (1), алгоритм просто відкидає їх і зупиняє свою роботу.

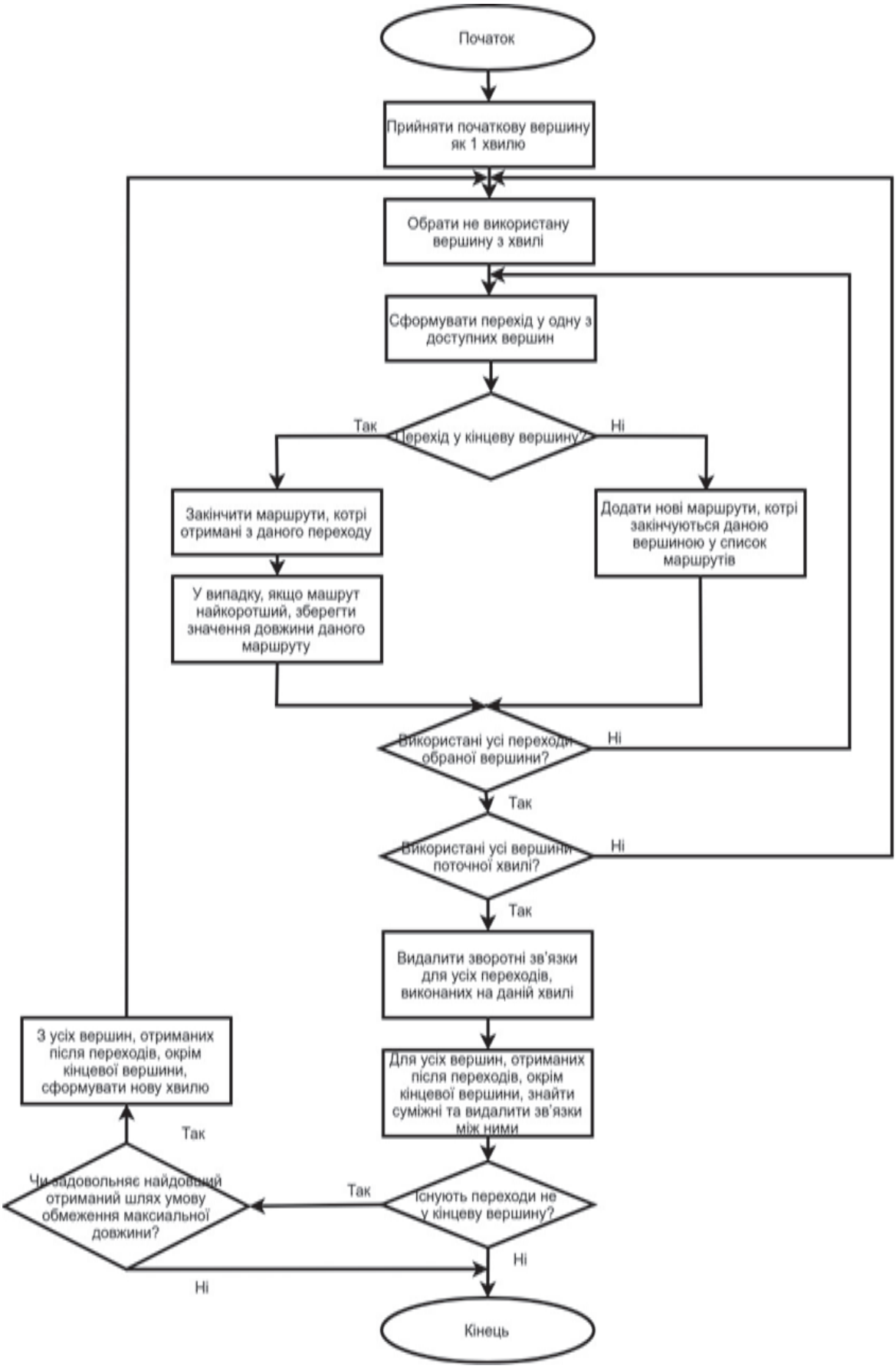


Рис. 1. Блок-схема розробленого алгоритму

Експерименти. Для проведення експериментів було розроблено моделюючу програму котра порівнює кількість елементарних операцій, виконаних розробленим алгоритмом у порівнянні з алгоритмом Дійкстри, модифікованим під пошук частково-пересічних шляхів. Для порівняння було виконано 100 запусків алгоритму на випадково згенерованих графах заданої розмірності. На рис. 6 зображено результати роботи.

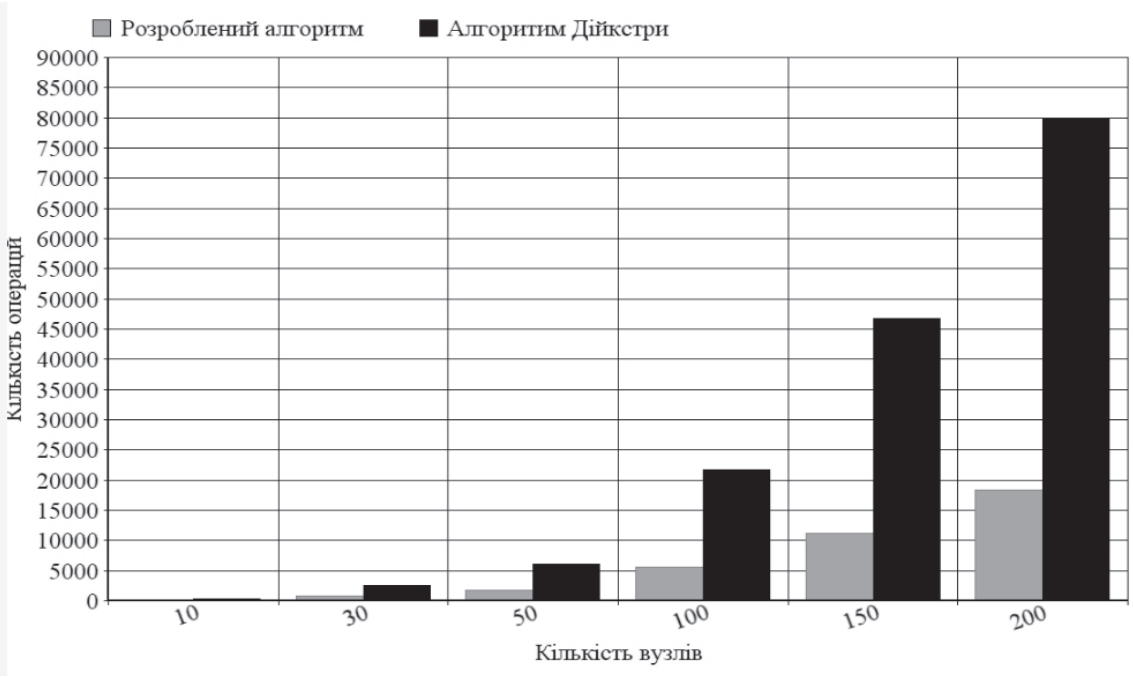


Рис. 2. Залежність кількості виконаних операцій від розмірності мережі

Висновок. Представлений у роботі алгоритм характеризується меншою часовою складністю у порівнянні з класичним алгоритмом, котрий використовується у сучасних мережах. Використання розробленого алгоритму дозволить значно покращити використання доступних ресурсів мережі при збереженні швидкості її роботи.

Список літератури

1. Lee, S. J., &Gerla, M. (2001). Split multipath routing with maximally disjoint paths in ad hoc networks. In Communications, 2001. ICC 2001. IEEE International Conference on (Vol. 10, pp. 3201-3205). IEEE.

2. Marina, M. K., & Das, S. R. (2001, November). On-demand multipath distance vector routing in ad hoc networks. In Network Protocols, 2001. Ninth International Conference on (pp. 14-23). IEEE.

3. Диброва, М. А., Коган, А. В., & Капорин, Р. М. (2016). Способы организации многопутевой маршрутизации.Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: збірник наукових праць. – Київ: Век+, – 2014. №64С. 21-26.

4. Кулаков, Ю. А., Коган, А. В., & Морозовский, Т. О. (2015). Способ организации многопутевой маршрутизации с помощью модифицированного метода ветвей и границ.Вісник Національного технічного університету України КПІ. Інформатика, управління та обчислювальна техніка, (№62), 27-31.

ДОВІДКА ПРО АВТОРІВ

Калюжний Олександр Олегович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kaliuzhnyi Oleksandr – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: sleeper253@gmail.com

Кулаков Юрій Олексійович – професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kulakov Yurii – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: ya.kulakov@gmail.com

Діброва Михайло Олександрович– аспірант, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Mykhailo Dibrova– PHD student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

O. Kaliuzhnyi, Y. Kulakov, M. Dibrova

A METHOD FOR ORGANIZING MULTIPATH ROUTING USING PARTIALLY-OVERLAPPING PATHS

Relevance of research topic. Expanding the scope of computer networks and increasing their load requires higher requirements for the quality of service and the using of available resources. This work is devoted to the method of organization of multipath routing as one of the ways to improve network performance.

Target setting. Improving existing multipath routing methods using partially-overlapping data paths.

Actual scientific researches and issues analysis. Modern methods of generating a set of paths are based on algorithms like search in depth, search in width, or combinatorial methods. Most such routing algorithms are designed to find non-overlapping routes, which in turn can not fully utilize all network resources.

Uninvestigated parts of general matters defining. This article is devoted to the method of organization of multipath routing using partial-overlapping routes. The research focuses on the routing algorithm, which can find partial-overlapping routes.

The research objective. The task is to develop a routing algorithm that will find a plurality of partially-overlapping paths in the network and will be optimal in time complexity.

The statement of basic materials. An analysis of existing multipath routing algorithms has been carried out. A modification of the wave algorithm is described, for the possibility of finding partial-overlapping paths. A block diagram of the developed algorithm is presented.

Conclusions. The algorithm presented in the work is characterized by less temporal complexity in comparison with the classical algorithm used in modern networks. Using the developed algorithm will significantly improve the use of available network resources while maintaining the speed of its work.

Key words: multipath routing, wave algorithm, partially-overlapping paths

УДК 004.514

Бугай Андрій, ОлійникВолодимир

МОДЕЛЬ ПОБУДОВИ ІНТЕРФЕЙСУ КОРИСТУВАЧА З ВИКОРИСТАННЯМ МІКРОСЕРВІСНОЇ АРХІТЕКТУРИ

MODEL OF USER INTERFACE DEVELOPEMNT USING MICROSERVICE ARCITECTURE

У статті пропонується модель побудови інтерфейсу користувача з використанням мікросервісної архітектури. Запропонована модель дозволяє об'єднати декілька невеликих інтерфейсів користувача у вигляді одного більшого інтерфейсу для використання у веб-застосунках.

Ключові слова: мікросервісна архітектура, мікросервіс, інтерфейс користувача, iframe

Рис.: 1.Бібл.: 2.

The article proposes a model for constructing a user interface using the microservice architecture. The proposed model allows aggregating several small user interfaces into one larger interface which is usefulfor modern web applications.

Key words: microservice architecture, microservice, user interface, iframe

Fig.: 1.Bibl.: 2.

Актуальність теми. У зв'язку з поширенням Cloud технологій для розробки веб-застосунків набула популярності мікросервісна архітектура [1]. Вона дозволяє створювати розподілені, здатні до масштабування застосунки, що складаються з окремих модулів розроблених з використанням різних технологій різними командами. Такий підхід надає переваги при розробці, тестуванні, розгортанні та впроваджені програмного забезпечення. Хоча такий підхід наразі здебільшого використовується для створення серверної частини веб-застосунків, аналогічні підходи можуть бути застосовані і для побудови інтерфейсної частини застосунків для задоволення вимог, що висуваються до сучасних інтерфейсів. Ці вимоги включають: використання в одному застосунку компонентів інтерфейсу створених на різних технологіях, агрегацію декількох існуючих інтерфейсів в рамках одного застосунку та інші. Дана робота присвячена створенню підходу для розробки інтерфейсів користувача з використанням принципів мікросервісної архітектури.

Аналіз існуючих рішень. Наразі найбільш розповсюдженим способом для об'єднання декількох мікросервісів інтерфейсу користувача в рамках однієї веб-сторінки є використання технології iframe [2]. Такий підхід має ряд недоліків таких як: фіксовані границі блоку iframe, ускладнена комунікація між частинами інтерфейсу, що знаходяться в різних блоках iframe, кожнийблок iframe має окремий контекст, що унеможливлює спільне використання будь яких даних між частинами інтерфейсу з різних блоків iframe.

Постановка задачі. Задачею поставлено запропонувати модель веб-застосунку, що реалізує принципи мікросервісного підходу для побудови інтерфейсу користувача. Дана модель має надавати можливості для агрегації розподілених інтерфейсів в єдиний веб-застосунок та дозволити уникнути недоліків описаного підходу з використанням технології iframe.

Викладення основного матеріалу. Основна складова веб-застосунку, що потребує використання мікросервісного підходу – це множина невеликих прикладних веб-застосунків з інтерфейсами користувача (мікросервісів), які необхідно об'єднати в рамках одного великого веб-застосунку. Ці інтерфейси користувача розташовані за різними URL-адресами та можуть бути розроблені з використання різних технологій. Задачею описаних нижче модулів є надання можливості для агрегації цих мікросервісів в один застосунок.

Окрім мікросервісів інтерфейсу користувача дана модель складається з модуля агрегатора мікросервісів, модуля маршрутизації і модуля розмітки сторінки, що зображено на рис. 1.

Агрегатор мікросервісів дозволяє створити єдину веб-сторінку з декількох мікросервісів. Отримавши відповідні URL-адреси мікросервісів, даний модуль отримує весь необхідний контент цієї сторінки, включаючи HTML, CSS, Javascript, зображення, іконки, шрифти та інші можливі данні і поміщає ці компоненти інтерфейсу користувача на єдину сторінку.

Модуль маршрутизації реалізує механізм знаходження відповідності між назвами мікросервісів та їх URL-адресами. Оскільки в Cloud системах URL-адреси мікросервісів можуть змінюватись, доцільно мати механізм, який відокремлює логіку взаємодії модулів (мікросервісів) даної моделі від фактичних URL-адрес, на яких розгорнуті дані модулі. Цей модуль може приймати запити і по імені мікросервісу, до якого направлений запит, перенаправляти його на відповідну адресу.

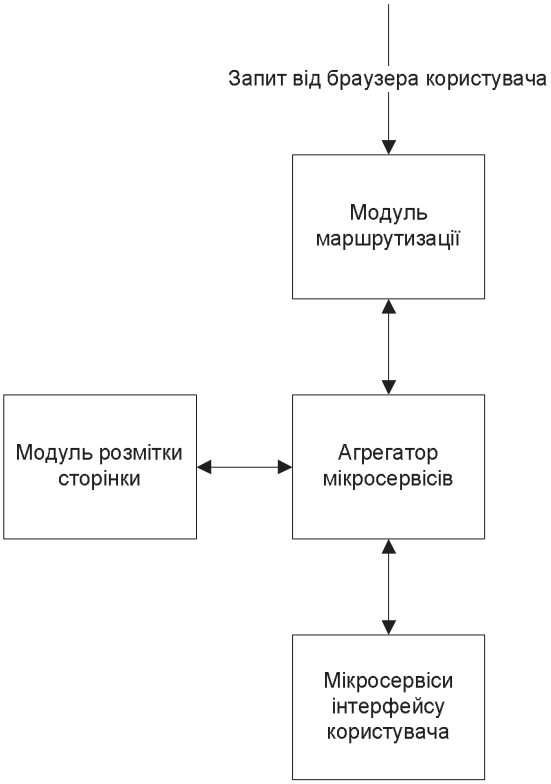


Рис. 1. Модель мікросервісної архітектури для побудови інтерфейсів користувача

Модуль розмітки сторінки – це сервіс, що взаємодії з агрегатором мікросервісів і дозволяє динамічно створювати розмітку сторінку для заповнення її інтерфейсами користувача відповідних мікросервісів.

В цілому, дане рішення працює наступним чином. Користувач звертається до застосунка, модуль маршрутизації знаходить мікросервіси, які необхідно відобразити користувачу та інші інфраструктурні модулі, що зображено на рис. 1. Агрегатор мікросервісів компонує сторінку з розміткою, що надає модуль розмітки сторінки, із відповідних мікросервісів. Отримана сторінка потрапляє для перегляду в браузер користувача.

Висновки. В роботі запропоновано підхід до вирішення задачі об'єднання інтерфейсів користувача у вигляді одного веб-застосунка. Запропонований підхід дозволяє реалізувати принципи мікросервісного підходу для побудови інтерфейсу користувача при цьому уникнути недоліків використання технології iframe.Це дозволяє розроблювати інтерфейси користувача розбиваючи його на невеликі модулі, отримуючи переваги мікросервісного підходу.

Список використаних джерел

1. Fowler M. Microservices - a definition of this new architectural term [Електронний ресурс] // martinfowler.com. 2014. URL: https://martinfowler.com/articles/microservices.html. Дата звернення: 28.3.2018

2. Pool J. The Inline Frame element [Електронний ресурс] // developer.mozilla.org. 2018. URL: https://developer.mozilla.org/en-US/docs/Web/HTML/Element/iframe. Дата звернення: 30.04.2018

ДОВІДКА ПРО АВТОРІВ

Бугай Андрій Андрійович – бакалавр, студент кафедри технічної кібернетики НТУУ "КПІ".

Bugai Andrii – bachelor, student, Department of Technical Cybernetics NTUU "KPI".

E-mail: andriibugai@yahoo.com

Олійник Володимир Валентинович – к.т.н., старший викладач кафедри технічної кібернетики, НТУУ «КПІ ім. Ігоря Сікорського»

Oliinyk Volodymyr Valentynovych – Ph.D., senior lecturer at Technical Cybernetics Department, NTUU "Igor Sikorsky Kyiv Polytechnic Institute"

E-mail: oliinyk.volodymyr@gmail.com

Andrii Bugai,
Oliinyk Volodymyr

MODEL OF USER INTERFACE DEVELOPMENT USING MICROSERVICE ARCHITECTURE

Target setting. Due to cloud technologies are becoming wide spread for the of web applications development, the microservice architecture is becoming popular as well. This approach has a number of benefits and is mainly used to develop server side of web applications. As soon as requirements for modern user interfaces are becoming more difficult, the same microservice approach is may be used to solve user interface issues. This work is devoted to the development of an approach for developing user interfaces using the principles of microservice architecture.

Analysis of existing solutions. Currently, the most common way to merge multiple user interfaces within a single web page is to use the iframe technology. This approach has a number of drawbacks.

Formulation of the problem. The task is to propose a web application model that implements the principles of a microservice approach for building a user interface that will allow to aggregate distributed interfaces into a single web application.

Main results. The model consists of a few main modules including a set of small applications with user interfaces (microservices) that need to be aggregated into one large web application, aggregator of microservices, router and template modules. The entire solution allows to create a single web page from several microservices. The routing module implements a mechanism for finding the correspondence between names of the microservices and their URLs. The template module allows to dynamicalcreation of a markup page to fill its user interfaces with the corresponding microservices.

Conclusions The paper proposes an approach for solving the problem of combining user interfaces in a single web application. The proposed approach allows implementing the principles of the microservice approach for building a user interface avoiding the disadvantages of using the iframe technology.

UDC 004.7

Кулаков Юрій,
Шудренко Євгеній,
Баринін Юрій

МОДИФІКАЦІЯ АЛГОРИТМУ ЕНЕРГОЗБЕРІГАЮЧОГО УПРАВЛІННЯ ТРАФІКОМ В ГІБРИДНИХ МАГІСТРАЛЬНИХ SDN/IP МЕРЕЖАХ

MODIFICATION OF ENERGY SAVING TRAFFIC MANAGEMENT ALGORITHM IN HYBRID MAINLINE SDN/IP NETWORKS

Дана стаття присвячена вивченню та модифікації алгоритмаHEATHE. Покращення сфокусовано на вихідному алгоритмі що представлений графом мережі за наявності функції пошуку сусідів по регіону та функції розподілу потоку.

Ключові слова: SDN, HEATHE, модифікація алгоритму, енергозбереження, трафік.

Рис.: 3. Табл.: 1. Бібл.: 13.

This article is devoted to the study and modification of the HEATHE algorithm. Improvement is focused on the output algorithm represented by the graph of the network in the presence of the search function of neighbors in the region and the distribution function of the flow.

Key words: SDN, HEATHE, modification of the algorithm, energy saving, traffic.

Fig.: 3. Tabl.: 1. Bibl.: 13.

Актуальність теми дослідження. У зв'язку зі стрімким поширенням хмарних технологій в приватний та бізнес сферах і, відповідно, високими темпами зростання мережевого трафіку, все більше мереж потребують централізованих, програмованих засобів управління для ефективного керування трафіком. Враховуючи також сектор мобільних пристроїв, які стали регулярно використовуватися для доступу до хмарних сховищ, постають проблеми ускладнення мереж та необхідності гнучкішого їх регулювання.

Постановка проблеми. В даній роботі буде розглянуто проблему енергозберігаючого керування трафіком в гібридних SDN/IP мережах, існуючі шляхи її вирішення та можливості щодо їх вдосконалення та прискорення. Загалом доведено, що поставлена задача є NP-повною при одночасній зміні як вагових коефіцієнтів так і коефіцієнтів розділення трафіку.

Аналіз останніх досліджень і публікацій.Одним із запропонованих евристичних алгоритмів є HEATHE [5], що базується на ідеї фіксації одного з параметрів, а саме оптимізації вагових коефіцієнтів OSPF каналів зв'язку, що дає можливість обрахувати коефіцієнт розділення трафіку за поліноміальний час. Основна мета алгоритму полягає у зниженні утилізації ліній передачі, що найменше завантажені з метою їх остаточного відключення і, як наслідок, економії енергії. При цьому здійснюється перевірка завантаженості залишкових каналів, беручи до уваги коефіцієнт максимально допустимої завантаженості β, на основі якого приймається рішення про статус чергової ланки зв'язку.

Виділення недосліджених частин загальної проблеми.Вихідний алгоритм [5] може бути представлений наступним чином:

Вхідні дані: граф мережі G(N, L), де N – набір вузлів, L – набір направлених фізичних ліній зв'язку; максимально допустима утилізація β та передача трафіку h_{vt}, де $v-$ початковий вузол, а t – вузол призначення; початкова вага W ліній зв'язку.

