

Міністерство освіти і науки України
Прикарпатський національний університет імені Василя Стефаника
Представництво "Польська академія наук" в Києві
Науково-технологічний університет "Гірничо-металургійна академія"
імені Станіслава Сташціа в Кракові
Вінницький національний технічний університет
Харківський національний університет радіоелектроніки
Національний авіаційний університет
Тернопільський національний економічний університет
Фінансово-економічний інститут Таджикистану
Економічна академія "Д.А.Ценов", Болгарія
Лудзький університет, Польща
Штуттгардський університет, Німеччина
Інститут інженерів з електротехніки та електроніки (ІЕЕЕ), Українська секція
Громадська організація "Івано-Франківський ІТ кластер"

**КОМП'ЮТЕРНІ НАУКИ,
ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ
ТА СИСТЕМИ УПРАВЛІННЯ**

**Матеріали
міжнародної науково-технічної конференції
молодих вчених, аспірантів та студентів**

**28–30 листопада 2018 року
Івано-Франківськ, Україна**

**COMPUTER SCIENCE, INFORMATION TECHNOLOGIES
AND MANAGEMENT SYSTEMS**

**Proceedings
of the International Scientific Young Scientists Conference**

CSYSC-2018

**2018, November , 28th to 30th
Ivano-Frankivsk, Ukraine**

**Івано-Франківськ
2018**

Електронне видання комбінованого
використання на CD-ROM

УДК 004+005

Наукові редактори: докт. техн. наук, проф. **Л.Б. Петришин** (ПНУ, АГН);
докт. техн. наук, проф. **П. Лебковський** (АГН)

Рецензенти:

д.т.н., проф. **В.В. Безкоровайний**;

д.т.н., проф. **П. Маслянка**;

д.т.н., проф. **В.М. Сінеглазов**.

Комп'ютерні науки, інформаційні технології та системи управління :
матеріали Міжнародної науково-технічної конференції студентів,
аспірантів та молодих вчених, м. Івано-Франківськ, 28–30 листопада 2018 року /
наук. ред. Л.Б. Петришин, П. Лебковський. – Електрон. дані. – Івано-Франківськ
: Прикарпатський національний університет імені Василя Стефаника, 2018.
– 1 електрон. опт. диск (CD-ROM); 12 см. – Назва з тит. екрану.

ISBN 978-966-640-448-3

Збірник містить матеріали Міжнародної науково-технічної конференції
з проблем комп'ютерних наук, інформаційних технологій, систем управління
та ігрового програмного забезпечення.

УДК 004+005

Зміст	3
СЕКЦІЯ 1	
КОМП'ЮТЕРНІ НАУКИ	7
FPGA for Implementations of Data Encryption and Cryptanalysis	8
Miroshnyk Maryna, Fedorin Dmytro, Miroshnyk Anatoliy	
Time series forecasting for 3D printing data	12
Oleh Kaida	
Приладові Панелі та їх Застосування у Бізнес-Аналітиці	17
Копп Андрій, Орловський Дмитро, Кука Данило	
Розробка Підходу до Оцінювання Якості Веб-Сайтів	21
Лютенко І.В, Курасов О.І.	
Аналіз Сучасного Стану Реалізації Методу Акустичної Емісії	24
Надія Івасишин	
Моделювання Топологічних Структур Мереж Логістики з Реверсивними Потоками	27
Володимир Безкоровайний, Аліна Зубенко, Світлана Мальцева	
Спосіб Формування Доменів Маршрутизації в SDN	30
Ю.О. Кулаков, А.В. Коган, Є.О. Павленкова	
Платформа для Локалізації Веб та Мобільних Додатків	32
Смолюц Дмитро, Мельник Жанна	
Моделі Надійності Алгоритмів, що Виконуються з Помилками Різних Типів	35
Олексій Козачко, Дмитро Солдатов, Артур Адаменко	
Інтегро-Диференційна Модель Ціноутворення на Енергоринку	37
Тетяна Заїка	
Непропорційне Масштабування Цифрового Зображення	39
Катерина Трифонова, Олександр Оголюк	
Оцінювання та Аналіз Навичок та Компетенцій Студентів ВНЗ	43
Кузьмін Олексій, Орловський Дмитро, Копп Андрій	
Комп'ютерне Моделювання Поверхневих Стоків	47
Іван Боровець, Ольга Мічута	
Засоби Псевдондетермінованого Потокового Шифрування	49
Караван Владислав, Баришев Юрій	
The Throughput Modeling of E-UTRA Network with Adaptive Modulation	52
Anna Artemenko-Didenko, Dmytro Makoveyenko	
Огляд Методів Моделювання Капілярних Поверхонь для Методу Висячої Краплі	55
Сусак Я.Я. , Малько О.Г.	
Вдосконалення Структур АЦП Шляхом Введення Надлишковості	58
Святослав Гриджан, Тарас Піварчук, Михайло Петришин	
Детектування Високочастотних Характеристик Цифрових Сигналів на Основі Трійкового Симетричного Вейвлет-Перетворення	60
Ігор Голуб, Артем Ізмайлов	
Застосування Трійкового Симетричного Вейвлет-Перетворення для Підвищення Ефективності Керування Процесом Буріння Нафтових і Газових Свердловин	63
Артем Ізмайлов	