1. Ініціалізувати набір вихідних зв'язків S вхідним набором L: $S \leftarrow L$.
2. Запуск функції *Пошуку сусідів по регіону* для оновлення вагових коефіцієнтів w_l
3. Для всіх вузлів $t \in N$ побудувати дерево найкоротших шляхів до всіх інших вузлів, знайти перший SDN-комутаторu на шляху і розрахувати для нього вхідний потік I_{ut}.

3.1. Запуск функції *Розподілу потоку*I_{ut}для SDN-комутатора u.

4. Видалити з набору активних зв'язок з найменшою утилізацією: $S = S - L$
5. Оновити утилізацію залишкових зв'язків, якщо жодна не перевищує граничного значення β, повернутися до пункту 2, інакше закінчити роботу алгоритму і повернути S.

Результат: набір активних лінійзв'язку S.

Функція пошуку сусідів по регіону:

Вхідні дані: Початковий набір вагових коефіцієнтів w_l, максимально можлива утилізація β.

1. Для заданої кількості ітераційi розрахувати:

1.1. Найкоротше дерево шляхів в мережі на основі вихідних коефіцієнтів w_l;

1.2. Утилізацію ліній зв'язку $u_l \leftarrow x_l/c_l$;

1.2.1. Якщо $u_l < 0.3 \cdot \beta$, або $u_l \geq \beta$, збільшити вагу лінії $w_l = = w_l + 1/i$.

Результат: Новий набір вагових коефіцієнтів w_l.

Функція розподілу потоку:

Вхідні дані: Найкоротший шлях від інших вузлів до вузла t, трафік x_l впродовж цього шляху, пропускна здатність c_lусіх ліній зв'язку цього шляху та перший SDN-сумісний комутатор u на ньому, вхідний потік трафіку I_{ut}, гранична утилізація лінії β.

1. Для першого SDN-комутатора u, якщо є kвихідних шляхів, обрахувати $k_{\text{і}}$ найкоротший шлях і зберегти його до набору P.
2. Обрахувати пропускну здатність усіх шляхів, що $\in P$.
3. Сортувати набір шляхів Pза зростанням пропускної здатності.
4. Розділити вхідний потік I_{ut} по лініям зв'язку з набору P у порядку від найменшої до найбільшої ємності, до тих пір, поки $I_{ut} > 0$.

Результат: Розподіл вхідного трафіку по kвихідним лініям зв'язку.

Постановка завдання. Нами запропонована модифікація цього алгоритму, що полягає у мінімізації обчислювальної складності алгоритму в цілому, та, зокрема, функції *Пошуку сусідів по регіону* за рахунок використання додаткової змінної в таблицях маршрутизації вузлів, що підтримують SDN. Ця змінна міститиме відомості про агреговану вагу шляху до пункту призначення пакетів, для кожного з доступних шляхів, динамічно оновлюючись на кожній

ітерації зміни ваг шляхів. Маючи таблицю усіх доступних шляхів від початкового вузла v до термінального вузла t, запропонована змінна дозволить легко та швидко оновлювати вартість чергового шляху у процесі функціонування алгоритму *Пошук сусідів по регіону*. На основі розширеної таблиці з сумарними вагами шляхів, SDN-комутатору v достатньо вибрати найкоротший шлях з таблиці, для доставки пакетів до вузла t. Таким чином виключається необхідність розрахунків дерева можливих шляхів на кожній ітерації циклу функції *Пошуку сусідів по регіону* , що суттєво зменшує кількість обчислень та дозволяє прискорити роботу алгоритму HEATHE в цілому. Останнє твердження очевидне, враховуючи різницю в часі пошуку значення в таблиці шляхів, та побудови чергового дерева шляхів за використання алгоритму Дейкстри.

Викладення основного матеріалу.Модифікована функція пошуку сусідів по регіону може бути представлена у вигляді алгоритму:

Вхідні дані: Початковий набір вагових коефіцієнтів w_l, максимально можлива утилізація β.

1. Розрахувати дерево найкоротших шляхів в мережі на основі початкових вагових коефіцієнтів w.

2. Для заданої кількості ітераційi розрахувати:

2.1. Утилізацію ліній зв'язку $u_l \leftarrow x_l/c_l$:

2.1.1. Якщо $u_l < 0.3 \cdot \beta$, або $u_l \geq \beta$, збільшити вагу лінії $w_l = w_l + 1/i$.

2.1.2. Оновити значення сумарної ваги шляху в таблицях маршрутизації.

Результат: Ініціалізовані остаточні значення ваг шляхів в таблицях маршрутизації.

За рахунок зменшення кількості обчислень найкоротших шляхів мережі в $i-1$ разів, обчислювальна складність функції *Пошуку сусідів по регіону* знижується в $\sim(i-1)^2$ разів, так як розрахункова складність побудови дерева найкоротших шляхів - $O(n^2)$, де, в даному випадку, $n-$ кількість вершин вхідного графу мережі.

Моделювання за допомогою програмних засобів показало прискорення виконання модифікованого алгоритму HEATHEна $20-30\%$в залежності від кількості вузлів у мережі. Графік залежності прискорення від розмірності вхідного графу показано на рис. 1. Окрім цього проведено також порівняльний аналіз часу виконання вихідного алгоритму та модифікованого варіанту, як відображено на рис. 2.

При цьому варто відмітити, що зберігання додаткового значення довжини шляху у форматі одинарної точності з плаваючою комою потребуватиме ~4 байти\комірку, що збільшить загальні витрати пам'яті на $\sim 2/n^2$, де n – кількість вузлів мережі. Дана залежність відображена на рис. 3, що є результатом прикладного тестування витрат пам'яті моделюючою програмою для різної величини мережі. Основою для побудови залежності є відношення витрат пам'яті за використання вихідного алгоритму HEATHEта модифікованої версії. Як видно з малюнка, при збільшенні кількості вузлів мережі, об'єм пам'яті, що виділяється під комірки для зберігання змінної-ваги шляху, є незначною величиною, котрою цілком можна знехтувати.

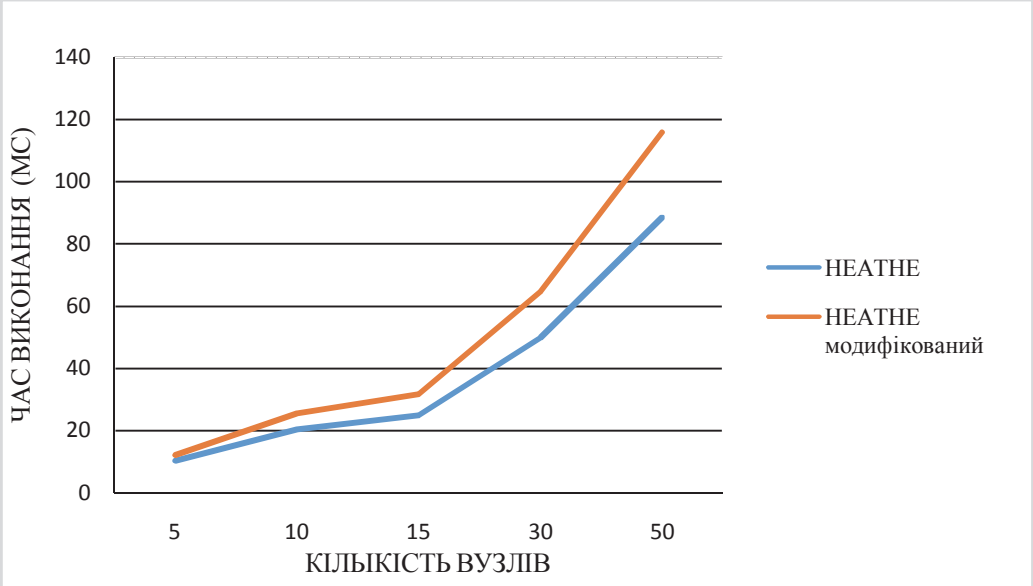


Рис. 1 Порівняння часу виконання обох версій алгоритму

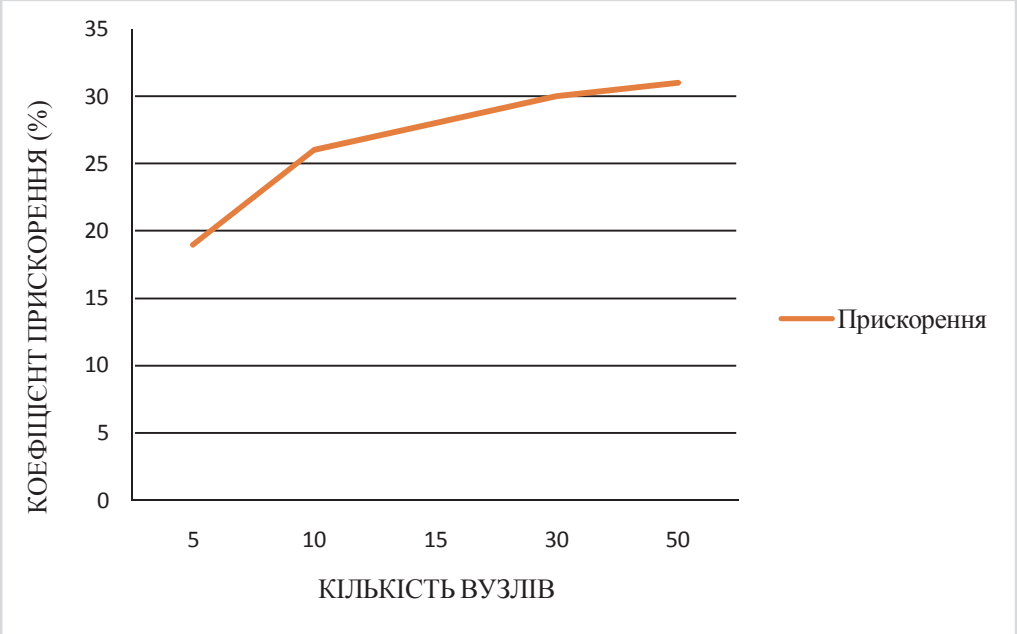


Рис. 2. Залежність прискорення алгоритму від величини мережі

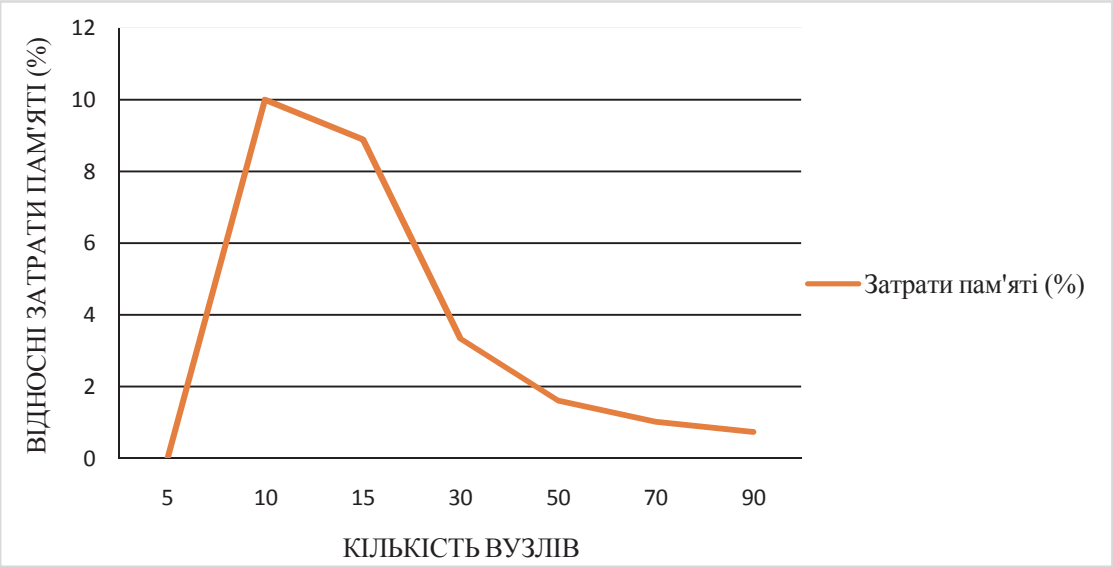


Рис. 3. Залежність зростання витрат пам’яті для модифікованого алгоритму

Висновки. В результаті проведених досліджень можна стверджувати, що застосування модифікованого алгоритму HEATHE з використанням змінної для зберігання сумарної ваги шляху в таблиці маршрутизації дає перевагу над вихідним алгоритмом в швидкодії, вимагаючи лише незначного об’єму пам’яті для зберігання даних та не ускладнюючи при цьому загальну концепцію алгоритму і не збільшуючи його обчислювальну складність. Враховуючи евристичний характер алгоритму HEATHE, можна також стверджувати про доцільність та ефективність застосування запропонованої модифікації алгоритму.

Список використаних джерел

1. Kamal B. Software-defined networking (SDN): a survey [Електронний ресурс] / B. Kamal, E. Abdeslam, E. E. Abdelbaki // John Wiley & Sons Ltd. – 2017. – Режим доступу до ресурсу: https://onlinelibrary.wiley.com/doi/full/10.1002/sec.1737.

2. Salsano S. Hybrid IP/SDN networking: open implementation and experiment management tools [Електронний ресурс] / S. Salsano, P. Ventre // IEEE Transaction of Network and Service Management. – 2015. – Режим доступу до ресурсу: https://pdfs.semanticscholar.org/1500/d5048d67b2095ff347439cc03a47c617babf.pdf.

3. Fernández-Fernández A. Energy Efficiency and Network Performance: A Reality Check in SDN-Based 5G Systems [Електронний ресурс] / A. Fernández-Fernández, C. Cervelló-Pastor, L. Ochoa-Aday // XIII Jornadas de Ingeniería Telemática (JITEL 2017). – 2017. – Режим доступу до ресурсу: www.mdpi.com/1996-1073/10/12/2132/pdf.

4. Design and implementation of SDN/IP hybrid space information network prototype [Електронний ресурс] / H.Lei, Z. Xiaoning, C. Zijing, J. Yajie // IEEE. – 2016. – Режим доступу до ресурсу: https://ieeexplore.ieee.org/document/7586705/.

5. Yunkai W. Energy Aware Traffic Engineering in Hybrid SDN/IP Backbone Networks / W. Yunkai, X. Lei, L. Supeng. // JOURNAL OF COMMUNICATIONS AND NETWORKS, VOL X. – 2016.

ДОВІДКА ПРО АВТОРІВ

Кулаков Юрій Олексійович – доктор технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

KulakovYurii – Doctor of Technical Sciences, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Politechnic Institute”.

E-mail: ya.kulakov@gmail.com

Шудренко Евгеній Олегович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

ShudrenkoYevhenii – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Politechnic Institute”. E-mail: kuoreal@gmail.com

Баринін Юрій Ігорович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Barynin Yurii – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Politechnic Institute”. E-mail:

KulakovYurii,
ShudrenkoYevhenii,
BaryninYurii

MODIFICATION OF ENERGY SAVING TRAFFIC MANAGEMENT ALGORITHM IN HYBRID MAINLINE SDN/IP NETWORKS

Relevance of research topic. Due to the rapid spreading of cloud-based technologies in the private and business sectors and, consequently, the high growth rates of network traffic, more and more networks require centralized, programmable management tools to effectively manage traffic. Taking into account also the sector of mobile devices that have become regularly used for access to cloud storage, there are problems of network complication and the need for more flexible regulation.

Formulation of the problem. Absence of determined, optimum solution to the problem of energy-saving traffic management in hybrid SDN/IP networks, numerous ways of available algorithms' improvement.

Analysis of recent research and publications. One of the proposed solutions is HEATHE, a heuristic algorithm, which provides an opportunity to calculate the traffic separation factor in polynomial time.

Selection of unexplored parts of the general problem. The original HEATHE is a heuristic algorithm, which gives an opportunity to consider ways to optimize it in different ways and directions.

Target setting. We propose a modification of this algorithm, which consists in minimizing the computational complexity of the algorithm in general, and the function of *searching neighbors in the region* in particular, by using an additional variable in the node routing tables, that are supporting the SDN.

The statement of basic materials. By excluding repetitive calculations of the shortest path tree from Neighboring Region Search Function, we significantly reduce amount of computations performed and thereby speed up the whole algorithm. Latter can be proven by software simulation, which has showed performance acceleration of the modified HEATHE algorithm by 20-30% depending on the network size.

Conclusion. As a result of the research, it can be argued that the application of the modified HEATHE algorithm using the variable for storing the total weight of the path in the routing table gives an advantage over the original algorithm in speed, requiring only a small amount of memory for data storage while decreasing its computational complexity. Although various tests and comparison have been taken to highlight the benefits of improved version, considering the heuristic nature of the HEATHE algorithm, one can also argue the expediency and effectiveness of the proposed modification of the original concept.

УДК 004.75

Ю. М. Виноградов,
М. О. Байдиков

ПИТАННЯ ЯКОСТІ МОБІЛЬНИХ ДОДАТКІВ НА ОСНОВІ ЇХ ТЕСТУВАННЯ

MOBILE APPLICATIONS QUALITY ISSUES BASED ON TESTING

У статті розглядається питання підвищення якості мобільних додатків. Для цього використовується повноцінне тестування, яке включає в себе тестування окремих модулів додатків, тестування поведінки всього додатку та тестування поведінки додатку під великим навантаженням. Для тестування використовуються підходи Given-When-Then, Behavior Testing, Fuzz testing та Performance Testing.

Ключовіслова: тестування, Given-When-Then, Behavior Tests, Fuzz tests, Performance Tests.

The paper deals with the issues of increasing quality of mobile applications using testing. As main testing types used Given-When-Then, Behavior Testing, Fuzz testing and Performance Testing. All these testing types cover testing separate modules of application, testing the behavior of whole application and testing application under unexpected cases of usage.

Key words: testing, Given-When-Then, Behavior Tests, Fuzz tests, Performance Tests.

Target setting. Due to growing number of mobile applications appeared last time, the question of their quality has become the main thing to think about.

Actual researches and issues analysis. During last time some big companies made this topic one of the most discussed at their conferences, but it’s still isn’t covered enough.

Uninvestigated parts of general matters defining. There is a huge lack of works, describing Fuzz and Performance testing of mobile applications. Moreover, even usual Given-When-Then and Behavior tests remain investigations. And the biggest gap in all these investigations in connections all this different testing types into one stable ecosystem.

The research objective. The purpose of this paper is to investigate mobile applications testing types and their combination. As a solution, the article will focus on creating one single ecosystem of different testing techniques which can show as much as possible bugs of mobile applications. This approach will help all the developers to increase the quality of their applications and reduce the number of application crashes.

The statement of basic materials. In general, testing is a process of writing some code, which runs app modules or the whole app, performs some actions, collects results and checks their equality to the expected result, provided by developers [1]. Testing can be divided into several types. The main testing types, applicable to the most popular mobile operating systems, iOS and Android, are:

1. Unit testing - tests individual parts of code for correctness, using an automated test suite. A good unit test assures the functionality expecting out of the unit is correct. A good unit test is repeatable, fast, readable, independent and comprehensive.

2. Given-When-Then testing - similar to unit testing, but declares strong format of test, which must be inherited to achieve the proper result.

3. UI Testing - takes a flow that the user might follow and ensures it produces the right output. Can be performed both on real devices or emulators and simulators. Given the same state, the UI test always must produce the same output. The input to the test in the user clicks and the output is the screen.

4. Behavior testing - very similar to UI testing, but instead of testing the whole user flow, test small user actions.

5. Fuzz testing - unit tests and UI tests are user to ensure that the expected output happens when the expected input is used. However, users can be unpredictable and can do unexpected things. For such cases comes the fuzz testing. It introduces a random stream of events into app and records result. It's a stress test which simulates random presses on the screen. This is very similar to UI testing, but it has no input and it doesn't check some defined user flow.

6. Performance testing- gathers performance metrics of the app to ensure that application doesn't take up device resources. This metrics include network usage, memory usage, battery drain.

7. Continuous Integration is a tool to collect all the above techniques into one flow. It is a command line tool, which is hosted on external server and observes application repository. When a developer performs commit, this tool takes the application from this commit and performs all the tests. If all of them are success, then the tool pushes all the changes to remote. In other case, it notifies the developer, that something is wrong with the app and must be fixed before pushing [2].

Table 1 illustrates all the testing types and native frameworks, used on iOS and Android platforms for these types of testing [3].

Table 1

Testing types and native frameworks

Testing type	iOS native framework	Android native framework
Unit testing	XCTest	JUnit
Given-When-Then testing	-	-
UI testing	XCUITest, Xcode UI Test Recorder	Espresso, Espresso Test Recorder
Behavior testing	-	-
Fuzz testing	-	-
Performance testing	Instruments (only for collecting data)	Android Monitor (only for collecting data)
Continuous Integration	-	-

As the table shows, native testing frameworks on each platform don't provide developers with all the required tools to performs full mobile application testing. Moreover, the don't provide anything to perform Given-When-Then testing and are not suitable for Behavior Testing [4].

The correct testing frameworks. The correct testing frameworks must include all the required tools to perform all the given testing types. To achieve these requirements, were developed next frameworks.

Table 2

Developed frameworks for testing

Testing type	iOS framework	Android framework
Unit testing	Quick+Nimble	JUnit
Given-When-Then testing	SwityGWT	KotlinGWT
UI testing	XCUITest, Xcode UI Test Recorder	Espresso, Espresso Test Recorder
Behavior testing	Quick+Nimble	KotlinBDD
Fuzz testing	UI AutoMonkey	monkeyrunner
Performance testing	New Relic Mobile	New Relic Mobile
Continuous integration	Jenkins	Jenkins

Experiments. To check developed frameworks, was developed a small notes application. After that, one it's copy was covered by 30% using native frameworks of each platform. The real usage of this app showed ten crashes per week for one thousand users. The result seems to be not bad, but testing app with developed frameworks leads to 50% code coverage and only one crash per week for one thousand users. The difference is big even at this point, but scaling app to ten thousands users leads to one hundred crashes per week for the version, tested by native frameworks, while the version, tested with developed frameworks achieved the result of ten crashes per week.