Застосування Трійкового Симетричного Ортогонального Перетворення для Підвищення Ефективності Функціонування Пристроїв Обліку та Контролю Параметрів Передачі Газу	65
Артем Ізмайлов	
Узагальнена Методологія Побудови Систем Предиктивного Аналізу на Прикладі Гірського Устаткування в Умовах BIG DATA	67
Андрій Купін, Родіон Івченко	
Метод Вибору Маршруту для Перенаправлення Трафіку в SDN	75
А.В. Коган, Є.О. Павленкова,	
Система Захисту Програмного Забезпечення з Прив'язкою до Сервера	77
Куперштейн Леонід, Войтович Олеся, Беляєв Ігор.	
Елементна База та Моделювання Роботи Прототипу Пристрою для Неінвазивного Контролю Рівня Глюкози в Крові	79
Михайло Котик, Ігор Когут	
Метод Вдосконалення Одношляхових Протоколів Динамічної Маршрутизації	82
Шевчук К., Захарченко С.М.	
Дослідження Математичної Моделі Процесу Теплопередачі та Розробка на її Основі Пристрою для Контролю Температури Рідини	85
Михайло Шевчук	
СЕКЦІЯ 2	
ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ	87
Інформаційне Забезпечення Діяльності Морського Агента в Сервісній Ергатичній Системі	88
Петров І.М., Рудніченко М.Д., Шибаєва Н.О , Шибаєв Д.С	
The Measure of the Information Systems Security Risk Assessment Technique Based on a Fuzzy-Multiple Approach	90
N.V. Romashchenko, O.V. Shmatko, I.S. Litvinova	
Детектування Наявності Додаткової Інформації в Цифрових Зображеннях за Умови Різного Ступеню Наповненості Контентів	92
Ахмаметьєва Г.В., Мпугу Крістофер Бвабва	
Аналіз Параметрів Алгоритму Цифрового Підпису на Основі Гешування W-OTS	95
Олександр Марухненко, Геннадій Халімов, Михайло Наконечний	
Trading in Real Time Using an Expert Advisor Using Parabolic SAR, MA, CCI Technical Analysis Methods	98
Oleg Motuzov	
Використання Моделей VGG-Типу для Класифікації Зображень із Застосуванням Бібліотеки Pytorch	99
Дмитро Миронюк	
Система Виявлення Атак на Локальну Комп'ютерну Мережу	102
Майк Максимов, Ірина Шпінарева	
Параметричний Синтез Моделі Багатокритеріального Оцінювання на Основі Нейронних Мереж	106
Януш Вероніка, Володимир Безкоровайний	
Технології Big Data у Великій Політичній Грі	109
Ковальчук Ольга, Масьонкова Марія	