Conclusion. The paper has demonstrated the ability of proper testing to decrease number of crashes of mobile application up to ten times even on the small audience. The developed frameworks can be used with any application and don't required lots of time to develop all the test types. Using such combination of frameworks together with continuous integration produces qualitative results.

There are still several directions for future work. One is to develop frameworks for UI Testing for both platforms to detect different UI bugs of applications. Another is to increase simplify setup of all these frameworks and continuous integration. These changes will definitely improve the result.

References

1. Daniel Knott (2015). Hands-On Mobile App Testing: A Guide for Mobile Testers and Anyone Involved in the Mobile App Business. Addison-Wesley Professional; 1 edition, May 28, 2015.

2. Paul Blundell, Diego Torres Milano (2015). Android Application Testing. Packt Publishing - ebooks Account, April 30, 2015.

3. Vijay Velu (2016). Mobile Application Penetration Testing. Packt Publishing - ebooks Account, March 18, 2016.

4. Mobile Testing: Complete Guide to Test your Mobile Apps. Update date: February 15, 2018. URL: https://www.guru99.com/testing-mobile-apps.html (application date: April 29, 2018).

ДОВІДКАПРОАВТОРІВ

Виноградов Юрій Миколайович – старший викладач, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Vinogradov Yurii – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

БайдиковМиколаОлександрович - студент, кафедраобчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Baidykov Mykola - student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

Email: nbaidikoff@gmail.com

Ю. М. Виноградов, М. О. Байдиков

ПИТАННЯ ЯКОСТІ МОБІЛЬНИХ ДОДАТКІВ НА ОСНОВІ ЇХ ТЕСТУВАННЯ

Актуальність теми дослідження. Проблема якості мобільних додатків стала актуальною в останні дні у зв’язку з різким збільшення їх кількості та зростанням попиту. Це спричинило появу великої кількості додатків за короткий період часу без належного тестування. Крім того, самі розробники середовищ створення додатків не дають достатньо ресурсів для написання якісних тестів. Дана стаття присвячена створенню оптимального набору ресурсів, які дозволять проводити повноцінне тестування додатків.

Постановка проблеми. Відсутність якісних ресурсів для повноцінного тестування мобільних додатків та застосування автоматизованого тестування.

Аналіз останніх досліджень і публікацій. Протягом останніх років з’являється все більше праць присвячених тестуванню мобільних додатків. Проте всі вони зосереджуються на якомусь одному методі тестування (найчастіше на Юніт тестуванні), у той час як матеріалів про загальне тестування зі застосуванням декількох різних методів майже немає.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена аналізу існуючих рішень та створенню нових для повноцінного тестування мобільних додатків. Дослідження сфокусовано на вивченні осонвних методик тестування, їх переваг та недоліків.

Постановка завдання. Завдянням є створити оптимальний набір інструментів для всеохоплюючого тестування мобільних додатків на двох сучасних платформах, iOS та Android.

Викладення основного матеріалу. Проведено аналіз стандартних засобів тестування для iOS та Android, виділено їх найголовніші недоліки. Описано вимоги до засобів тестування, які повинні забезпечити можливість повноцінного покриття тестами будь-якого додатку. Створено необхідні засоби.

Висновки. Проаналізовано створені засоби тестування на невеликому додатку з різною кількістю користувачів. Наведені результати експериментів та аналіз вище вказаних кроків.

Ключові слова: тестування, Given-When-Then, BehaviorTests, Fuzz tests, PerformanceTests.

УДК 681.3.06

Демчик Валерій,
Корочкін Олександр

ЗАСТОСУВАННЯ ДРІБНОЗЕРНИСТОГО ПАРАЛЕЛІЗМУ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПАРАЛЕЛЬНИХ ТА РОЗПОДІЛЕНИХ ОБЧИСЛЕНЬ

APPLICATION OF FINE-GRAINED PARALLELISM FOR INCREASING THE EFFICIENCY OF PARALLEL AND DISTRIBUTED COMPUTING

В статті наводяться результати досліджень можливості підвищення ефективності обчислень в багатоядерних комп'ютерних системах за рахунок застосування дрібнозернистого паралелізму як окремо, так і в складі середньозернистого паралелізму.

Ключові слова: ядро, потік, паралелізм, зернистість, fork-join.
Рис.: 4. Табл.: 2. Бібл.: 5.

The paper deals with the issues of increasing the efficiency of computations in multi-core computer systems due to the application of fine-grained parallelism, separately or as part of middle-grained parallelism.

Key words: core, thread, parallelism, granularity, fork-join.
Fig.: 4. Tabl.: 2. Bibl.: 5.

Актуальність теми дослідження. На даному етапі розвитку обчислювальної техніки, коли питання жорсткої обмеженості доступних апаратних ресурсів не стоїть настільки гостро, одним з найважливіших показників, який визначає ефективність роботи конкретної обчислювальної системи, є час, який буде витрачено даною системою на рішення нею конкретної задачі. Мінімізація часових затрат на обчислення є важливим питанням, до якого спрямований значний інтерес як наукового товариства, так і бізнесу. Одним з шляхів, який дозволяє досягти значних скорочень часових затрат є організація паралельних та розподілених обчислень. І одними з ключових питань, які вирішуються при цьому, є вибір зернистості паралелізму, оскільки застосування паралелізму тієї чи іншої зернистості може виказати значний вплив на ефективність різноманітних обчислень, та вибір засобів організації паралелізму.

Постановка проблеми. В найпопулярніших засобах організації паралельних та розподілених обчислень, наряду з традиційними механізмами організації потоків, тільки відносно нещодавно з'явились надійні засоби підтримки дрібнозернистого паралелізму. Тому відсутні порівняльні дослідження проблеми вибору паралелізму того чи іншого ступеню зернистості для вирішення конкретних математичних задач, які потребують значних обчислень, а також засобів, які реалізують паралелізм.

Аналіз останніх досліджень і публікацій. Протягом останніх років з'являється все більше статей присвячених паралелізму різного ступеню зернистості. Проте підходи для застосування дрібнозернистого паралелізму все ще недостатньо вивчені.

Виділення недосліджених частин загальної проблеми. Дана стаття присвячена вивченню, порівнянню та аналізу питань ефективності застосування

паралелізму різного ступеню зернистості, організованого різними засобами при застосуванні його в багатоядерних комп'ютерних системах. Також проводиться дослідження ефективності власного запропонованого підходу комбінованого паралелізму для організації паралельних та розподілених обчислень, який ґрунтується на одночасному застосуванні (комбінуванні) середньо- та дрібнозернистого паралелізму.

Постановка завдання. Завданням є розробка пакету програм для матричних операцій з використанням бібліотеки OpenMP, мов C# та Java. Кожну операцію реалізувати у вигляді чотирьох програм П1-П4. В програмі П1 використати традиційний підхід, який базується на програмуванні декількох потоків та низькорівневих примітивів їх взаємодії (середньозернистий паралелізм), в програмі П2 – застосувати лише розпаралелювання обчислювальних циклів (дрібнозернистий паралелізм), в програмі П3 – організувати розпаралелювання на декілька потоків, в кожному з яких провести розпаралелювання обчислювальних циклів (комбінований паралелізм). Програма П4 повинна являти собою послідовну реалізацію, необхідну для розрахунку коефіцієнтів прискорення програм П1-П3. Провести тестування розроблених програм.

Аналіз засобів реалізації. Засоби реалізації дрібнозернистого паралелізму реалізовано в бібліотеці OpenMP та мовах Java і C# [1- 4].

Бібліотека OpenMP. Дрібнозернистий паралелізм представлений спеціальною директивою препроцесору #pragma omp parallel for**,** яка відноситься до директив розділення роботи (work-sharing directive). Такі директиви застосовуються не для паралельного виконання коду, а для логічного розподілу групи потоків, щоб реалізувати вказані конструкції керуючої логіки. Директива #pragma omp for повідомляє, що при виконанні циклу в паралельному режимі ітерації циклу повинні бути розподілені між групою потоків. Взаємодія потоків здійснюється через неявну бар'єрну синхронізацію в кінці кожного блоку #pragma omp for.

Мова Java. Розробникам пропонується надзвичайно гнучкі та потужні засоби реалізації дрібнозернистого паралелізму, в основу яких покладено так механізм «Fork-Join». За фактичну реалізацію цього механізму в Java обрано рекурсивний алгоритм, робота якого полягає в наступному:

1. Здійснюється перевірка можливості розподілу дій даного потоку на дві менші задачі;

2. Якщо перевірка успішна, то виконується розподіл (Fork), шляхом створення нових потоків для кожної нової задачі. В кожному новому потоку алгоритм починається заново. Потік, який виконав розподіл блокується до тих пір, поки обидва створені ним потоки не закінчать свою роботу, а після виконує остаточний збір результату;

3. Якщо перевірка не успішна (досягнуто межі так званого «зерна паралелізму»), то виконуються задані обчислення в цьому потоці, по закінченню яких відбувається з'єднання з породившим потоком (Join).

Таку реалізацію містять два класи: RecursiveAction та RecursiveTask. Головна різниця між ними полягає в тому, що коли необхідно порахувати якесь конкретне числове значення певної великої функції (наприклад суму елементів вектора) то зручно використовувати RecursiveTask, а у випадку загальних операцій, результатом роботи яких не є конкретне число, краще використати RecursiveAction.

Наслідуючи дані класи та описавши власне перевантаження методу compute розробник може налаштувати роботу під вирішення конкретної задачі, що є надзвичайно ефективним.

Мова C#. В мові C# також з'явилась можливість реалізації дрібнозернистого паралелізму. Весь необхідний для цього функціонал вмістив в себе статичний клас System.Threading.Tasks.Parallel, а саме три його основні методи Parallel.For(), Parallel.ForEach(), Parallel.Invoke() та різні їх перевантаження. Parallel.For та Parallel.ForEach забезпечують паралельне виконання циклів for та foreach відповідно. В основу кожного зі згаданих методів покладено механізм, аналогічний тому, який застосовується в мові Java. Але при цьому його значно спрощено шляхом застосування механізму делегування та використанням анонімних методів.

Комбінований паралелізм. Запропонований в роботі підхід базується на використанні в паралельної програмі двох видів паралелізму: середньозернистого і дрібнозернистого. При цьому паралельна програма включає набір традиційних потоків по кількості ядер БКС. Кожен з цих потоків додатково реалізує внутрішній (дрібнозернистий) паралелізм шляхом створення підпотоків за допомогою відповідних засобів типу Fork-Join. Початкова кількість традиційних потоків може зменшуватися з метою забезпечення дрібнозернистого паралелізму вільними процесорними ресурсами.

Результати тестування. Тестування програм проводилось на КС, оснащеною шестиядерним процесором AMD Phenom II та модулем ОЗУ об'ємом 4 ГБ типу DDR3. Програмне забезпечення: операційна система Windows 7, бібліотека OpenMP 3.1, віртульна машина JVM 1.8, .NET Framework 4.7.

В таблиці 1 наведені отримані результати тестування паралельних програм для операції множення матриць для різних значень N (розмірність матриць).

Таблиця 1

Результати тестування паралелізму різного ступеню зернистості

N	Час виконання (сек.)								
	Середня зернистість			Дрібна зернистість			Змішана зернистість		
	OpenMP	Java	C#	OpenMP	Java	C#	OpenMP	Java	C#
516	1.7	0.2	1.2	1.6	0.2	1.1	1.5	0.2	1.1
1032	13.3	3.3	9.8	12.8	2.9	9.8	12.4	3.1	9.5
1548	46.7	14.0	35.6	45.9	12.1	34.1	44.9	13.0	33.8
2064	111.7	38.5	79.8	110.8	33.5	78.1	108.1	35.5	74.8

На рисунках 1-3 зображено графіки, які демонструють виявлену залежність коефіцієнтів прискорення (Кп) від значень N, для всіх трьох рівнів паралелізму.

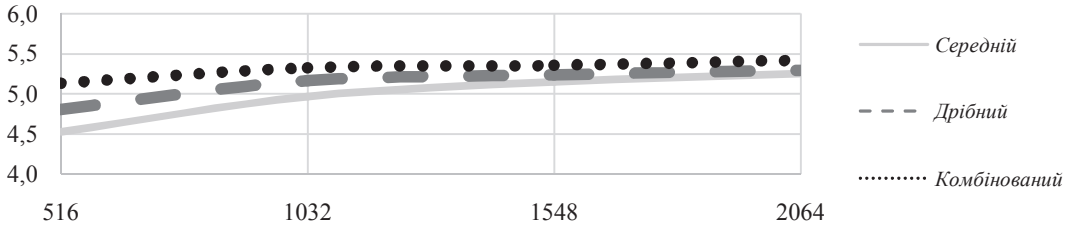


Рис.1. Графік залежності Кп від N. Бібліотека OpenMP

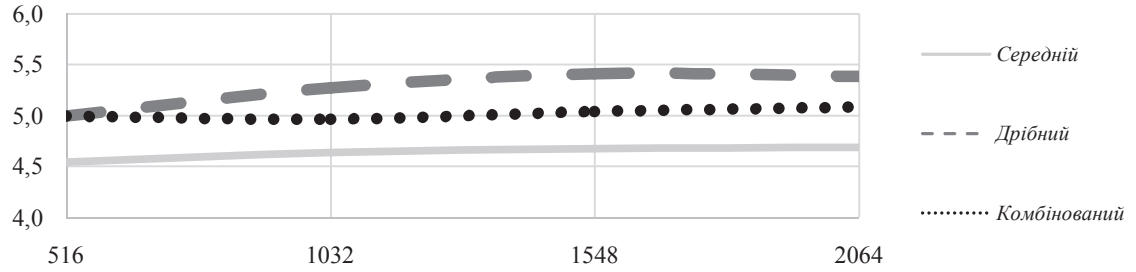


Рис.2. Графік залежності Кп від N. Мова Java

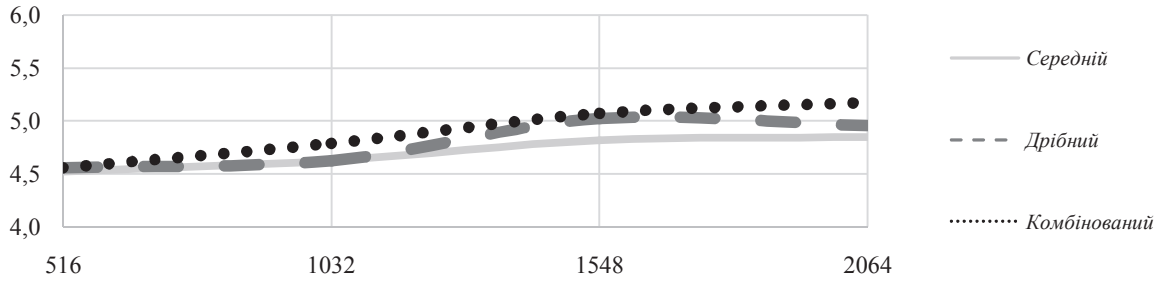


Рис.3. Графік залежності Кп від N. Мова C#

Окрім того, додатково було проведено окреме, більш детальне тестування дрібнозернистого паралелізму, з метою виявлення шляхів підвищення його ефективності. На цьому етапі тестувалися ті програми із розробленого пакету, які були написані з використанням лише дрібнозернистого паралелізму. Проводилися багаторазові заміри часу виконання операції множення двох матриць розмірністю 1500*1500 елементів. Перевірялася залежність часу виконання від кількості програмно-реалізованих потоків. На рисунку 4 наведено графік, який демонструє виявлену залежність часу роботи від кількості потоків (P). Детальні результати наведено в таблиці 2.

Таблиця 2

Результати тестування дрібнозернистого паралелізму

Кількість потоків	**Час виконання (сек.)**		
	C#	**OpenMP**	**Java**
2	71.309	133.944	29.67
3	55.1	101.41	24.12
4	42.099	77.393	16.894
5	37.202	60.942	14.167
6	34.222	47.532	13.101
10	34.113	45.551	12.24
12	32.841	45.117	12.485
15	31.405	44.37	11.195
20	30.404	44.33	9.899
25	30.511	43.101	9.636
30	28.329	41.9	9.211
50	27.683	40.88	9.5

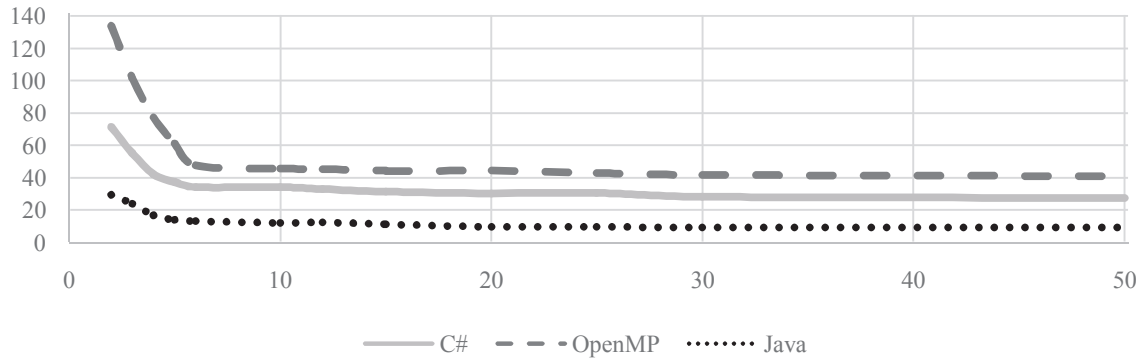


Рис. 4. Дрібнозернистий паралелізм.
Графік залежності часу роботи від кількості потоків

Висновки. Результати тестування показали ефективність багатоядерних комп'ютерних систем при реалізації рішення розглянутої математичної задачі засобами мов JavaiC# та бібліотеки OpenMP. При цьому можливе скорочення часу виконання програм при застосуванні паралелізму будь-якого ступеню зернистості (значення Кп лежать в межах 4.5-5.5). Найкращий результат по часу виконання програми отримано для мови Java.

Середньозернистий паралелізм показав достатню ефективність, але мав найгірший результат. При цьому прискорення, досягнуте засобами OpenMP постійно зростає зі збільшенням обсягу оброблюваних даних, а засобами Java та C# залишається приблизно на одному і тому ж рівні.

Застосування дрібнозернистого паралелізму також виявилося ефективним, в цьому випадку коефіцієнт прискорення стабільно зростає зі збільшенням обсягу оброблюваних даних. Найбільш ефективним цей тип паралелізму виявився в мові Java, завдяки потужному механізму підтримки fork-join.

Окрім того, проведене додаткове тестування дрібнозернистого паралелізму виявило спадаючу експоненціальну залежність між кількістю потоків, які дозволено створити програмі, та часом її роботи. Головною причиною цього є те, що чим більша кількість потоків, тим більше від них звернень до спільних ресурсів, що призводить до значних простоїв внаслідок вирішення задачі взаємного виключення. Оптимальна кількість потоків дрібнозернистого паралелізму, незалежно від того, якими засобами його було організовано, лежить в межах від 5 до 20.

Запропонований в роботі підхід, який ґрунтується на комбінованому паралелізмі, показав свою ефективність і дозволів збільшити Кп при використанні в мові C# та бібліотеці OpenMP. При цьому спостерігається зростання Кп при збільшенні обсягу оброблюваних даних, що є одним з найважливіших аргументів доцільності застосування даного підходу в БКС.

Можна припустити, що ефективність використання комбінованого паралелізму буде збільшуватися при зростанні кількості ядер в БКС, де :

- з'являться додаткові процесорні ресурси для його реалізації;
- можна зменшити розмір дрібної зернистості
- можна знайти оптимальне співвідношення між кількістю потоків і підпотоків.

Крім того, ефективність реалізації комбінованого паралелізму в OpenMP можна покращити шляхом більш ефективної реалізації потужного механізму управління підпотоками, що закладений в бібліотеці, аналогічно тому, як це зроблено в моделі Fork-Join.

Таким чином, можна стверджувати що застосування комбінованого паралелізму в більшості випадків є ефективним підходом до реалізації об'ємних паралельних обчислень на багатоядерних комп'ютерних системах.

Список посилань

1. Lea, Doug. A Java Fork/Join Framework, In Proceedings of ACM Java Grande 2000 Conference (San Francisco, California, June 3-5, 2000)

2. Цилькер Б.Я. Организация ЭВМ и систем: учебник для вузов / С.А. Орлов, Б.Я. Цилькер. - СПб.: Питер, 2011. - 688 с.

3. Жуков І.А., Корочкін О.В. Паралельні та розподілені обчислення: Навч. посібник [Текст]. – К.: Корнійчук, 2005. – 226 с. – ISBN 996-7599-36-1.

4. Канг Су Гэтлин, Пит Айсенси. OpenMP и C++. [Електронний ресурс]. – Режим доступу: https://msdn.microsoft.com/ru-ru/library/dd335940.aspx

5. Julien Ponge.Fork and Join: Java Can Excel at Painless Parallel Programming Too![Електронний ресурс]. — Режим доступу: http://www.oracle.com/technetwork/articles/java/fork-join-422606.html— Загл. з екрану.

ДОВІДКА ПРО АВТОРІВ

Демчик Валерій Валентинович – студент, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Demchyk Valerii –student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: kirintor830@gmail.com

Корочкін Олександр Володимирович – доцент, кандидат технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

KorochkinAleksandr– docent, candidate of Technical Sciences, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: avcora@kpi.ua

Demchyk Valerii, Korochkin Aleksandr

APPLICATION OF FINE-GRAINED PARALLELISM FOR INCREASING THE EFFICIENCY OF PARALLEL AND DISTRIBUTED COMPUTING

Relevance of the research. Solving the issue of organizing efficient parallel and distributed computing is one of the key stages in designing modern software for multi-core computer systems. At the same time, one of the key points in the solution of this issue is the choice of the grains of the applied parallelism, which can have a significant impact on the time of work.

Target setting. The lack of comparative studies on selection of a parallelism degree of granularity to solve specific mathematical problems that require significant computation, and tools that implement parallelism.

Actual scientific researches and issues analysis. In recent years, there are more articles devoted to the parallelism of varying degrees of graininess. However, approaches to apply fine-grained parallelism is still not well understood.

Uninvestigated parts of general matters defining. This article is devoted to the comparison and analysis of the issues of the efficiency of the application of parallelism of different degrees of grains, organized by various means when applied to multicore computer systems. The research of the effectiveness of the proposed proposed approach of combined parallelism is also being conducted.

The research objective. The task is to develop a package of programs for matrix operations using the OpenMP library, the C # languages and Java, to implement each operation in several variants, using a parallelism of varying degrees of grainy. Conduct testing of developed programs.

The statement of basic materials. The comparative characteristic of the main means of organizing parallelism of different degree of grains is carried out. The description of its own proposed type of parallelism is given. Testing of programs for parallel computer systems with application of different granularity of parallelism is carried out.

Conclusions. The application of fine-grained parallelism proved to be the most effective approach to the organization of parallel computing. The approach suggested in the work has shown its efficiency and permissions to increase the acceleration factor. At the same time there is an increase in speedup coefficient with an increase in the volume of processed data, which is one of the most important arguments of the expediency of this approach in multi-core systems.

Key words: core, thread, parallelism, granularity, fork-join.

УДК 004.75

Юрій Кулаков,
Євгеній Лопушен,
Берест Руслана

СПОСІБ МОНІТОРИНГУ ОБЧИСЛЮВАЛЬНИХ РЕСУРСІВ В ПРОГРАМНО-КОНФІГУРОВАНИХ МЕРЕЖАХ

WAY OF COMPUTATIONAL RESOURCES MONITORING IN SOFTWARE-DEFINED NETWORKS

Анотація. В статті розглядається кластер, в якому виконується розподілений додаток. Кластер знаходиться в програмно-конфігурованій мережі (SDN). Для отримання даних про наявні та зайняті ресурси використовуються загальноприйняті інструменти моніторингу серверів та їх REST API для передачі даних в контролер SDN. Контролер SDN отримує дані і на їх основі приймає рішення про зміну конфігурації мережі.

Ключові слова: програмно-конфігуровані мережі, SDN контролер, моніторинг, агент моніторингу, масштабування,

Рис.: 2, Табл.: 0, Бібл.: 5

Annotation. The paper deals with the cluster, in which there is a running distributed application. The cluster is located in the software-defined network (SDN). Widely adopted tools for server monitoring are used for data retrieval about free and occupied resources and their REST API is used for transferring the data to the SDN controller. SDN controller receives the data and makes the decision based on this data about changing the network configuration.

Key words: software-defined networks, SDN controller, monitoring, monitoring agent, scaling.

Fig.:2, Tabl.: 0, Bibl.: 5

Актуальність проблеми. В сучасних системах є актуальною проблема стійкості web-додатків до зміни навантаження, що спричиняється зміною кількості користувацьких запитів. Постійне утримання ресурсів, необхідних для роботи додатку під піковими завантаженнями часто є невиправдано дорогим, оскільки більшість часу кількість запитів є значно меншою. Кількість запитів від користувачів додатку не є сталою, наприклад, в робочі години навантаження значно зростає, що може призвести до затримок в роботі. Водночас, при меншому навантаженні не є доцільним утримання надлишку ресурсів. Тому, доречно динамічно масштабувати додаток, використовуючи лише необхідну кількість вузлів. Це дозволить оптимізувати витрати і вивільнити ресурси під інші потреби.

Огляд існуючих рішень. Контролер SDN має засоби відстеження трафіку, що надходить до кожного з хостів в мережі.Обсяг трафіку часто не є показником завантаження вузла: необхідно мати інформацію про використання ОЗУ, диску, процесора тощо. Підтримка SDN-контролером додаткових плагінів дозволяє налаштувати отримання даних про завантаження безпосередньо з хоста. Існуючі способи моніторингу кластера [1], [2] не підтримують роботу з

програмно-конфігурованими мережами, проте мають можливість інтеграції з контролером через API.

Постановка завдання. Контролер SDN відповідає за маршрутизацію трафіку в програмно-конфігурованій мережі. Проте, для прийняття рішення про маршрутизацію на той чи інший хост, контролеру потрібно мати інформацію щодо його завантаження. Необхідно розробити спосіб моніторингу наявних ресурсів в програмно-конфігуровних мережах SDN, що надасть можливість отримувати актуальні дані про завантаження хостів в мережі. Наявність таких даних дає можливість налаштувати маршрутизацію на контролері таким чином, щоб розподілити навантаження рівномірно і запобігти відмові хоста, а за відсутності потреби в ресурсах – відключити хост від мережі, тим самим звільнивши ресурси під інші потреби.

Викладення основного матеріалу

Засоби реалізації способу моніторингу в SDN. Запропонований спосіб базується на основних перевагах програмно-конфігурованих мереж. По-перше, можливість реконфігурації мережі шляхом редагування таблиці маршрутизації через API контролера. По-друге, можливість розширення функціоналу контролера з допомогою плагінів. Контролер програмно конфігурованої мережі здійснює управління таблицею маршрутизації трафіку між хостами, проте, не має вбудованих засобів для прийняття рішення щодо реконфігурації мережі. Ідея запропонованого способу полягає в реалізації двох складових: агента моніторинга ресурсів та плагіну SDN-контролера для реконфігурації мережі.

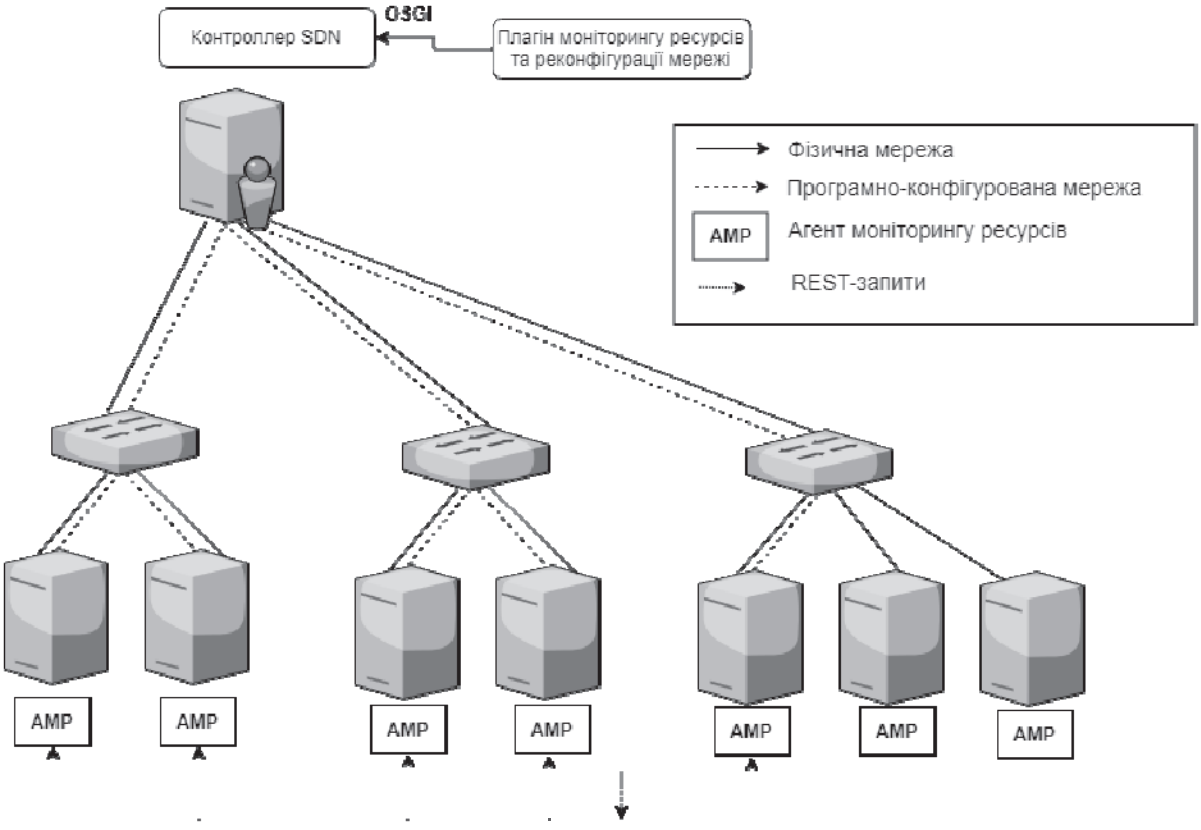


Рис. 1. Приклад кластеру в SDN з агентами та плагіном моніторингу.

Плагін може бути реалізований на будь-якій мові програмування, оскільки взаємодія з контролером можлива через REST API [3]. У випадку, наприклад, контролера Opendaylight, плагін реалізується на мові програмування Java з використанням технології OSGI [4], що дозволяє динамічно підключати плагіни без зупинки контролера. Агент моніторинга ресурсів виконується безпосередньо на тому ж вузлі, що і екземпляр розподіленого додатку. Для отримання даних про завантаження вузла використовуються розповсюджені інструменти, наприклад Zabbix [1] чи Nagios [2]. Основною задачею агента моніторинга ресурсів є надання актуальної інформації про навантаження на вузли кластера через API, доступний для контролера SDN.

Алгоритм реконфігурації SDN. При досягненні певного рівня навантаження, контролер приймає рішення про зміну топології мережі: перебудову таблиць маршрутизації таким чином, що трафік розподіляється на менш завантажені вузли. При падінні значення критерію завантаження вузлів до певних значень змінюється кількість вузлів в кластері: вузли відключаються від програмно-конфігурованої мережі [5]. Таким чином, підвищується ефективність використання обчислювальних ресурсів: звільняються сервера, які можуть бути використані для інших задач.



Рис. 2. Алгоритм реконфігурації SDNна основі даних моніторингу.

Висновки

Використовуючи запропонований спосіб може бути вдосконалений шляхом реалізації вдосконалених алгоритмів маршрутизації трафіку в плагіні, що підключається до контролера. Ще одним напрямком подальшої роботи є збір та аналіз інформації про обсяги трафіку, що надходить в мережу в певні проміжки часу, та реалізація способу прогнозування навантаження на хости в мережі. Це надасть можливість завчасно підключати додаткові ресурси до мережі, таким чином запобігаючи перевантаженню хостів.

Список використаних джерел

1. Imamagic E., Dobrenic D. Grid infrastructure monitoring system based on nagios //Proceedings of the 2007 workshop on Grid monitoring. – ACM, 2007. – С. 23-28.

2. Tader P. Server monitoring with Zabbix //Linux Journal. – 2010. – Т. 2010. – №. 195. – С. 7.

3. Medved J. et al. Opendaylight: Towards a model-driven sdn controller architecture //World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2014 IEEE 15th International Symposium on a. – IEEE, 2014. – С. 1-6.

4. BARAK L. Implementing a prototype for the Deep Packet Inspection as a Service Framework. – 2016.

5. Кулаков Ю. А., Лопушен Е. Ю. Способ масштабирования распределённых приложений в программно-конфигурируемых сетях с использованием гибридного облака //Міжнародний науковий журнал" Науковий огляд". – 2017. – Т. 8. – №. 40.

Довідка про авторів

Кулаков Юрій Олексійович - д.т.н., професор кафедри обчислювальної техніки НТУУ «КПІ»

Kulakov Jurij Oleksijovych - D.S., Professor, Department of Computer technics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Email: ya.kulakov@gmail.com

Лопушен Євгеній Юрійович - аспірант кафедри обчислювальної техніки НТУУ «КПІ» ім Ігоря Сікорського.

Lopushen Jevghenij Jurijovych – Ph.D. student, Department of Computer technics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Email: lopushen@gmail.com

Берест Руслана Юріївна- асистент кафедри обчислювальної техніки НТУУ «КПІ» ім Ігоря Сікорського.

Berest Ruslana Yuriivna – assistant professor, Department of Computer technics, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

Jurij Kulakov,
Jevghenij Lopushen,
Ruslana Berest

WAY OF COMPUTATIONAL RESOURCES MONITORING IN SOFTWARE-DEFINED NETWORKS

Target setting. The goal of the paper is to develop a universal way to monitor available resources in the SDN, which will allow the controller to obtain such data.

Actual scientific researches and issues analysis. The number of requests from users of the application is not constant, for example, during work hours, the load increases significantly, which can lead to delays in the work. At the same time, with less load it is not expedient to hold excess resources. **Uninvestigated parts of general issues defining.**Existing cluster monitoring methods do not support work with software-defined networks, but can integrate with the controller through the API.

The research objective. There is a need to come up with a way to monitor available resources in the software-defined networks, which will provide an opportunity to receive up-to-date data on the load of hosts on the network. The availability of such data makes it possible to configure the routing on the controller in such a way as to distribute the load evenly and prevent the host to reject the traffic, and in case of absence of resource demand - to disconnect the host from the network.

Ways of monitoring implementation in SDN. The proposed method is based on the main advantages of software-defined networks: the ability to reconfigure the network by editing the routing table through the controller API, and the ability to expand the controller functionality using plug-ins.

SDN reconfiguration algorithm. When a certain level of load is reached, the controller decides to change the topology of the network: the reconfiguration of the routing tables in such a way that traffic is allocated to less loaded nodes. When the value of the criteria of the nodes load drops to certain values, the nodes are disconnected from the software-defined network.

Conclusions. The paper suggested a way of resources monitoring in the cluster that resides in software-defined network.

Key words: software-defined networks, SDN controller, monitoring, monitoring agent, scaling.

УДК 004.75

Ріпневський Олександр

ВИКОРИСТАННЯ СУБД ІЗ АРХІТЕКТУРОЮ IN-MEMORY DATA GRID

Анотація. В даній статті розглядається концепція надшвидкої обробки даних за допомогою СУБД з архітектурою In-Memory Data Grid. Ця технологія є більш ефективною для багатьох задач, які мають у своїй основі саме вирішення питань швидкодії та продуктивності. Також розглянуті деякі проблеми, що виникають при використанні цієї технології, та запропоновані шляхи їх розв’язання.

Ключові слова: СУБД, Великі Дані, швидкість обробки, система обробки даних у оперативній пам'яті, динамічне завантаження класів.

Аннотация. В данной статье рассматривается концепция сверхбыстрой обработки данных с помощью СУБД с архитектурой In-Memory Data Grid. Эта технология является более эффективной для многих задач, которые имеют в своей основе именно решение вопросов быстродействия и производительности. Также рассмотрены некоторые проблемы, возникающие при использовании этой технологии, и предложены пути их решения.

Ключевые слова: СУБД, Большие Данные, скорость обработки, система обработки данных в оперативной памяти, динамическая загрузка классов.

Summary. This article discusses the concept of ultra-fast data processing with the help of DBMS with the structure of In-Memory Data Grid. This technology is more effective for many tasks, which are based on the decision of the issues of speed and performance. Also, some arisen problems with the use of this technology are considered, and ways of their solution are suggested.

Keywords: DBMS, Big Data, processing speed, data processing system in operational memory, dynamic load of classes.

У ряді завдань може знадобитися дуже висока швидкість обробки даних. Наприклад, біржовим гравцям іноді потрібно миттєво прийняти рішення, ґрунтуючись на великій кількості даних про стан ринку, - за пару секунд ситуація вже може змінитися. Існує також цілий ряд завдань, коли рішення потрібно приймати в реальному часі, наприклад обробка біометричних даних, якінеобхіднообробляти в режимы реального часу, звіряючисьіз базою даних про зловмисників. Дуже велика швидкість надходження даних характерна для багатьох наукових завдань. Наприклад, проект по запуску гігантського радіотелескопа з сумарною площею антен 1 км^2, передбачає передачу сигналів з однієї антени зі швидкістю 160 Гбіт / с, що в 10 разів перевищує весь нинішній інтернет-трафік [1].

Багато бізнес-завдання і наукові експерименти вимагають спільної обробки даних різних форматів - це можуть бути табличні дані в СУБД, ієрархічні дані, текстові документи, відео, зображення, аудіофайли та т.д.

В останній час все більше компаній звертається до обробки даних безпосередньо в оперативній пам’яті. Це є наслідком стрімкого падіння вартості оперативної пам'яті (RAM), в результаті чого стає можливим зберігання всього

набору операційних даних в пам'яті, збільшуючи швидкість їх обробки більш ніж в 1000 разів, порівняно із обробкою даних на жорстких дисках. Інструментарій для побудови таких рішень надають продути In-Memory Compute Grid та In-Memory Data Grid.

По суті, In-Memory Data Grid - це розподілене сховище об'єктів. Воно забезпечує швидку обробку інформації і надвисоку доступність за допомогою розподіленого зберігання даних в оперативній пам'яті. Це радикально контрастує з підходом традиційних СУБД, які розроблені для зберігання даних на стабільних носіях. Оскільки процеси обробки даних у оперативній пам'яті протікають швидше, ніж звернення до файлової системи та зчитування інформації з неї, СУБД в пам'яті забезпечує на порядок більш високу продуктивність програмних додатків. Оскільки конструкція СУБД в пам'яті набагато простіша традиційних, то вони накладають набагато менші вимоги до об'єму пам'яті та характеристик центрального процесора.

Традиційні СУБД для прискорення роботи широко використовують кешування. Кешування це процес, в рамках якого традиційні СУБД зберігають часто використовувані записи в пам'яті для швидкого доступу до них. Однак, кешування прискорює лише процес пошуку необхідної інформації, а не її обробки. Відповідно, виграш у продуктивності істотно менший.

Є ряд функціональних особливостей, які відрізняють IMDG від інших продуктів, таких як IMDB, NoSql або NewSql бази даних. Одна з основних - по-справжньому масштабується секціонування даних (Data Partitioning) в кластері. IMDG по суті являє собою розподілену хеш-таблицю, де кожен ключ зберігається на строго певному сервері в кластері. Чим більше кластер, тим більше даних можна в ньому зберігати. Принципово важливим у цій архітектурі є те, що обробку даних слід проводити на тому ж сервері, де вони розташовані (локально), виключаючи (або зводячи до мінімуму) їх переміщення по кластеру. Фактично, при використанні добре спроектованого IMDG, переміщення даних буде повністю відсутнє за винятком випадків, коли в кластер додаються нові сервера або видаляються існуючі, змінюючи тим самим топологію кластера і розподіл даних в ньому.

Зберігання даних в IMDG – це лише половина функціоналу, необхідного для in-memory архітектури. Дані, що зберігаються в IMDG, необхідно обробляти паралельно та з високою швидкістю. Зазвичай in-memory архітектура секціонує дані в кластері за допомогою IMDG, а згодом виконуваний код відправляється саме на ті сервера, де знаходяться необхідні йому дані. Оскільки виконуваний код (обчислювальна задача) зазвичай є частиною обчислювальних кластерів (Computer Grids), і повинен бути правильно розвернутий (deployment), бути збалансованим при навантаженні (load-balancing), бути відмовостійким (fail-over), а також мати можливість запуску по розкладку (scheduling), інтеграція між Compute Grid та IMDG дуже важлива. Найбільш ефективною є система, в якій IMDG та Compute Grid є частинами одного і того же продукту та використовують одні, й ті ж сами API, це дозволяє добитися найбільшої продуктивності та надійності in-memory рішення.

Серед рішень, які на ринку використовуються найчастіше – рішення Oracle, IBM Cognos TM1, SAP HANA, Microsoft PowerPivot, QlikView і Pentaho

Business Analytics. Такі платформи добре використовувати при необхідності аналізу даних в режимі реального часу з урахуванням того, що в процесі аналізу дані можуть бути змінені в будь-який момент. Також такі рішення добре підходять у випадках, коли немає можливості створення багатовимірного сховища даних і потрібно проводити аналіз даних облікової системи без її модифікації. Системи пропонують різні способи горизонтального масштабування, як з використанням засобів самої платформи, так і з застосуванням додаткового ПЗ.

Важливим аспектом In-Memory Data Grid (IMDG) реалізації розподілених сховищ даних є динамічне завантаження класів.

Динамічне завантаження – це завантаження під час виконання певного процесу.

Будь який клас, що використовується в середовищі виконання, був так чи інакше туди завантажений будь-яким завантажником Java. Зазвичай класи завантажуються за необхідністю їх використання. Окрім цього базові класи завантажуються під час старту програми.

Існує два варіанти завантаження класів:

Статичне завантаження класів – це звичайне завантаження, що проводиться автоматично. При старті програми завантажник класів рекурсивно завантажує усі класи, що зустрічаються у програмі починаючи з основного класу. Об'єкти таких класів створюються стандартним способом за допомогою оператора new;

Динамічне завантаження класів відтворюється за допомогою використання Classloader. Динамічне завантаження має сенс, коли необхідно завантажити клас під час виконання програми, коли необхідно замінити клас, змінив якусь певну логіку, і коли при цьому перезапуск програми не є раціональним. На сьогодні склалась ситуація, коли жодна із систем, жодна з існуючих систем з архітектурою In-Memory Data Grid повністю не реалізує динамічне завантаження класів. Це призводить до необхідності перезавантаження системи, або окремої частини, де і буде використаний необхідний клас, і, відповідно, при необхідності внесення певних змін, неможливо буде це зробити без втрати даних. Крім того, такий підхід потребує часу на перезапуск.

Підхід динамічного завантаження дає змогу уникнути цих проблем. Перезавантаження будь-якого класу може бути ініційовано в будь-який момент часу, за умови що в цей час вже не відбувається перезавантаження цього класу. Під час операції завантаження тимчасово блокується можливість змінювати вже існуючі об'єкти старої версії класу, коли у випадку перезавантаження кластера доступ був би втрачений повністю.

Ідея динамічного завантаження класів, як і ідея звичайного завантаження, полягає у тому, що коли певні класи зазнали змін, що впливають на виконання програми, то необхідно перекомпілювати саме змінений файл для отримання актуальних результатів після проведених змін. Відмінністю є той факт, що при звичайному завантаженні є можливість призупинити сервіс або програму, що залежить від зміненого класу, перекомпілювати необхідний ресурс, а згодом

запустити всю систему з нуля, або перезапустити лише ту частину, на яку вплинули ці зміни.

Для забезпечення простоти та можливості використання інструменту динамічного завантаження в різних сервісах, було вирішено розробити модель, що повинна визначати зміни у файлах, що задіяні, а також самостійно завантажувати змінені класи. Маючи на увазі те, що перезавантаження має відбуватись в IMDG, було вирішено позиціонувати програму як доповнення до розподіленого сховища даних, що починає перезавантаження при надсиланні клієнтом класу.

Таким чином, приходимо до висновку, що сучасний стан функціонування СУБД із архітектурою In-Memory Data Grid потребує розробки алгоритму реалізації інструменту динамічного завантаження. На нашу думку цей алгоритм повинен складатися із 4-х етапів.

1) Ініціалізація: завантаження класу, конверторів, створення мапи зав'язків полів та конвертерів на основі описуючих фалів;
2) Підготовка: генерація конвертерів;
3) Обробка: створення нових об'єктів на основі старих;
4) Заміщення: заміщення старих даних новими.

Розробка вказаного алгоритму та програмного забезпечення на його основі значно підвищить швидкодію та надійність обробки Великих Даних у СУБД із архітектурою In-Memory Data Grid.

Література

1. Найдич А. Big Data: проблема, технология, рынок [Електронний ресурс]. Доступ: https://compress.ru/article.aspx?id=22725.

2. Обухов А. In-Memory. База данных в оперативной памяти [Електронний ресурс]. ECM Journal. Доступ: https://ecm-journal.ru/post/In-Memory-Baza-dannykh-v-operativnojj-pamjati.aspx

3. Underwood, Jen. Exploring 2017 gartner bi magic quadrant results [Електронний ресурс]. Доступ:http://www.jenunderwood.com/2017/02/22/2017-gartner-bi-magic-quadrant-results/

ДОВІДКА ПРО АВТОРІВ

Ріпневський Олександр – студент, кафедра СПіСКС, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського». Ripnevkij Oleksandr – student, Department of SP&SCS, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

E-mail: alexripnesky@gmail.com

Ріпневський Олександр

ВИКОРИСТАННЯ СУБД ІЗ АРХІТЕКТУРОЮ IN-MEMORY DATA GRID

Актуальність теми. В даній роботі розглянуті використання СУБД у розподілених сховищ даних. Описано основні функції та розглянуто механізми обробки даних в IMDG. Окремо розглянутий механізм завантаження класів до сховищ в оперативній пам'яті, та описано існуючі рішення його оптимізації.

Постановка проблеми. Відсутність добре інтерпретованого методу оптимізації механізму та створення СУБД до розподілених сховищ даних в оперативній пам'яті.

Постановка завдання. Завданням є описати використання баз даних і систем їх управління у використанні їх у розподілених сховищах даних та розробка алгоритму, що зможе використовувати можливості даного рішення максимально.

Викладення основного матеріалу. Проведено аналіз існуючих проблем у використанні звичайних баз даних у роботі систем розподілених даних. Описано можливі варіанти покращення їх роботи. Результати виявились добре інтерпретованими та змістовними.

Висновки. Проаналізовано параметри, проблематику та вирішення проблем, що мають місце під час роботи з розподіленими сховищами даних, які показали найкращі результати для генерації текстів. Було розроблено алгоритм, що дозволить використати можливості СУБД у розподілених сховищах в більшому степені.

УДК 004.074.32

Олександр Марковський,
Вікторія Максимук

МЕТОД СТРОГОЇ АВТЕНТИФІКАЦІЇ ВІДДАЛЕНИХ АБОНЕНТІВ НА ОСНОВІ АЛГЕБРИ ПОЛІВ ГАЛУА

METHOD FOR STRONG REMOTE USERS BASED ON GALOIS FIELDS ALGEBRA

Представлений новий метод реалізації теоретичної строгої перевірки віддалених користувачів на основі концепції «нульового знання». Запропонований підхід полягає в застосуванні математичної операції експоненціювання на полях Галуа замість модульної експоненти. Це дозволяє прискорити процес ідентифікації для апаратної реалізації. Технологія математичних перетворень, що надаються запропонованим методом, викладена чітко. Наведено числовий приклад для розроблених процедур реєстрації користувача.

Ключові слова: концепція «нульових знань», ідентифікація віддалених користувачів, автентифікація користувачів, незворотні перетворення на полях Галуа.

Бібл.: 5.

The new method for implementation of theoretical strong autentification of remote users based on zero-knowledge conception is presented. The proposed method consist of using of mathematical operation of exponentiation on Galois fields instead of modular exponentiation. It allows to speed up of identification process for hardware implementation. The technology of mathematical transformations whose are provided by proposed method are set forth clearly. A numerical example for designed procedures of user registration.

Key words: zero-knowledge conception, remote users identification, users authentication, irreversible transformation in Galois fields.

Bibl.: 5.

Актуальність теми дослідження. З появою технологій комп'ютерних мереж прогрес в більшості областей людської діяльності значною мірою визначається інтеграцією інформаційних ресурсів. Можливість доступу до якісно більш широких об'ємів інформації дозволяє значно підвищити якість прийняття рішень, прискорити та здешевити їх розробку. Необхідною умовою інформаційної інтеграції є застосування ефективних механізмів контролю доступу до даних, важливе місце серед яких займають засоби ідентифікації та автентифікації віддалених абонентів.

Поява та динамічний розвиток хмарних технологій надає користувачам значні за обсягом обчислювальні ресурси, які можуть бути використані потенційними зловмисниками для зруйнування існуючих механізмів контролю доступу до інформації. Це об'єктивно вимагає здійснення заходів для підвищення надійності механізмів автентифікації віддалених користувачів. Разом з тим, поява хмарних технологій має наслідком значне зростання кількості користувачів, для якісного обслуговування яких потрібно радикально прискори-

ти процедури контролю їх доступу до інформаційних та обчислювальних ресурсів.

Таким чином, задача підвищення ефективності засобів автентифікації віддалених користувачів є актуальною та важливою з огляду особливостей сучасного етапу розвитку інформаційних технологій.

Постановка проблеми. В сучасних умовах і у найближчій перспективі потрібний для практики рівень надійності автентифікації може бути забезпечений лише методами, що спираються на концепцію нульових знань [2]. Ця концепція передбачає наступні елементи:

- Пароль користувача змінюється при кожному зверненні до системи.
- Користувач має спеціальний механізм генерації правильних паролей.
- Система не може генерувати правильних паролей, але має механізм перевірки їх правильності.

Для побудови ідентифікації на основі концепції нульових знань потрібне використання спеціальних криптографічних незворотних та неоднозначних перетворень. Ефективність ідентифікації визначається рівнем не оберненості перетворення та об'ємом часових ресурсів на його реалізацію. Проблема побудови ефективної схеми ідентифікації зводиться до найбільш вдалого розв'язання компромісу між вказаними чинниками.

Аналіз основних досліджень і публікацій. Найбільш відомими методами, які реалізують цю концепцію є FFSIS(Feige Fiat Shamir Identification Scheme)[3], методи Шнора (Schnorr) та Гіллоу—Квіскатера (Guillou—Quisquater) [1]. Базовими обчислювальними операціями для FFSIS є $A^2B \bmod m$, *а для методів Шнора і Гіллоу*—Квіскватера—$A^eB^v \bmod m$. Враховуючи те, що задля забезпечення потрібного рівня захищеності розрядності чисел становить 2048 або 4096, швидкість ідентифікації обмежуються значною обчислювальною складністю цих операцій і, відповідно, виникає необхідність в її зменшенні.

Таким чином, основний недолік існуючих методів реалізації концепції нульових знань полягає в значній обчислювальній складності їх базових обчислювальних процедур.

Виділення недосліджених частин загальної проблеми. Аналіз існуючих систем строгої ідентифікації показав, що основним недоліком існуючих протоколів та механізмів є значна обчислювальна складність. В сучасних умовах зростання розрядності цей недолік має наслідком порушення часових рамок протоколів мережевого захисту. Тому важливим завданням є розробка нових незворотних перетворень обчислювальна реалізація яких потребує менше часу. Одним з найбільш перспективних напрямків є перехід до альтернативних алгебраїчних базисів.

Постановка завдання. Метою дослідження є зменшення обчислювальної складності процесу криптографічно строгої автентифікації на основі концепції нульових знань при забезпеченні високого рівня захищеності за рахунок застосування альтернативного алгебраїчного базису – скінченних полів Галуа.

Теоретичне підґрунтя. Аналогічно тому, як у традиційній алгебрі в якості базової операції широкого кола механізмів криптографічного захисту використовується модулярне експоненціювання $A^E \bmod M$, в алгебрі скінчених полів Галуа для криптографічних застосувань в якості базової застосовується експоненціювання $A^E \operatorname{rem} M$ полях [4].

На практиці обчислення експоненти $R_1 = A^E \text{ rem } M$ реалізується рекурсивною процедурою, яка для n-розрядної експоненти $E = e_1 + 2\cdot e_2 + 2^2\cdot e_3 + \ldots + 2^{n-1}\cdot e_n$, $e_1, e_2, \ldots, e_n \{0,1\}$, передбачає послідовне, починаючи з R_n=1 і j=n, обчислення значень $R_{n-1}, R_{n-2}, \ldots, R_1$ з використанням наступної формули:

$$R_{j-1} = (R_j \otimes R_j) rem\ M \otimes (A\cdot e_j \oplus e_j \oplus 1)\ rem\ M \quad (1)$$

де символом ‘’ позначена операція поліноміального множення, або множення без переносів (Multiplication Without Carry -MWC), а абревіатурою “rem” – операція редукції на полі Галуа, тобто віднаходження залишку при поліноміальному діленні результату множення без переносів на утворюючий поліном M поля.

Якщо $g(x)$ – простий поліном ступеню d, то для будь-якого $u(x)$,що є елементом поля GF(2^d) і якому співвідноситься d-розрядне двійкове число u, таке, що 0<u h, де h=2^d-1 виконується [4] :

$$u\ |^{h+1}\ rem\ g = u \quad (2)$$

Якщо утворюючий поліном $M(x)$ ступеню r являє собою добуток двох простих поліномів $p(x)$ ступеню v і $g(x)$ ступеню d : $M(x)=p(x)g(x)$, $r = v+d$, причому поліному $M(x)$ співвідноситься двійкове число m то для будь-якого $u(x)$, що належить полю GF(2^v) і з яким співвідноситься число u, u $h=2^d$-1 справедливо:

$$(u \otimes p)\ |^{h+1}\ rem\ m = u \otimes p \quad (3)$$

Наприклад, якщо $p(x) = x^3+x^2+1$ простий поліном ступеню v=3, якому відповідає число p=13, для якого $l = 2^3$-1=7, і $g(x)=x^5 + x^4 + x^3 + x + 1$ простий поліном ступеню d= 5, якому відповідає число g= 59, для якого $h = 2^5$-1 = 31, то добуток цих поліномів має вигляд $M(x) = p(x)\ g(x) = x^8 + x^3 + x^2 + x + 1$, і співвідноситься з числом m=271. Якщо вибрати число u таким чином, що 0<u h-1, зокрема, u = 28, то up = 140. Тоді $(up)|^{h+1}$ rem m = $(2813)|^{32}$ rem 271 = 140.

Аналогічно, для будь-якого поліному $w(x)$, що належить полюGF(2^d) та співвідноситься з числом w, так, що wl=2^v-1 справедливо:

$$(w \otimes g)\ |^{l+1}\ rem\ m = w \otimes g \quad (4)$$

Якщо в рамках наведеного прикладу вибрати довільне w менше за l: 0 <w<l, наприклад, w= 5, то добуток wg = 559 = 215. Відповідно, $(wg)^{l+1}$ rem m = $(5\ 59)^8$ rem 271 = 215, тобто дорівнює згідно (3) w g. Доведення формул (3) і (4) виконано в роботі [5]

Таким чином, знання однією зі сторін процесу ідентифікації періоду повторення експоненціювання дозволяє будувати ефективні алгоритми ідентифікації в рамках криптографічно строгої концепції “нульових знань”.

Метод строгої автентифікації абонентів з використанням алгебри полів Галуа. Запропонований в попередньому розділі підхід до реалізації ідентифікації віддалених користувачів в рамках криптографічно строгої концепції “нульових знань” з використанням алгебри полів Галуа може бути модифікованим для реалізації задачі автентифікації віддалених користувачів.

Процедура автентифікації, на відміну від ідентифікації, дозволяє не тільки гарантувати, що система взаємодіє з авторизованим користувачем, але надає користувачеві механізми, які дозволяють впевнитися в тому, що він взаємодіє дійсно з системою.

Пропонована процедура реєстрації передбачає наступну послідовність дій:

1) Користувач отримує від системи її відкритий закриваючий ключ $К_с$.

2)Користувач довільним чином вибирає пару простих поліномів $p(x)$та $g(x)$з різними степенями: $p(x)= x^v + p_{v-1}\cdot x^v + \ldots + p_1\cdot x + p_0$ степені v та $g(x)= x^d +$

$+ g_{d-1} \cdot x^{d-1} + \ldots + g_1 \cdot x + g_0$ степені d, де $p_0, p_1, \ldots, p_{v-1}$ {0,1}, $g_0, g_1, \ldots, g_{d-1}$ {0,1}, причому $d>v$.

3) Користувач формує поліном $M(x)$ у вигляді поліноміального добутку вибраних двох простих поліномів $p(x)$та $g(x)$: $M(x) = p(x)g(x)$. Число m, з яким співвідноситься поліном $M(x)$ являє собою першу компоненту відкритого ключа користувача.

4) Користувач вибирає випадкове число : $0 << 2^d$ та обчислює другу компоненту відкритого ключа користувача у вигляді: $= p$ rem m.

5) Користувач вибирає довільне число $<2^{d+v}-2^{d+1}$. Це число являє собою третю компоненту ключа користувача.Користувач обчислює значення коду $= 2^{d+v}-2^d-2^v + 1 -$; значення зберігається в пам'яті користувача.

6) Всі три компоненти m , та ключа користувача шифруються відкритим закриваючим ключем $К_с$ і відсилаються системі.

6) Система з використанням секретного відкриваючого ключа $К_о$ відновлює значення трьох компонентів m, та ключа користувача після чого зберігає їх в захищеній пам'яті.

Пропонована процедура реєстрації може бути ілюстрована наступним прикладом. У відповідності з п.2 Користувач вибирає простий поліном ступеню 3 (v=3) $p(x)= x^3 + x^2 + 1$, яке співвідноситься з числом p=11, та простий поліном степені 5 (d=5) $g(x) = x^5 + x^4 + x^3 + x + 1$, якому співвідноситься число g = 59. Згідно з п.2 користувач формує утворюючий поліном $M(x)$ як поліноміальний добуток вибраних простих поліномів: $M(x) = p(x)g(x) = x^8 + x^7 + x^4 + x^2 + 1$, який співвідноситься з числом m = 405. Згідно з п.4 процедури реєстрації користувач випадковим чином обирає число = 18 та обчислює другу компоненту ключа: $= p$ rem $m = 11^{18}$ rem 405 = 49. Згідно п.5 запропонованої процедури користувач вибирає довільне менше $2^{d+v}-2^{d+1} = 192$. Нехай, вибране значення дорівнює 18, тобто =18. Користувач обчислює значення $= 2^{d+v}-2^d-2^v + 1 - = 218 - 18 = 200$.

Трійка компонентів ключа користувача < 405, 49, 18 > шифрується відкритим ключем $К_с$ системи та відправляється в систему.

Розроблена процедура одного циклу автентифікації віддаленого абонента передбачає виконання наступної послідовності дій:

1) Користувач ініціює звернення до системи.

2) Система формує випадкове число r, з використанням компонентів m та ключа користувача обчислює значення $= r$ rem m. Коди r та відсилає користувачеві.

3) Користувач приймає від системи код r та . Для перевірки того, що він працює дійсно з системою, обчислює $= r$ rem m. Якщо $= r$, тобто поліноміальний добуток обчисленого на отриманий код дорівнює отриманому від системи числу r, то користувач впевнюється, що сторона, яка надіслала йому коди r та дійсно знає компоненти його ключа m та , тобто є системою.

4) Впевнившись, що комутація відбувається дійсно з системою, користувач віднаходить число w таке, для якого виконується умова: $r \cdot w \bmod (2^d - 1) =$.

5) Користувач обчислює сеансовий пароль у вигляді $= p|$ rem m після чогошифрує за допомогою відкритого ключа $К_с$ отриманий код та надсилає його системі.

6) Система, отримавши зашифрований сеансовий пароль , розшифровує його своїм секретним відкриваючим ключем $К_о$ та обчислює $y = |^r$ rem M.

Отримане значення у порівнюється з другою компонентною відкритого ключа цього користувача: якщо ці коди співпадають, тобто $y =$, то сеанс автентифікації вважається успішним і користувачеві надається доступ до ресурсів системи.

Робота запропонованої процедури одного сеансу ідентифікації може бути ілюстрована наступним прикладом, який фактично є продовженням попереднього.

У відповідності з п.2 процедури, у відповідь на звернення користувача, система випадковим чином вибирає число $r = 28$ і обчислює $= r \text{ rem } m = 28^{18} \text{ rem } 405 = 38$, після чого пересилає його користувачеві $r=28$ та $=38$.

Користувач приймає від системи надіслані йому коди $r = 28$ та $=38$. Для того, щоб впевнитися, що сторона, яка надіслала йому ці коди дійсно знає його ключ, тобто є системою користувач згідно п.3 обчислює $= r \text{ rem } m = 28^{200} \text{ rem } 405 = 117$. Далі користувач обчислює поліноміальний добуток $= 11738 = 28$. Оскільки значення обчисленого добутку співпадає з отриманим кодом $r = 28$, користувач фактично ідентифікує сторону, що веде з ним інформаційний обмін як систему, що знає його реєстраційний ключ.

У відповідності з п.4 описаної вище процедури, користувач віднаходить таке w, для якого виконується умова: $28w \bmod 31 = 18$. Ця умова виконується для $w= 25$. Згідно з п.5 користувач обчислює сеансовий пароль у вигляді $= p| \text{ rem } m = 11|^{25} \text{ rem } 405 = 245$. Цей сеансовий пароль користувач посилає системі. Згідно з п. 6, система отримавши зашифрований сеансовий пароль , розшифровує його своїм секретним відкриваючим ключем K_o та обчислює $y = |^r \text{ rem } m = 245|^{28} \text{ rem } 405 = 49$. Це значення порівнюється з другою компонентою ключа користувача. Оскільки обчислене значення $y = 49$ співпадає зі значенням другої компоненти реєстраційного ключа цього конкретного користувача = 49, то система успішно його ідентифікує.

Оцінка ефективності. Оцінку рівня захищеності доцільно виконувати з двох позицій: з позиції третьої сторони, яка може мати доступ до каналу обміну ідентифікаційною інформацією, але не знає коду утворюючого поліному $M(x)$ поля Галуа; з позицій зловмисника, що має доступ до даних, що зберігаються в системі, і зокрема коду $M(x)$. $p(x)$та $g(x)$

Очевидно, що для того, щоб виконати успішний підбір коректного паролю користувача потрібно знати простий поліном $g(x)$, а також степінь v іншого простого поліному $p(x)$. Найдоцільніша технологія підбору компоненти паролю $p(x)$ для стороннього зловмисника полягає в перехопленні коду $q(x) = k(x)p(x)$ та розкладенні його на співмножники. Важливим моментом при цьому є те, що у зловмисника практично відсутній критерій коректності розкладення. Ще одна особливість задачі розкладання коду $q(x)$ на співмножники полягає в тому, що один з них не є простим, що потенційно збільшує обсяг перебору. Кількість можливих простих поліномів швидко зростає зі збільшенням їх степеня. Та уже при степені 29 кількість простих поліномів перевищує $18 \cdot 10^6$. На практиці степінь поліному становить сотні і, відповідно, задача перебору простих поліномів далеко виходить за рамки технічних можливостей сучасних комп'ютерів.

Для зловмисника, що має доступ до ідентифікаційної інформації, яка зберігається в системі, задача відтворення коректного паролю зводиться до задачі розкладання відомого поліному $M(x)$ на два простих множника $p(x)$ та $g(x)$ з різними ступенями. Як уже зазначалося, вище вирішення такої задачі шляхом перебору для реальних степенів поліномів виходить далеко на межі технічних можливостей сучасних комп'ютерних систем.

Висновки. На основі отриманих теоретичних результатів запропоновано методи ідентифікації та автентифікації віддалених користувачі, що реалізують криптографічно строгу концепцію “нульових знань” в алгебрі полів Галуа. Оскільки ці операції виконуються на один-два порядки швидше за мультиплікативні операції модулярної арифметики, що лежать в основі відомих методів криптографічно строгої ідентифікації, використання запропонованих методів дозволяє суттєво прискорити процес автентифікації відлених користувачів.

Список літератури

1. Schneier B. Applied Cryptography. Protocols. Algorithms and Source codes in C. -Ed.John Wiley, 1996 - 758 pp.

2. Feige U. Zero knowledge proofs of identity / Feige U.,Fiat A.,Shamir A.// Journal of Cryptology.- v.1.- №.2.- 1988, pp.77-94.

3. Markovskyy O. Fast subscriber identification based on the zero knowledge principle for multimedia content distribution/ O. Markovskyy, N. Bardis, N. Doukas // International Journal of Multimedia Intelligence and Security 2010 - Vol. 1,- pp.78-82.

4. Mukhin, V.E., Zacharioudakis Leftherios, Gerasimenko, O.Yu. and Kozeratskiy, M.S. (2017), “Method of zero-knowledge identification of remote users based on the conception of “zero knowledge”, Telecommunikatsiyni ta informatsiyni tekhnologii, Vol. 54, no. 1, pp. 37-45.

5. Марковський О.П. Використання алгебри полів Галуа для реалізації концепції «нульових знань» при ідентифікації та автентифікації віддалених користувачів /О.П. Марковський, Захаріудакіс Лефтеріс, В.Р. Максимук // Электронное моделирование.-том 39.-№6.-2017.-С.33-46.

ДОВІДКА ПРО АВТОРІВ

Марковський Олександр Петрович – кандидат технічних наук, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

Oleksandr Markovskyi – Associate Professor, PhD, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Максимук Вікторія Романівна – студентка, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Viktoriia Maksymuk – student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

Марковський О. П.,
Максимук В. Р.

METHOD FOR STRONG REMOTE USERS BASED ON GALOIS FIELDS ALGEBRA

Acute problem. A particularly acute problem, is the problem of fast algorithms for identification architectures based on the concept of "zero – knowledge" in mobile devices and embedded microcontrollers, with limited energy consumption possibilities, that are commonly end user devices. In this context, the problem of acceleration of user identification based on zero knowledge is imminent and has a wide range of practical applications.

Target setting. FFSIS is a relatively simple and at the same time sufficiently effective scheme for the identification of the subscribers of multi-user systems, on the basis of which a number of more practical to use modified algorithms have been proposed. In the attempts of using this scheme in practice, the main disadvantage of FFSIS is the need for a large number of data exchanges during the user identification process, which noticeably loads the communication channels used. Other existing identification schemes, which implement the zero knowledge concept, require a substantially smaller volume of data transfers, but the procedures provided by them involve large computational complexity, since instead of the operation of squaring, they use the modular exponential operation.

Actual scientific researches and issues analysis. In literature identification methods, which satisfy the first three of the given requirements are classified as "strict", in contrast the remaining schemes that are classified as "weak". In the class of the weak schemes belong, for example, the procedure of identification which used in the UNIX operating system. This procedure involves the storage in the system of only the hash value of the passwords of users, that, with the use of the one way hash functions, excludes the possibility of the reproduction of password of the system; however, passwords themselves do not change, which makes it sufficient simple to intercept them. The class of strict procedures is principally composed by methods of identification that are based the concept "zero knowledge". The most commonly known of these methods are the FFSIS (Feige Fiat Shamir Identification Scheme) [3], Guillou- Quisquater [4] and Schnorr identification schemes [5].

Uninvestigated parts of general matters defining. The main disadvantage of known identification schems consist of that schemes demand a significantly smaller volume of data exchanges compared to the FFSIS, but their implementation involves a significantly large computational volume, as the squaring operation has been replaced by the modular exponentiation. The FFSIS is considered more economic in terms of the volume of the calculations involved, but its application demands several cycles of information exchange.

The research objective. The purpose of this research is the development of a modified architecture for zero knowledge user identification, which involves significantly smaller computational complexity and increases the speed of identification with software and hardware implementations.

The statement of basic materials. The proposed method, contrary to existing techniques uses a single session for the registration between the user and the system.

Conclusions. In this paper, zero – knowledge user authentication using non – reversible transformation in Galois fields was investigated and a scheme for such authentication was proposed. For this purpose a theoretical study was conducted of the properties of specific cycles that appear during exponentiation in the context of special cases of Galois fields. The theoretical results enabled the development of procedures for the registration and authentication of remote users. It was theoretically determined that the level of security attained does not differ from existing zero knowledge systems. The principal advantage of the proposed method is the possibility of achieving significantly higher rates of authentication. This increased rate is of vital importance in current applications with ever growing numbers of system users and the necessity for remote information processing. Maximum effectiveness of the proposed method is attained by implementation in hardware.

Key words: zero-knowledge conception, remote users identification, users authentication, irreversible transformation in Galois fields.

УДК 004.056.5

Марковський О. П.,
Білашевська А. В.

МЕТОД КОРЕКЦІЇ БАГАТОКРАТНИХ ПОМИЛОК В КАНАЛАХ ПЕРЕДАЧІ ЦИФРОВИХ ДАНИХ ЗІ СПЕКТРАЛЬНОЮ МОДУЛЯЦІЄЮ

METHOD FOR MULTIPLE ERRORS CORRECTION IN SPECTRAL MODULATION DATA TRANSMISSION CHANNEL

Стаття присвячена проблемі підвищення ефективності виправлення двократних помилок передачі даних в каналах зі спектральною модуляцією за рахунок спрощення та прискорення обчислень, пов'язаних з корекцією помилок.

В роботі запропоновано, теоретично обґрунтовано і досліджено підхід до корекції багатократних помилок в каналах зі спектральною модуляцією на основі позиційних корегуючих сум, що дозволяє визначати позиції спотворених символів та вектори їх спотворення без перебору всіх символів блоку.

Ключові слова: багатократні помилки передачі даних, виявлення помилок, виправлення помилок, кодування помилок, зважені контрольні суми, циклічні і лінійні корегуючи коди.

The paper is dedicated to the efficiency improving of double error correction of data transmission in channels with spectral modulation by simplifying and accelerating the computations associated with error correction.

The paper presents a theoretically grounded and investigated approach of double errors correction in channel with spectral modulation based on positional correcting sums that allows to determine the positions of distorted symbols and their distortion vectors without enumeration through all the symbols of block.

Key words: multiple data transmission errors, error detecting, error correction, coding errors, weighted checksums, cyclic and linear redundancy correcting codes.

Актуальність теми дослідження. Розвиток розподілених комп'ютерних систем в останнє десятиліття стимулювався, головним чином, потужним прогресом засобів передачі даних між їх компонентами. Чільне місце в сучасних технологій передачі даних в комп'ютерних мережах та розподілених системах займають канали передачі зі спектральною модуляцією. Цифрові дані в таких каналах передаються посимвольно і на фізичному рівні передача символу співвідноситься з відповідною стрибкоподібною зміною фази та амплітуди несучого синусоїдального сигналу [1].

Продуктивність розподіленої обробки даних значною мірою визначається швидкістю обміну даними між компонентами комп'ютерних систем. Тому, в найближчій перспективні, швидкість передачі даних буде зростати, в тому числі і за рахунок зменшення надійності, маючи на увазі те, що воно, в принципі, може бути скомпенсоване використанням більш потужних методів виправлення помилок [2].

Таким чином, наукова задача прискорення підвищення ефективності корекції помилок даних в каналах передачі даних зі спектральною модуляцією з огляду на сучасний стан розвитку технологій комп'ютерної обробки інформації є актуальною.

Постановка проблеми. Зростання швидкості обміну цифровими даними між компонентами комп'ютерних систем диктує необхідність вдосконалення існуючих і створення нових методів та засобів виправлення помилок, що виникають при передачі цифрових даних.

Аналіз існуючих способів корекції помилок в каналах зі спектральною модуляцією. Вважаючи на практичну важливість проблеми виправлення помилок передачі даних в бездротових каналах, на які вливають зовнішні електромагнітні завади, до теперішнього часу створено ряд методів корекції помилок передачі символів [3].

Ефективність засобів корекції помилок характеризується через три групи критеріїв:

– функціональні, що характеризують здатність вирішувати задачу виявлення та виправлення помилок;

– обчислювальні, що характеризують реалізацію корекції на програмному та апаратному рівнях;

– інформаційні, що характеризують рівень інформаційної надлишковості для виправлення помилок.

До функціональних критеріїв відносяться:

1. кратність помилок, що можуть бути гарантовано виправлені без повторної передачі;
2. ймовірності виправлення помилок кратності, більшої за гарантовану;
3. кратність помилок, які гарантовано виявляються.

До розряду обчислювальних критеріїв відносяться:

1. обчислювальна та часова складність обчислювальних процедур виявлення та корекції помилок,
2. часова складність обчислювальних процедур виявлення та корекції помилок,
3. складність апаратних засобів корекції помилок.

В якості інформаційного критерію ефективності засобів корекції помилок найчастіше виступає кількість контрольних розрядів, що передаються разом з блоком даних.

Динаміка розвитку технологій передачі даних має наслідком зміщення акцентів значимості наведених критеріїв ефективності методів контролю помилок. Багатократне підвищення швидкості передачі зумовлює зростання значимості оцінки обчислювальної складності операцій декодування, виявлення помилок та їх корекції, яка визначає можливість виконання цих операцій в темпі передачі. Цей же чинник знижує значимість, як критерію ефективності, числа контрольних розрядів, що додатково передаються.

Стосовно каналів передачі даних зі спектральною модуляцією засоби корекції помилок мають локалізувати спотворені при передачі дані, а також визначити вектори спотворення кожного із пошкоджений символів.

Якщо вважати, що інформаційний блок B містить $n=2^k-1$ m-розрядних канальних символів $B=\{X_1, X_2,...,X_n\}$, кожен j-тий ($j\in\{1,..,n\}$ з яких X_jпередається одним канальним сигналом іскладається з m бітів: $X_j=\{x_{j1},x_{j2},...,x_{jm}\}$, $\forall i\in\{1,..,m\}$: $x_{ji}\in\{0,1\}$, то теоретичний мінімум k_m кількості контрольних розрядів, потрібних для виправлення h спотворених символів становить:

$$k_m = h \cdot (log_2 n + log_2 m) \quad (1)$$

Класичним засобом виправлення багатократних помилок в каналах зі спектральною модуляцією є корегуючі коди Ріда-Соломона [4]. Ці коди являють собою недвійкові циклічні коди, що здатні виправляти довільну кількість h спотворених при передачі символів з використанням $2 \cdot h$ контрольних символів. Використання кодів накладає обмеження на довжину блоку:

$$n \leq 2^m - 1 \quad (2)$$

Таким чином, при урахуванням обмеження (2), кількість контрольних символів в кодах Ріда-Соломона відповідає теоретичному мінімуму.

Найбільш суттєвим недоліком кодів Ріда-Соломона вважається [3] висока обчислювальна та часова складність. Це зумовлено тим, що для розв'язання системи нелінійних рівнянь на полях Галуа, до якого в математичному сенсі, зводиться локалізація помилок та визначення векторів спотворень, використовується авторегресійна технологія, яка фактично реалізується перебором всіх n символів блоку.

Час T_0 виконання виявлення помилок для кодів Ріда-Соломона визначається формулою:

$$T_0 = 2 \cdot h \cdot n(t_m + t_{XOR}) \quad (3)$$

де t_m- час виконання множення на полях Галуа, t_{XOR}- час виконання операції додавання за модулем 2. Враховуючи, що операція множення складається з m циклів зсуву та додавання за модулем 2, обчислювальна складність виявлення помилок складає $O(8 \cdot h \cdot n \cdot m)$.

Процес корекції помилок включає розв'язання двох систем з h лінійних рівнянь та обчислення синдрому наявності помилки для кожного з символів блоку, так, що загальний час корекції оцінюється формулою [3]:

$$T_s = (n \cdot h + 8 \cdot h^2) \cdot (t_m + t_{XOR}) \quad (4)$$

Відповідно, обчислювальна складність операцій, пов'язаних з корекцією помилок становить $O(n \cdot m \cdot h^2)$.

Виділення недосліджених частин загальної проблеми. В сучасних умовах збільшення швидкості передачі, розрядності символів та довжини блоку, використання кодів Ріда-Соломона не забезпечує для багатьох важливих для практики застосувань можливість корекції в темпі передачі даних.

Постановка завдання. Ціллю дослідження є розробка методу прискореної корекції помилок в каналах зі спектральною модуляцією та спрощення апаратної реалізації корекції.

Метод швидкої корекції багатократних помилок в каналах зі спектральною модуляцією. Контрольований інформаційний блок B містить $n=2^k-1$ m-розрядних канальних символів $B=\{X_1, X_2,...,X_n\}$, кожен j-тий ($j \in \{1,..,n\}$ з яких X_jпередається одним канальним сигналом іскладається з m бітів: $X_j=\{x_{j1},x_{j2},...,x_{jm}\}$, $\forall i \in \{1,..,m\}: x_{ji} \in \{0,1\}$.

Для виявлення багатократних помилок і гарантованого виправлення двох помилок пропонується метод, яких передбачає передачу разом з інформаційним блоком контрольного блоку, який складається з $2 \cdot k+1$ компонентів: . $C=\{C_0,C_1,C_2,...,C_k, S_1,S_2,...,S_k\}$.

Спосіб формування контрольного коду. Нульову компоненту контрольного коду C_0 пропонується обчислювати у вигляді суми за модулем 2 всіх символів блоку:

$$C_0 = X_1 \oplus X_2 \oplus \cdots \oplus X_n \quad (5)$$

Кожен j-тий номер символу X_j блоку B може бути представлений у вигляді двійкового k-розрядного коду: $j = j_1 + j_2 \cdot 2 + j_3 \cdot 2^2 + \ldots + j_k \cdot 2^{k-1}$. Друга компонента C_1 пропонованого контрольного коду являє собою суму за модулем 2 тих символів блоку B, молодший розряд номерів котрих дорівнює одиниці:

$$C_1 = {}_{j=1}^{n}\oplus j_1 \cdot X_j \quad (6)$$

Аналогічно, кожна l-та компонента контрольного коду C_l, $l = 1,\ldots,k$ обчислюється як сума по модулю 2 лише тих символів блоку даних B, l-тий розряд номери яких дорівнює одиниці:

Компоненти контрольного коду $S_1,S_2,\ldots,S_k$ пропонується формуватиз використанням лінійної згортки.

$$C_l = {}_{j=1}^{n}\oplus j_l \cdot X_j \quad (7)$$

Лінійна згортка $\lambda(Y)$ m-розрядного коду $Y=\{y_1,y_2,\ldots,y_m\}$, де $\forall j \in \{1,2,\ldots,m\}$: $y_j \in \{0,1\}$ являє собою h-розрядний ($h=\lfloor \log_2 m \rfloor$) код $Z=\{z_1,z_2,\ldots,z_h\}$, причому, молодший розряд z_1 коду згортки формується як сума за модулем два всіх двійкових розрядів коду Y:

$$Z_1 = {}_{j=1}^{m}\oplus y_j \quad (8)$$

Другий розряд z_2 коду лінійної згортки Z формується як сума за модулем два всіх розрядів коду Y, номери яких мають одиницю в другому розряді, третій розряд z_3 коду згортки являє собою суму за модулем 2 всіх тих розрядів коду Y, номери яких мають одиницю в третьому розряді. Загалом, l-тий розряд z_l, $l \in \{2,3,\ldots,h\}$ коду лінійної згортки формується згідно наступної формули:

$$Z_1 = {}_{j=1}^{m}\oplus y_j \cdot ((j mod 2^{l-1})/2^{l-2}) \quad (9)$$

Наприклад, якщо $Y=\{1,0,1,1,0,1,0,1\}$, то $z_1 = 1\oplus0\oplus1\oplus1\oplus0\oplus1\oplus0\oplus1=1$, $z_2 = 0\oplus1\oplus1\oplus1 = 1$, $z_3 = 0\oplus1\oplus0\oplus1 = 0$, $Z=\lambda(Y) = \{1,1,0\}$.

Головна властивість коду лінійної згортки $\lambda(Y)$ полягає в тому, щоб при будь-якій зміні коду Y зазнає змін його лінійна згортка.

Кожну l-ту компоненту S_l контрольного коду, l=1,...,kпропонується обчислювати як суму за модулем 2 поліноміальних добутків лінійних згорток символів блоку даних B, l-тий розряд порядкових номерів в блоці яких дорівнює одиниці, на їх номери:

$$S_j = {}_{j=1}^{m}\oplus(\lambda(X_j) \otimes j) \cdot j_l) \quad (10)$$

Запропонований спосіб формування контрольного коду може бути ілюстровано наступним прикладом. Нехай передається блок з 7-ми символів, що нумеруються від одиниці до семи: B = {0101, 1011, 1100, 1001, 1111, 1101, 0011}. Для цього прикладу розрядність символів m=4, кількість символів в блоці $n=2^3$-1=7, розрядність порядкового номеру символів в блоці k=3.

Нульова компонента контрольного коду C_0 згідно (1) обчислюється як логічна сума: C_0 = 0101⊕1011⊕1100⊕1001⊕1111⊕1101⊕ 0011 = 1010. Компонента C_1 обчислюється згідно (2) як логічна сума тих символів, молодший розряд номерів яких дорівнює одиниці: $C_1 = X_1\oplus X_3\oplus X_5\oplus X_7$ = 0101⊕ 1100⊕1111⊕0011 = 0101. Аналогічно, компонента C_2 обчислюється як логічна сума тих символів другий розряд номеру яких дорівнює одиниці: $C_2 = X_2\oplus X_3\oplus X_6\oplus X_7$ = 1011⊕ 1100⊕1101⊕0011 = 1001. Код компоненти C_3 являє собою суму за модулем 2 символів, старший розряд номеру яких дорівнює одиниці: $C_3 = X_4\oplus X_5\oplus X_6\oplus X_7$ =

1001⊕ 1111⊕1101⊕0011 = 1000. Лінійні згортки символів блоку утворюють множину: { 00,01,10,10,00,11,10}. Відповідно з (2.15) код S_1 формується як сума поліноміальних добутків згорток символів, молодший розряд номеру яких дорівнює одиниці на самі ці номери, тобто: $S_1=\lambda(X_1)\otimes 1 \oplus\lambda(X_3)\otimes 3 \oplus\lambda(X_5)\otimes 5 \oplus\lambda(X_7)\otimes 7$ = 00⊗001 ⊕ 10⊗011 ⊕ 00⊗101 ⊕ 10⊗111 = 0110 ⊕ 1110 = 1001. Аналогічно, $S_2=\lambda(X_2)\otimes 2 \oplus\lambda(X_3)\otimes 3 \oplus\lambda(X_6)\otimes 6 \oplus\lambda(X_7)\otimes 7$ = 01⊗010 ⊕ 10⊗011 ⊕ 11⊗110 ⊕ 10⊗111 = 010 ⊕ 110 ⊕ 1010 ⊕ 1110 =0; $S_3=\lambda(X_4)\otimes 4 \oplus\lambda(X_5)\otimes 5 \oplus\lambda(X_6)\otimes 6 \oplus\lambda(X_7)\otimes 7$ = 10⊗100 ⊕ 00⊗101 ⊕ 11⊗110 ⊕ 10⊗111 = 1000 ⊕ 1010 ⊕ 1110 = 1100.

Всі описані вище компоненти контрольного коду обчислюються на стороні передавача і передаються разом з інформаційним блоком. Позначимо компоненти контрольного коду, обчислені на стороні передавача як $C_{S,0}$, $C_{S,1}$, $C_{S2},...,C_{S,k}$, $S_{S,0}$, $S_{S,1}$, $S_{S2},...,S_{S,k}$, а компоненти контрольного коду, обчислені приймачем за прийнятим блоком, як $C_{R,0}$, $C_{R,1}$, $C_{R2},...,C_{R,k}$, $S_{R,0}$, $S_{R,1}$, $S_{R2},...,S_{R,k}$. На стороні приймача обчислюються різниці контрольних кодів передавача та приймача:

$$\begin{gathered}\Delta_0 = C_{S,0}\oplus C_{R,j}\\ \forall j = 1,\dots,k\colon \Delta_j = C_{S,j}\oplus C_{R,j}\\ \delta_j = S_{S,0}\oplus S_{R,j}\end{gathered} \qquad (11)$$

Очевидно, що в разі, коли жоден з символів блоку не зазнав спотворення при передачі, всі різниці дорівнюють нулю: $\forall j$=0,...,k: Δ_j=0.

В разі спотворення лише одного q-го символу блоку, $q\in\{1,2,...,n\}$, на стороні приймача код цього символу $X_{R,q}$ відрізняється від відповідного символу на передавачеві $X_{S,q}$, так, що вектор спотворення $\Delta X_q = X_{R,q}\oplus X_{S,q}\neq 0$. Очевидно, що коди різниць в разі спотворення одного символу визначаються наступним чином: $\Delta_0 = \Delta X_q$, $\forall j\in\{1,2,..,k\}$: Δ_j= 0, якщо j-тий розряд q_j номеру q дорівнює нулю: q_j = 0 і Δ_j= ΔX_q, якщо j-тий розряд q_j номеру q дорівнює одиниці: q_j = 1. Таким чином, за кодами перших k+1 різниць $\Delta_0\Delta_1,...,\Delta_k$ однозначно визначаються код вектору спотворення символу: ΔX_q =Δ_0 та розряди $q_1,q_2,...,q_k$ номеру q спотвореного при передачі символу: $\forall j\in\{1,2,..,k\}$: q_j = 0 якщо Δ_j=0, q_j= 1, якщо $\Delta_j\neq 0$. Відповідно, виправлення спотвореного символу виконується у вигляді: X_q = $X_{R,q}\oplus\Delta X_q$. Сама ситуація виникнення однократної помилки визначається на умовою: $\forall j\in\{1,2,..,k\}$: $\Delta_j\in\{0,\Delta_0\}$.

В разі спотворення при передачі пари символів, що мають в блоці порядкові номери q та p, причому p <q, постає задача визначення як самих позицій q і p, так і векторів спотворень обох символів ΔX_q та ΔX_p.

Якщо вектори спотворень обох пошкоджених при передачі символів відрізняються, тобто $\Delta X_q\neq\Delta X_p$, ця задача вирішується доволі просто. Код різниці нульових компонентів контрольного коду в цій ситуації дорівнює в цій ситуації значення: Δ_0 = $\Delta X_p\oplus\Delta X_q\neq$ 0. Кожна j-та різниця, $j\in\{1,2,..,k\}$, наступних k компонентів контрольного коду можуть приймати лише 4 значення:

– Δj = 0, якщо j-ті розряди qj і pj номерів q та p дорівнюють нулю: qj = pj = 0;

– Δj = Δ0 = ΔXp ⊕ΔXq, якщо j-ті розряди qj і pj номерів q та p дорівнюють одиниці: qj = pj = 1;

– Δj = ΔXp якщо j-тий розряд qj номеру q дорівнює нулю, а j-тий розряд pj номеру p дорівнює одиниці: qj = 0, pj = 1;

– Δj = ΔXq якщо j-тий розряд qj номеру q дорівнює одиниці, а j-тий розряд pj номеру p дорівнює нулю: qj = 0, pj = 1.

Процедура відновлення номерів пошкоджених символів. Відповідно, процедура відновлення номерів *q* та *p* пошкоджених символів представлена у вигляді наступної послідовності дій:

1. Встановити індекс j поточного розряду номеру в k: j = k, прапорець f виявлення першого розряду, в якому різняться коди q та p встановлюється в нуль: f=0.

2. Якщо Δj = 0, то j-ті розряди p та q встановлюються в нулі, pj = 0, qj = 0. Перехід на п.6.

3. Якщо Δj = Δ0, то j-ті розряди p та q встановлюються в одиниці, pj = 1, qj = 1. Перехід на п.6.

4. Якщо Δj ≠Δ0, Δj ≠ 0 і f = 0, то j-тий розряд q встановлюється в одиницю, j-тий розряд p встановлюється в нуль, qj = 1, pj = 0, ΔXq = Δj і f=1 Перехід на п.6.

5. Якщо Δj ≠Δ0,Δj ≠ 0 і f = 1, то, якщо Δj = ΔXq , j-тий розряд q встановлюється в одиницю, j-тий розряд p встановлюється в нуль, qj = 1, pj = 0, інакше j-тий розряд p встановлюється в одиницю, а j-тий розряд q встановлюється в нуль, pj = 1, qj = 1. Перехід на п.6.

6. Декремент індексу поточного розряду номеру j = j – 1. Якщо j > 0, то повернення на п.2.

7. Визначення вектору ΔXp = Δ0 $\oplus$ ΔXq. Корекція пошкоджених символів: Xp= XR,p⊕ΔXp; Xq= XR,q⊕ΔXq. Кінець.

Конструктивність розробленої процедури визначається тим, що в силу того, що номери *q* та *p* обов'язково розрізняються, причому *q*>*p*, то в старшому розряді, в якому ці коди відмінні, відповідний розряд *q* дорівнює нулю, а однойменний розряд *p* дорівнює одиниці. Це значить, що обов'язково виконується п.4 наведеної вище процедури, який визначення вектор спотворення *q*-го символу - ΔX_q.

Наведена процедура корекції двох спотворених при передачі символів блоку за умови, що вектори їх спотворення відрізняються може бути ілюстрована наступним чином. Якщо припустити, що в рамках наведеного вище прикладу спотворені 4-й та 6-й символи (*q*=6, *p*=4), причому вектор першого із спотворених символів ΔX_p= 1100, а другого ΔX_q = 1001, то прийнятий блок B={ 0101, 1011, 1100, **0101**, 1111, **0100**, 0011}. Значення обчислених на стороні приймача перших *k*+1 компонентів контрольного коду становлять:

$C_{R,0}$= 0101⊕1011⊕1100⊕**0101**⊕1111⊕**0100**⊕0011=1111,

$C_{R,1}$=0101⊕ 1100⊕1111⊕0011 = 0101;

$C_{R,2}$=1011⊕1100⊕**0100**⊕0011 = 0000;

$C_{R,3}$= **0101**⊕ 1111⊕**0100**⊕0011 = 1101.

Відповідно, Δ_0= 0101; Δ_1=0000; Δ_2=1001; Δ_3=0101.

Згідно п.1 наведеної вище процедури, індекс *j*=3, *f*=0. Так як Δ_3=0101=Δ_0, то згідно з п. 3 процедури старші розряди номерів пошкоджених символів дорівнюють одиниці: p_3=1 і q_3=1. При *j*=2 виконуються умови п.4 процедури: $\Delta_2 \neq \Delta_0$, $\Delta_2 \neq 0$ і *f* = 0: відповідно 2-тий розряд *q* встановлюється в одиницю, 2-тий

розряд p встановлюється в нуль, $q_j = 1$, $p_j = 0$, $\Delta X_q = \Delta_2 = 1001$ і f=1. При j=1 виконується умова п.2 процедури: Δ_1=0, відповідно згідно з п.2 p_1=1 і q_1=1. Таким чином, позиції пошкоджених при передачі визначено: q=6, p=4. Останнім п.7 визначається вектор спотворення 4-го символу: $\Delta X_4 = \Delta_0 \oplus \Delta X_6 = 0101 \otimes 1001 = 1100$ та відновлюються спотворені символи: $X_4 = X_{R,4} \oplus \Delta X_4 = 0101 \oplus 1100 = 1001$, $X_6 = X_{R,6} \oplus \Delta X_6 = 0100 \oplus 1001 = 1101$.

Якщо вектори спотворень обох пошкоджених символів однакові, тобто $\Delta X_p = \Delta X_q$, то Δ_0=0 і для корекції використовуються різниці $\delta_1, \delta_2, ..., \delta_k$ компонентів $S_1, S_2, ..., S_k$. Процедура корекція пари символів з порядковими номерами p та q, причому $q>p$ полягає в виконанні наступної послідовності дій:

1. Встановити індекс j поточного розряду номеру в k: $j = k$, прапорець f виявлення першого розряду, в якому різняться коди q та p встановлюється в нуль: f=0.

2. Якщо $\delta j = 0$, то j-ті розряди p та q встановлюються в нулі, pj = 0, qj = 0. Перехід на п.6.

3. Якщо $\delta j \neq 0$, а $\Delta j = 0$, то j-ті розряди p та q встановлюються в одиниці, pj = 1, qj = 1. Перехід на п.6.

4. Якщо $\delta j \neq 0$, $\Delta j \neq 0$ і f = 0, то j-тий розряд q встановлюється в одиницю, j-тий розряд p встановлюється в нуль, qj = 1, pj = 0, $\Delta Xq = \Delta j$, збереження значення δj в змінній d: $d = \delta j$, установка в одиницю прапорця f=1 Перехід на п.6.

5. Якщо $\delta j \neq 0$, $\Delta j \neq 0$ і f = 1, то, якщо $\delta j = d$, j-тий розряд q встановлюється в одиницю, j-тий розряд p встановлюється в нуль, qj = 1, pj = 0, інакше j-тий розряд p встановлюється в одиницю, а j-тий розряд q встановлюється в нуль, pj = 1, qj = 0. Перехід на п.6.

6. Декремент індексу поточного розряду номеру j = j – 1. Якщо j > 0, то повернення на п.2.

7. Визначення вектору ΔXp = ΔXq. Корекція пошкоджених символів: Xp= XR,p⊕ΔXp; Xq= XR,q⊕ΔXq. Кінець.

Процедура корекції двох однаково спотворених символів блоку може бути ілюстрована наступним чином. Якщо припустити, що в рамках наведеного вище прикладу спотворені 3-й та 6-й символи (q=6, p=3), причому вектори спотворення обох символів обох символів дорівнюють $\Delta X_p = \Delta X_q$ =1001, то прийнятий блок має вигляд B={ 0101,1011,**0101**,1001,1111, **0100**, 0011}. Значення обчислених на стороні приймача $2 \cdot k$+1 компонентів контрольного коду становлять: $C_{R,0}$= 1010, $C_{R,1}$= 1101; $C_{R,2}$=1001; $C_{R,3}$= 0001, $S_{R,1}$=1111, $S_{R,2}$= 1010, $S_{R,3}$= 0. Відповідно, Δ_0= 0; Δ_1=1001; Δ_2=0; Δ_3=1001, δ_1 = 110; δ_2 =1010; δ_3 = 1100.

Згідно п.1 наведеної вище процедури, індекс j=3, f=0. Так як $\delta_3 = 1100 \neq 0$, $\Delta_3 = 1001 \neq 0$ і f=0, то згідно з п. 4 процедури старший розряд більшого номеру q дорівнює одиниці: q_3=1, старший розряд меншого номеру p дорівнює нулю: p_3=1; в змінній d фіксується значення δ_3 = 1100, f=1. При j=2 виконуються умови п.3 процедури: Δ_2=0, $\delta_2 \neq 0$: відповідно p_2=1 і q_2=1. При j=1 виконується умова п.5 процедури: $\delta_1 = 110 \neq 0$, $\Delta_1 = 1001 \neq 0$ і $f = 1$, при цьому $\delta_1 = 110 \neq d$=1100, відповідно молодший розряд p встановлюється в одиницю, а молодший розряд q встановлюється в нуль, $p_j = 1$, q_j=0. Таким чином, позиції пошкоджених при передачі визначено: q=6, p=3. Останнім п.7 визначається вектор спотворення

3-го символу: $\Delta X_3 = \Delta X_6 = 1001$ та відновлюються спотворені символи: $X_3 = X_{R,3} \oplus \Delta X_3 = 0101 \oplus 1001 = 1100$, $X_6 = X_{R,6} \oplus \Delta X_6 = 0100 \oplus 1001 = 1101$.

Загальна довжина l контрольного коду C та S обчислюється як сума розрядів всіх k її компонентів. Очевидно, що розрядність перших k компонентів коду C дорівнює розрядності символу – m, компонентів коду S – $\log_2 k + \log_2 m$ -1. Таким чином, загальна довжина L контрольного коду визначається формулою:

$$L = k \cdot (m + \log_2 m + k - 1) + m. \quad (12)$$

Наприклад, для типових для практики значень розрядності символу m = 8 (QAM-256) і довжини блоку n =1024, довжина контрольного коду становить L =208, або 26 символів, що становить 2.5% об'єму інформаційного блоку.

З виразу (12) слідує, що кількість контрольних символів, що передаються додатково і використовуються для корекції спотворених при передачі символів для запропонованому методі помітно перевищує аналогічний показник для циклічних кодів і, зокрема, коду Ріда-Соломона. Проте для сучасних швидкостей передачі даних, яка вимірюється Мегабайтами за секунду, зменшення на 1-2 порядки часу корекції помилок вагоміший за передачу декількох додаткових символів.

Висновки. В результаті проведених досліджень запропоновано метод прискореної корекції помилок в каналах зі спектральною модуляцією, кратність не більше 2-х.

За рахунок використання більшої кількості контрольних розрядів, в $4 \cdot n$ разів зменшено обчислювальну складність корекції в порівнянні з кодами Ріда-Соломона, що дозволяє на порядки прискорити процес корекції та спростити апаратну реалізацію.

Метод орієнтовано для швидкісних каналів передачі даних між компонентами розподілених систем комп'ютерного управління, що працюють у режимі реального часу.

Список літератури

1. Reed I.S. Polynomial codes over certain finite fields. Journal of the Society for Industrial and Applied Mathematics.-1960.- № 8(2). - pp. 300–304.

2. Wickers S.B.. Reed-Solomon Codes and Their Applications. IEEE Press. Piscataway, New Jersey.-1983.- p.433

3. Ирвин Дж. Передача данных в сетях: инженерный подход. СПб.: БХВ-Петербург, 2002.- 448 с.

4. Марковський О.П. Метод виправлення трьохкратних помилок передачі даних в двійкових симетричних каналах Вісник Національного технічного університету України "КПІ" Інформатика, управління та обчислювальна техніка, – Київ: ВЕК+ – 2014. – № 60. - С.33-40

5. Морелос-Сарагоса Р. Искусство помехоустойчивого кодирования. Методы, алгоритмы, применение, 2005.- 319 с.

6. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. Издательский дом "Вильямс".- 2004.- 1104 с.

ДОВІДКА ПРО АВТОРІВ

Марковський Олександр Петрович – доцент, кафедра обчислювальної техніки,Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Markovskyi O. P. – associate professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv PolytechnicInstitute”.

Білашевська Анастасія Володимирівна – студентка, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Bilashevska Anastasia Volodymyrivna - student, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: harasymovych.hv@gmail.com

Markovskyi O. P.,
Bilashevska A. V.

METHOD FOR MULTIPLE ERRORS CORRECTION IN SPECTRAL MODULATION DATA TRANSMISSION CHANNEL

Topicality of the research.

The current dynamic expansion of the use of mobile telecommunications has as a consequence the multiple increase of the intensity of high-frequencies in the environment. This correspondingly increases the level of radio interference which produces an increase in the errors within the digital information transmission channels.

This specifies the urgency and the practical importance of the development of new methods and means to increase the reliability of the errors control in the telecommunications network systems with spectral modulation.

Problemstatement.

The urgent task is to develop means to increase the effectiveness of error detection in digital information channels with effective spectral modulation, which channels are constantly increasing in the telecommunication networks.

Analysis of recent researches and publications.

The basic criteria for effective error control in the computer networks are the reliability of error detection in multiple cases and the response time for the implementation of the calculations connected with error detection during the data transmission.

In the computer networks the information is most frequently transferred by blocks and as a result the CRC (Cyclic Redundancy Code) and CS (Check Sum) methods are widely used for error control.

Uninvestigated parts of general matters defining.

The essential drawback of the CRC error detection method is its low speed implementation of the remainders calculated on a per bit basis. For eliminating this

deficiency in recent years a number of tabular methods have been proposed for the remainder calculation on per bytes and per words basis. However, the complexity of realizing the tabular methods is sufficiently high.

In particular, they do not ensure guaranteed detection of signal channels of the most common after the single errors, which are the dual and triple errors. Furthermore, the use of CRC is connected with the problem of effective computational realization of the operations of control at the rate of information transmission, since the procedure of the calculation of the error syndrome in CRC is done principally in a sequential manner.

The research objective.

The scope of the present research is the development of the means to increase the reliability of the repeated errors detection during transmission of the symbols in the digital information channels with the use of effective spectral modulation. For this it is necessary to ensure the possibility of the high-speed realization of the calculations, which are connected with the data integrity during transmission, with the aim these calculations to be implemented in parallel.

Main body.

The paper is dedicated to the efficiency improving of double error correction of data transmission in channels with spectral modulation by simplifying and accelerating the computations associated with error correction.

The paper presents a theoretically grounded and investigated approach of double errors correction in channel with spectral modulation based on positional correcting sums that allows to determine the positions of distorted symbols and their distortion vectors without enumeration through all the symbols of block.

Conclusions.

For data channels with spectral modulation a based on the proposed approach method of accelerated errors correction has been developed, which allows to determine the positions of distorted symbols and their distortion vectors without enumeration through all the symbols of block.

Key words: multiple data transmission errors, error detecting, error correction, coding errors, weighted checksums, cyclic and linear redundancy correcting codes.

УДК 004.75

Міщенко Л. Д.,
Виноградов Ю. М.

МЕТОД ПАРАЛЕЛЬНОГО ОБЧИСЛЕННЯ МОДУЛЯРНОЇ ЕКСПОНЕНТИ З ВИКОРИСТАННЯМ ПЕРЕДОБЧИСЛЕНЬ

METHOD FOR PARALLEL MODULAR EXPONENTIATION BY USING PRECOMPUTATION

В статті запропонована організація паралельного виконання модулярного експоненціювання. Доведено, що на рівні операцій модулярного множення трьопотоковий паралелізм є найбільш доцільною формою паралельного виконання модулярного експоненціювання. Наведено математичне обґрунтування запропоновано підходу. Запропонована процедура паралельного обчислення модулярної експоненти детально викладена та ілюстрована чисельним прикладом. Виконано порівняльний аналіз продуктивності запропонованого методу обчислення модулярної експоненти Теоретично та експериментально доведено, що запропонований метод забезпечує прискорення обчислення модулярної експоненти приблизно втроє.

Ключові слова: комп'ютерна арифметика, паралельні обчислення, модулярне множення, модулярне експоненціювання, мережові протоколи захисту даних.

In article the organization of modular exponentiation parallel executing are presented. It has been shown that on modular multiplication level the three stream parallelism is best suited for parallel modular exponent calculation. The mathematical background of the proposed approach is presented. The proposed procedure of parallel modular exponent calculation are described in details and illustrated by numerical example. Performed comparative analysis of the proposed methods of modular exponent calculation has been executed. By the theoretical and experimental ways it is proved that the proposed method provides an acceleration of modular exponentiation by approximately three times.

Key words: computer arithmetic, parallel calculation, modular multiplication, modular exponentiation, data security protocols.

Актуальність теми дослідження. З плином часузростає важливість забезпечення високої продуктивності реалізації термінальними пристроями комп'ютерних систем моніторингу та управління існуючих мережевих протоколів. Найбільш критичними з точки зору обчислювальної реалізації протоколів мережевого обміну є операції модулярної арифметики, що виконуються над числами великої розрядності, яка значно перевищує розрядність процесорів. На теперішній час, для досягнення прийнятного для більшості застосувань рівня захищеності, необхідна довжина чисел становить 1024-2048 розрядів з перспективою її зростання в найближчі роки до 4096.

Тенденція зростання пропускної здатності каналів передачі даних комп'ютерних мереж вимагає адекватного зростання швидкості реалізації протоколів захисту інформації як на комп'ютерах загального призначення, так і на малорозрядних мікроконтролерах.

Наведені фактори визначають розробку нових підходів до прискорення програмної реалізації операцій модулярної арифметики при їх застосування в протоколах захисту інформації, як важливу та актуальну проблему, від вирішення якої значною мірою залежить ефективність локальних та глобальних комп'ютерних мереж.

Постановка проблеми. Базовою обчислювальною операцією широкого кола мережевих протоколів захисту інформації є модулярне експоненціювання, тобто обчислення $A^E \bmod M$, де A, E і M – числа розрядністю n, значно більшою за розрядність процесора.

Процес модулярного експоненціювання зводиться до послідовного виконання n циклів, у кожному із яких виконується операція модулярного піднесення до квадрату результату операції попереднього циклу і, залежно від поточного біта степені E, здійснюється операція модулярного множення. Залежно від порядку, в якому організована обробка розрядів коду експоненти E можна існують два базових алгоритми експоненціювання [1]: зі старших та молодших розрядів коду експоненти.

Алгоритм обробки розрядів експоненти зі старших розрядів в нотаціях мови C++ алгоритм має такий вигляд:

1. $R = 1$.
   2. for (j=n-1; j>=0; j --)
   {
      2.1. $R = R \cdot R \bmod M$
      2.2. if (e^j == 1)
$R = R \cdot A \bmod M$
   }
3. Результат: R.

Прицьому, підчас кожної ітерації циклу виконується модулярне піднесення числа вквадраті множення на постійне число, рівне A, що створює потенційні передумови для підвищення швидкості множення. Недоліком є те, що всі операції виконуються строго послідовно й лежать на критичному шляху [2].

Алгоритм, який передбачає обробку розрядів степені E починаючи із молодших розрядів в нотаціях мови C++ має вигляд:

1. $R = 1$, $Q = 1$.
2. for (j=0; j <n; j ++)
 {
2.1. $R = R \cdot R \bmod M$
2.2. if (e_j = = 1)
$Q = Q \cdot R$
}
3. Результат: Q.

Аналіз останніх досліджень і публікацій. Аналіз обох базових алгоритмів показує, що час їх реалізації визначається сумою: $n \cdot t_{sq} + 0.5 \cdot n \cdot t_m$, де t_{sq} - час виконання операції модулярного піднесення до квадрату, а t_m – час виконання операції модулярного множення [3] .

В рамках другого з розглянутих алгоритмів модулярного експоненціювання існує потенційна можливість розпаралелювання обчислень. В роботі [4]

варіант такого розпаралелювання. Показано, що на рівні операцій модулярного множення максимальний рівень паралелізму не перевищує 2-х. Відповідно, прискорення обчислювання модулярної експоненти досягається за рахунок використання двох процесорів, один з яких реалізує потік операцій модулярного піднесення до квадрату, а інший – модулярного множення.

Проведений аналіз обчислювальних процедур модулярного експоненціювання показав, прискорення їх виконання може бути досягнуто за рахунок розпаралелювання на різних рівнях. Більшість робіт [5-7], присвячених вирішенню проблеми прискорення операції модулярного експоненціювання орієнтовані на рівень процесорних операцій, на які розкладаються операції модулярного піднесення до квадрату і модулярного множення. Зокрема, при експоненціювання 2048-розрядних чисел на 32-розрядних процесорів виконується 128 операцій процесорного множення, які можуть виконуватися паралельно [6] .

У свою чергу, час виконання модулярного множення визначається двома складовими: часом, необхідним для реалізації власне множення і часом, який витрачається на модулярну редукцію, тобто залишку від ділення результату множення на модуль М. У класичному множенні модулярна редукція реалізується з використанням операції ділення і, відповідно, друга складова відіграє значну роль. Значна ефективність обчислювальної реалізації модулярного множення досягається при використанні алгоритму Монтгомері [8], в якому модулярна редукція зводиться до зсуву на k розрядів. Інший підхід до зменшення часу редукції запропоновано в роботі [3]. Цей підхід базується на тому, що на практиці, модуль, що є частиною відкритого ключа, практично не змінюється. Це надає змогу виділити операції, що залежать від модуля, обчислити їх результати один раз і використовувати при кожному модулярному експоненціюванні.

На сьогодняшній день розроблено ряд методів прискореної реалізації модулярного експоненціювання [8,9], які реалізують можливості розпаралелювання прискорення обчислення модулярної експоненти на рівні процесорних операцій.

Виділення недосліджених частин загальної проблеми. На основі досліджених літературних джерел можна зробити наступні висновки. Основним резервом підвищення швидкості обчислювальної реалізації базової операції мережевих протоколів захисту інформації є організація паралельної обробки. Проте цей підхід не може бути використаний для прискорення модулярного експоненціювання на малопотужних мікроконтролерах – термінальних пристроях широкого кола мереж, що використовуються на практиці. Таким чином, існуючі методи прискорення обчислення модулярної експоненти не вирішують цю проблему для вельми широкого класу обчислювальних пристроїв, які мають підтримувати протоколи мережевого захисту даних.

Постановка задачі. Мета досліджень полягає в прискоренні виконання критичної для протоколів мережового захисту інформації операції модулярного експоненціювання.

Метод паралельного обчислення модулярної експоненти з використанням передобчислень. Для досягнення поставленої мети, пропонується n-розрядний код експоненти $E = \{e_1,e_2,\ldots,e_n\}$, $\forall j \in \{0,1\}$, розділити на sm-розрядних фрагментів: $f=\{e_1,e_2,\ldots,e_m\},\ldots,f_s = \{e_{n-m}, \ldots, e_n\}$, очевидно, що $s=n/m$.

До безпосередніх обрахунків модулярної експоненти пропонується виконати $2^m - 2$ передобчислень значень $A^2 \bmod M$, $A^3 \bmod M$, …, $A^c \bmod M$, де $c = 2^{m-2}$ Передобчислення виконуються запропонованим деревовидним алгоритмом по заданому значенню A. На першому кроці обчислюється значення $A^2 \bmod M = A \cdot A \bmod M$. На наступному кроці паралельно на двох процесорах обраховуються значення $A^3 \bmod M = A^2 \bmod M \cdot A$ та $A^4 \bmod M = A^2 \bmod M \cdot A^2 \bmod M$. За третій крок одночасно обчислюються чотири значення: $A^5 \bmod M = A^3 \bmod M \cdot A^2 \bmod M$, $A^6 \bmod M = A^4 \bmod M \cdot A^2 \bmod M$, $A^7 \bmod M = A^4 \bmod M \cdot A^3 \bmod M$, $A^8 \bmod M = A^4 \bmod M \cdot A^4 \bmod M$. Таким чином, усі передобчислення виконуються за m кроків на 2^{m-1} процесорах. Результати передобчислень зберігаються в таблиці: $T[0] = 1$, $T[1] = A$, $T[2] = A^2 \bmod M$, … , $T[2^m - 1] = A^c \bmod M$.

Таким чином, теоретично, час передобчислень $T_{\text{тп}}$ визначається добутком $T_{\text{тп}} = m \cdot t_{\text{m}}$ де t_{m} – час виконання операції модулярного множення.

Безпосереднє обчислення модулярної експоненти пропонується виконувати на s-1 процесорах, кожен з яких оброблює відповідний фрагмент коду експоненти.

На j-тому процесорі, $j \in \{1,2,\ldots,s-1\}$, що оброблює j-ий фрагмент f_jкоду експоненти E пропонується виконувати:

- піднесення коду $T[f_j]$ до степені $2^{n-j.m}$;
- формування результату r_j обробки j-того фрагменту f_jшляхом модулярного множення отриманого коду на результат r_{j+1} обробки наступного фрагменту f_{j+1} коду експоненти: $r_j = r_{j+1} \cdot T[f_j]^{2^{n-j.m}} \bmod M$. Для (s-1)-го фрагменту $r_{j+1} = r_s = T[f_s]$.

Запропонований метод розпаралелювання модулярної експоненти з виконанням передобчислень може бути показано наступним прикладом.

Нехай потрібно обчислити $48^{2542} \bmod 57 = 6$, тобто $A = 48$, $E = 2542$, $M = 57$.

До початку безпосередніх обрахунків виконуються передобчислення. Так як $E = 2542_{10} = 100111101110_2$, $n = 12$, запропоноване $m = 3$, тоді $s = n/m = 4$. Таблиця передобчислень заповнюється наступним чином.

Таблиця 1.

Таблиця передобчислень
для прикладу обчислення $48^{2542} \bmod 57 = 6$

f_j	$A^f \bmod M$	$T[f_j]$
2	$48^2 \bmod 57 = 24$	24
3	$48^3 \bmod 57 = 12$	12
4	$48^4 \bmod 57 = 6$	6
5	$48^5 \bmod 57 = 3$	3
6	$48^6 \bmod 57 = 30$	30
7	$48^7 \bmod 57 = 15$	15

Позначимо інтервал часу модулярного піднесення до квадрату як τ, тоді інтервал часу модулярного множення дорівнює $t_m = 2 \cdot \tau$.

Таблиця 2.

Часова діаграма обчислення прикладу $48^{2542} \bmod 57 = 6$.

τ	Процесори		
	1	**2**	**3**
1	$3^2 \bmod 57 = 9$	$15^2 \bmod 57 = 54$	$6^2 \bmod 57 = 36$
2	$9^2 \bmod 57 = 24$	$54^2 \bmod 57 = 9$	$36^2 \bmod 57 = 42$
3	$24^2 \bmod 57 = 6$	$9^2 \bmod 57 = 24$	$42^2 \bmod 57 = 54$
4	$30 \cdot 6 \bmod 57 = 9$	$24^2 \bmod 57 = 6$	$54^2 \bmod 57 = 9$
5		$6^2 \bmod 57 = 36$	$9^2 \bmod 57 = 24$
6		$36^2 \bmod 57 = 42$	$24^2 \bmod 57 = 6$
7		$9 \cdot 42 \bmod 57 = 36$	$6^2 \bmod 57 = 36$
8			$36^2 \bmod 57 = 42$
9			$42^2 \bmod 57 = 54$
10			$36 \cdot 54 \bmod 57 = 6$
11			

Цілком очевидно, що в запропонованому способі паралельного обчислення модулярної експоненти час експоненціювання визначається часом обробки першого фрагменту коду експоненти $T_1 = (n\text{-}m) \cdot t_q + t_m$, t_q – час модулярного піднесення до квадрату. Приймаючи до уваги, що $t_q \approx t_m/2$, сумарний час T_e обчислення модулярної експоненти з врахуванням часу передобчислень, визначається формулою:

$$T_e = m \cdot t_m + (n-m) \cdot t_q + t_m \approx n \cdot t_q + m \cdot t_q$$

Аналіз ефективності. Основною перевагою запропонованого способу модулярного експоненціювання є прискорення обчислення шляхом розпаралелювання. Тому, ефективніть запропонованого способу доцільно оцінювати коефіцієнтом прискорення, який розраховується відношенням часу виконання модулярного експоненціювання класичним способом – T_0, до часу виконання запропонованого способу паралельного обчислення модулярної експоненти – T:

$$\beta = \frac{T_0}{T}.$$

Для випадку, коли кількість процесорів не обмежена, коефіцієнт прискорення обраховується за формулою:

$$\frac{3 \cdot n \cdot t_q}{n \cdot t_q + m \cdot t_q} = \frac{3 \cdot n}{n + m}$$

Враховуючи, що$m \ll n$, реальне значення коефіцієнту прискорення при відсутності обмежень на кількість процесорів, дорівнює 3: $\beta \approx 3$.

Висновки. В результаті проведених досліджень, направлених на пошук шляхів прискорення виконання важливої для реалізації протоколів захисту інформації операції модулярного експоненціювання над числами, розрядність яких значно перевищує розрядність процесора можна зробити такі висновки.

Проведений аналіз обчислювальних процедур модулярного експоненціювання показав, прискорення їх виконання може бути досягнуто за рахунок виконання передобчислень.

Для практичної реалізації цієї можливості запропонована організація прискореного обчислення модулярної експоненти на багатоядерному процесорі. Теоретично та експериментально доведено, що розроблена організація забезпечує практично трьократне прискорення виконання операції модулярного експоненціювання для розрядностей 2048 і 4096.

Більш значне прискорення виконання модулярного експоненціювання може бути досягнуто при переході на рівень операцій процесорного множення.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Самофалов К.Г. Ускоренная реализация модулярного экспоненцирования на малоразрядных микропроцессорах и встроенных микроконтроллерах / К. Г. Самофалов, Рамзи Анвар Салиба Сунна, Д. Ю. // Проблеми інформатизації та управління. Збірник наукових праць: Випуск 4(15).-К.,НАУ, 2005.- С.144-153.

2. Can Xiang. Verifiable and Secure Outsourcing Schemes of Modular Exponentiations Using One Untrusted Cloud Server and Their Application // IACR Cryptology ePrint Archive 2014: PP.500 .- https://eprint.iacr.org/2014/500.pdf

3. Markovskyi O.P. Secure Modular Exponentiation in Cloud Systems./ Oleksandr P. Markovskyi, Nikolaos Bardis, Nikolaos Doukas, Sergej Kirilenko // Proceedings of The Congress on Information Technology, Computational and Experimental Physics (CITCEP 2015), 18-20 December 2015, Krakow, Poland, C. 266-269.

4. Брей Б. Микропроцессоры Intel. Архитектура, программирование и интерфейсы. Пер.с англ.- СПб:БХВ-Петербург.-2014.- С.1328.

5. Марковський О.П. Спосіб прискореного обчислення модулярної експоненти / О.П.Марковський, Л.Д. Міщенко // Прикладна математика та комп'ютинг ПМК-2017. Збірник тез доповідей 9-ї наук. конференції магістрантів та аспірантів, Київ, 19-21 квіт.2017.- К.:Просвіта,2017 – С.200-203.

ДОВІДКА ПРО АВТОРІВ

Виноградов Юрій Миколайович – старший викладач, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Vinogradov Yurii – associate professor, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute".

Міщенко Людмила Дмитрівна – студентка, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Mishchenko Liudmyla - student, Department of Computer Engineering, National Technical University of Ukraine "Igor Sikorsky Kyiv PolytechnicInstitute".

E-mail: liudamishchenko@gmail.com

Mishchenko Liudmyla,
Vinohradov Yurii

METHOD FOR PARALLEL MODULAR EXPONENTIATION BY USING PRECOMPUTATION

Topic relevance. Development new ways of acceleration programmatic implementation of modular arithmetics. These can be used in information encryption protocols and can affect both local and global computer networks.

Problem setting. Computation of the modular exponent, which is the spreadest operation in communication protocols. This means calculations of A^E mod M where A, E and M are n-digit numbers, much greater than CPU bit depht.

Actual researches and issues analysis. The executed analysis of modular exponent calculation procedures showed, that acceleration of execution can be achieved by using parallelization on different levels. The majority of works in this topic are oriented on the level of CPU operations, which are the components of modular squaring and multiplication.

Uninvestigated parts of general matters defining. Relying on the investigated references, the next conclusions can be done. The main source of acceleration computation speed of modular exponent is organization parallel processing. But this solution cannot be implemented on slow microcontrollers. Thereby existing methods do not solve the problem for huge amount of devices, which require communication protocols of information encryption.

Target setting. The investigation endpoint is acceleration modular exponent calculation operation, which is critical for communication protocols of information encryption.

The statement of basic materials. The proposed solution is to split n-digit exponent code into s m-digit components. Computation of modular exponent is proposed to calculate using s-1 CPUs. Each is responsive for calculation of corresponding exponent fragment.

Conclusion. The executed analysis of calculation showed that acceleration can be achieved using precalculations. More significant acceleration of modular exponent calculation can be achieved with transition to the level of CPU multiplication operations.

Key words: computer arithmetic, parallel calculation, modular multiplication, modular exponentiation, data security protocols.

УДК: 621.391

Елена Жданова,
Алла Коган,
Юрий Кулаков,
Майя Сперкач

О РАЗБИЕНИИ ГРАФА НА ПОДГРАФЫ

ON THE PARTITION OF THE GRAPH INTO SUBGRAPHS

В статье рассматривается задача разбиения графа на множество подграфов с определением разреза графа и граничных вершин. Определены критерии оценки качества подразрезов и наложены некоторые ограничения (граничные условия), которые должны быть учтены при нахождении разбиения.

Ключевые слова: граф, разрез графа, подграф, разбиение графа.

Рис.: 0. Табл.: 0. Бібл.: 3

In this paper we consider the problem of partitioning a graph into a set of subgraphs with the definition of a cut of the graph and border nodes. The criteria for assessing the quality of sub-divisions are defined and some restrictions (boundary conditions) imposed, which must be taken into account when finding the subdivision.

Keywords: graph, section of graph, subgraph, partition of graph.

Fig.: 0. Tabl.: 0. Bibl.: 3

Актуальность темы исследования. На сегодняшний день самыми актуальными компьютерными сетями являются беспроводные, так как дают свободу перемещений.

Вследствие увеличения объема мобильного трафика в сети, который растет в геометрической прогрессии, контролировать потоки информации становится сложнее. В свою очередь, для каждого типа трафика выдвигаются свои требования к качеству обслуживания (QoS) [1], такие как требования к пропускной способности, задержки, надежности и относительной загруженности узла.

Постановка проблемы. Возникает необходимость в разбиении беспроводной компьютерной сети на зоны маршрутизации для уменьшения времени конструирования трафика и формирования множества путей для передачи информации.

Анализ последних публикаций и исследований. В работе [1] предложен метод многоуровневого и сбалансированного распределения косвенных графов (Multi-level spectral graph partitioning method (MSGP). Используя собственные векторы матрицы Лапласа графов, метод имеет спектральный подход, который имеет преимущество перед методами с глобальной способностью к делению. Бисекция, которая является спектральным методом, может разделить граф с помощью вектора Фидлера, а рекурсивная версия метода может быть разделена на несколько уровней.

В работе [2] предлагается подход многоуровневого разбиения сети на зоны маршрутизации, и выделяют три фазы разбиения графа: укрупнение, начальное разбиение и уточнение. Также предложена реализация алгоритма Metis на основе OpenMP и сравнение его с MPI на трех разных многоядерных архитектурах. Многопоточное разделение обеспечивает более чем вдвое ускоре-

ние по сравнению с другими разбиениями, но также использует значительно меньше памяти.

В работе [3] расширяется задача разбиения графа на подграфы путем добавления произвольного количества ограничений балансировки. В работе, каждому из вершин присваивается вектор весов, и цель состоит в том, чтобы создать k-образное разбиение таким образом, чтобы разбиение удовлетворяло уравнению балансировки, связанному с каждым весом, при попытке минимизировать граничные узлы.

Выделение неисследованных частей общей проблемы. В рассмотренных способах разбиения графа на подграфы не определены критерии оценки качества подразрезов сети, которое непосредственно связано с безопасностью всей сети.

Постановка задачи. Традиционные алгоритмы разбиения графов вычисляют k-образное разбиение графа таким образом, что количество ребер (разрез), минимизируется, и каждый разрез имеет равное количество граничных вершин. Задача минимизации граничного разреза может рассматриваться как цель, и требование о том, чтобы разрезы имели одинаковый размер, можно рассматривать как ограничение.

Рассмотрим неориентированный граф $G=(V,E,W,C,P)$,

где V – не пустое конечное множество вершин, $V=\{1,...,n\}$, n - порядок графа; $E=\{(i,j)\in V\times V\}$ – множество ребер; $W:V\to R$ – весовая функция, которая каждой вершине ставит в соответствие действительное число ($w_i>0$–вес вершины $i\in V$); $C:E\to R$ – весовая функция, которая каждому ребру ставит в соответствие действительное число ($c_{ij}>0$–вес ребра $(i,j)\in E$); $P:E\to(0,1]$– весовая функция, которая каждому ставит в соответствие действительное число из полуинтервала (0,1] ($0<p_{ij}\le 1$–вероятность безотказной работы ребра $(i,j)\in E$).

Разбиение графа на подграфы означает представление исходного графа $G=(V,E)$ в виде множества подмножеств вершин (блоков) $V_1, V_2, ..., V_r$ ($r>1$), таких, что $V_1\cup V_2\cup...\cup V_r=V$, $V_k\cap V_l=\varnothing$, $k,l=1,...,r$, $k\ne l$.

Вершина $i\in V_k$, которая связана ребром с вершиной $j\in V_l$, $k\ne l$ называется **граничной вершиной**. Ребро, соединяющее вершины разных блоков, называются **ребром разреза**. Множество $E_{kl}=\{(i,j)\in E|\ i\in V_k, j\in V_l\}$ - множество ребер разреза между блоками V_k и V_l (некоторые из множеств E_{kl} могут быть пустыми), где V_k и V_l - множество E_{kl} которые называются когранице блоков.

Множества E_{kl} далее будем называть **подразрезами**.

Подграфом G_k будем называть граф с множеством вершин V_k и множеством ребер $E_k=\{(i,j)\in E\,|\,i\in V_k, j\in V_k\}$, то есть $G_k=(V_k,E_k)$.

Качество разбиения графа может оцениваться по одному или нескольким критериям, которые являются функциями параметров подразрезов E_{kl}, подмножеств $V_1, V_2, ..., V_r$ и/или подграфов $G_1, G_2, ..., G_r$:

$$\sum_{k<l} f(E_{kl}) \to \min(\max),$$

$$\sum_{k=1}^{p} g(V_k) \to \min(\max);$$

$$\sum_{k=1}^{p} h(G_k) \to \min(\max).$$

Возможны такие критерии оценки качества подразрезов, $f(E_{kl})$:

– суммарное количество ребер подразреза, $f(E_{kl}) = |E_{kl}|$;

– суммарный вес ребер подразреза (стоимость разреза), $f(E_{kl}) = \sum_{(i,j)\in E_{kl}} c_{ij}$;

– вероятность выхода из строя всех ребер подразреза $f(E_{kl}) = \prod_{(i,j)\in E_{kl}} (1 - p_{ij})$

(предполагается, что события «выход ребра из строя» являются независимыми);

– минимальная из вероятностей безотказной работы ребер подразреза $f(E_{kl}) = \min_{(i,j)\in E_{kl}} p_{ij}$;

– критерий, являющийся комбинацией параметров подразрезов, например сумма произведений весов ребер на вероятность выхода их из строя. $f(E_{kl}) = \sum_{(i,j)\in E_{kl}} c_{ij}(1 - p_{ij})$ (сводится к задаче разбиения с пересчитанной весовой функцией ребер).

Возможны такие критерии оценки качества получаемых подграфов $g(V_k)$, $h(G_k)$:

– плотность подграфа, $h(G_k) = \frac{2|E_k|}{|V_k|(|V_k| - 1)}$;

– максимальная из степеней вершин подграфа, $h(G_k) = \max_{i\in V_k}\{s_i\}$, где s_i – степень вершины $i \in V_k$ подграфа $G_k = (V_k, E_k)$;

– диаметр $d(G_k)$ подграфа $G_k = (V_k, E_k)$; диаметр графа (подграфа) - расстояние между двумя наиболее отдаленными вершинами в графе (подграфе);

– нарушение баланса между суммарными весами вершин подграфов (цель – найти такое разбиение, чтобы сумма весов вершин $g(V_k) = \sum_{i\in V_k} w_i$ в каждом подграфе были как можно близки друг к другу):

$$\max_{1\le k\le r}\{g(V_k)\} - \min_{1\le k\le r}\{g(V_k)\} \to \min;$$

– характеристики граничных вершин подграфов: $G_k = (V_k, E_k)$.

В задаче на получаемое решение также могут быть наложены некоторые ограничения (граничные условия), которые должны быть учтены при нахождении разбиения. К ним относятся ограничения (снизу и/или сверху) на следующие величины:

– количество подграфов разбиения;

– количество вершин, входящих в подграфы – так называемое *ограничение баланса*, согласно которому все подмножества $V_1, V_2, ..., V_r$ должны иметь примерно одинаковый размер, точнее, это ограничение требует, чтобы $|V_k| \le (1+\varepsilon)\left\lceil \frac{|V|}{r} \right\rceil$ $k = 1..r$ ($\varepsilon > 0$ - называется *параметром дисбаланса*);

– максимальный дисбаланс между количествами вершин подграфов разбиения $\max_k |V_k| / \left\lceil \frac{|V|}{r} \right\rceil$;

– суммарный вес вершин подграфов;

– размер дисбаланса между суммарными весами вершин подграфов;

– количество ребер, входящих в разрез;

– суммарный вес ребер разреза;

– вероятность выхода из строя всех ребер подразреза E_{kl};

– среднее количество ребер входящих в подразрезы E_{kl};

– средний вес ребер входящих в подразрезы E_{kl};

– характеристики граничных вершин подграфов $G_k = (V_k, E_k)$:

- степень d_i граничной вершины i;
- количество d_i^{cut} дуг, инцидентных граничной вершине i, входящих в подразрез ;
- вероятность выхода из строя «внутренних» и «внешних» ребер граничной вершины i.

Выводы. В работе предложен способ разбиения сети на зоны маршрутизации. Предложены и определены критерии оценки качества подразрезов сети, также наложены некоторые ограничения (граничные условия), которые должны быть учтены при нахождении разбиения.

Список использованной литературы.

1. Muhammed Fatih Talu. (2017, September) Multi-level spectral graph partitioning method . Journal of Statistical Mechanics: Theory and Experiment, (Volume 2017, pp. 5-11).

2. Dominique Lasalle, George Karypis (2013). Multi-threaded Graph Partitioning. Proceedings of the 2013 IEEE 27th International Symposium on Parallel and Distributed Processing (IPDPS '13) (pp. 225-236).

3. George Karypis, Vipin Kumar. (1998, November 07 - 13) Multilevel algorithms for multi-constraint graph partitioning. Proceedings of the 1998 ACM/IEEE conference on Supercomputing. (SC '98) (Pp. 1-13).

ДОВІДКА ПРО АВТОРІВ

Кулаков Олексій Юрійович – професор, кафедра обчислювальної техніки, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kulakov Yurii – professor, Department of Computer Engineering, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: ya.kulakov@gmail.com

Коган Алла Вікторівна – старший викладач, кафедра автоматизованих систем обробки інформації і управляння, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Kohan Alla - Senior lecturer, Department of Computer-Aided Management And Data Processing Systems, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: a.v.kohan433@gmail.com

Сперкач Майя Олегівна - старший викладач, кафедра автоматизованих систем обробки інформації і управляння, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Sperkach Maiia - Senior lecturer, Department of Computer-Aided Management And Data Processing Systems, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: sperkachmaya@gmail.com

Жданова Олена Григорівна – доцент, кафедра автоматизованих систем обробки інформації і управляння, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського».

Zhdanova Olena- associate professor, Department of Computer-Aided Management And Data Processing Systems, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”.

E-mail: zhdanova.elena@hotmail.com

Elena Zhdanova,
Alla Kohan,
Yurii Kulakov,
Maya Sperkach

ON THE PARTITION OF THE GRAPH INTO SUBGRAPHS

Relevance of research topic. The task of distributing the graph to smaller podgraphs was relevant many years ago and today is relevant in connection with the constant change in technology associated with the organization of computer networks. The task of distributing graphs to pieces in conventional networks is of the same nature, and when organizing the same process on the network there are many difficulties that need to be addressed.

The posing of problems. There is a need to break up the wireless computer network into routing zones in order to reduce the time of designing traffic and to form a multitude of paths for transferring information.

Analysis of recent research and publications. In recent years, the issue associated with the organization of the computer network sdn is of great interest, as it gives many opportunities in solving this problem.

Isolation of unexplored parts of a common problem. In the considered ways of partitioning a graph into subgraphs, there are no criteria for estimating the quality of subdivisions.

Formulation of the problem. The task is to partition the sdn network into the routing area, so that the minimization of the boundary cut is considered as a goal, and the requirement that the cuts have the same size can be considered as a limitation.

Presentation of the main material. The analysis of the methods of network splitting into zones is carried out and the relative criteria and difficulties are revealed. It is proposed to establish certain quality criteria for the breakdown of the graph. They can be evaluated according to one or more criteria, which are functions of the parameters of sub-divisions E_{ij}, subsets $V_1, V_2, ..., V_r$ and / or subgraphs $G_1, G_2, ..., G_r$

Conclusions. The method of splitting the network into routing zones is proposed. Proposed and defined criteria for assessing the quality of sub-divisions, also imposed some restrictions (boundary conditions), which must be taken into account when finding a partition.

Keywords: graph, section of graph, subgraph, partition of graph.

Наукове видання

Безпека. Відмовостійкість. Інтелект

Збірник праць міжнародної науково-практичної конференції
ICSFTI2018
10–12 травня 2018 р.

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Свідоцтво про державну реєстрацію: ДК № 5354 від 25.05.2017 р.,
просп. Перемоги, 37
м. Київ, 03056

Підп. до друку 12.07.2018. Формат 60×84$^{1}/_{16}$. Папір офс. Гарнітура Times.
Спосіб друку – ризографічний. Ум. друк. арк. 24,88. Обл.-вид. арк. 41,28. Наклад 30 пр.
Зам. № 18-093

Видавництво «Політехніка» КПІ ім. Ігоря Сікорського
вул. Політехнічна, 14, корп. 15
Київ, 03056
тел. (44) 204-81-78