Digital Watermark Resistant to JPEG	111
Alexander Kozin, Alla Kobozeva, Mariia Kozina, Natalia Loginova, Elena Trofimenko	
Модифікація Засобів Open-Source для Підвищення Ефективності Розробки Програмного Забезпечення	113
Залізецький Василь	
Виявлення Штучних Змінень в Зображеннях за допомогою Нейронних Мереж	116
Глайборода Марина	
Система Підтримки Роботи Ситуаційного Центру на Основі Інтелектуальних Хмарних Технологій	119
Алла Штокал, Олександр Восьмушко, Михайло Кренцін, Леонід Куперштейн	
Задача Моделювання Мотивації Членів Інтелектуальної Команди	121
Пономарьова С.В., Калита Н.І.	
Інформаційно-Аналітична Служба як Елемент Кібербезпеки Держави	124
Миرونюк Віталій	
Обробка Просторових Даних	126
Новожилова М.В., Умріхін Ілля	
Паралельна Функціональна Обробка Інформації на LISP	128
Орест Гейко, Артур Марцінковський	
Обробка Великих Даних за допомогою ELK-стеку	131
Ольга Заверуха, Володимир Кобзев	
Експертна Система Моніторингу і Попередження Природних Надзвичайних Ситуацій на Автомобільних Дорогах	135
Леонід Нефьодов, Філь Наталія, Кононіхин Олександр, Біньковська Анжела	
Особливості Моделювання Процесів Оцінювання Поточного і Прогнозованого Технічного Стану Автотранспорту при Діагностуванні	139
О. В. Лапіна, К. Р. Лапін, Д. М. Андріца, Я. Ю. Алексенко	
Computer Modeling of Text Information Classification	143
Vadym Yakovenko, Yuliya Ulianovska, Oleksii Kaliaka	
Застосування Згорткових Нейронних Мереж в Задачах Класифікації Акустичних Даних	145
Олексій Кудін, Анастасія Кривохата	
Застосування Алгоритму CART для Автоматичної Класифікації Кардіограм	147
Олег Федоришин	
Використання Автоматизованих Інформаційних Систем Класу Learning Management System	150
М. І. Шевчук, М. В. Семаньків	
Аналіз Методів Обробки Зображень в Системі Визначення Патологій Шкіри Людини	153
Наталія Герасименко, Володимир Кобзев	
Вибір Алгоритму Стемінгу для Ефективного Повнотекстового Пошуку	156
Марина Кальметьєва, Володимир Кобзев	
CQRS-підхід в Розробці Програмного Забезпечення	159
Ілля Шутєєв, Володимир Кобзев	
Реалізація Технології Data Mining у Аналізі Ризиків Страхової Компанії	162
Михайло Лазар, Володимир Кобзев	
Особливості Технології Захисту Інформації в Сучасних Месенджерах	164
Збаражський Кирило, Давидова Вероніка, Мороз Дмитро	

Розробка Програмного Забезпечення Вибору Транспортних Засобів для Автотранспортних Підприємств за допомогою Веб-Технологій	167
Вячеслав Асаян, Леонід Нефьодов, Дмитро Маркозов	
СЕКЦІЯ 3	
СИСТЕМИ УПРАВЛІННЯ	170
Параметрування Швидкісних Лічильників PLC Simatic S7-1200 для Систем Управління з Імпульсними Перетворювачами	171
Николайчук М.Я, Гедзик Д.Я.	
Система Антипомпажного Захисту Газоперекачувального Агрегату на Базі Апаратно-Програмних Засобів Siemens	173
Татарин Т.М., Іванюк Н.І., Заміховський Л.М	
Розроблення Автономного Пристрою Моніторингу Температури на Базі Мк Esp8266 та Середовища Lua У Межах Концепції "Інтернету Речей" (Iot)	176
Юрій Паньків, Олександр Танюк	
Розробка Концепції Інформаційної Системи Зберігання і Обробки Інформації про Реабілітацію Осіб з Особливими Потребами	178
Панфьорова Ірина, Кислинський Єгор	
Впровадження Маркетингових Інформаційних Систем (МІС) на Підприємстві для Задач Прийняття Рішень	180
Владислав Шевченко, Марія Козуля	
Система Тестування Психофізіологічних Властивостей Водіїв	183
Поворознюк А.І., Стрюков А.О., Черних О.П., Двараковська О.В.	
Модель Адаптивного Контролера на Основі Нечіткої Логіки для Управління Системою "Розумний Дім"	185
Ємельянов Г.С, Крапівний Ю.М	
Математична Модель Задачі Планування Групових Розміщень в Системах Управління Готельним Бізнесом	187
Оксана Мазурова, Юлія Грушко.	
Розробка САД-Системи для Моделювання Динаміки Штучних Біо-Структур	191
Катерина Газдюк, Сергій Остапов, Денис Куц, Володимир Жихаревич	
Експертна Система Моніторингу і Попередження Природних Надзвичайних Ситуацій на Автомобільних Дорогах	194
Леонід Нефьодов , Філь Наталія, Кононіхин Олександр ,Біньковська Анжела	
Автоматизована Система для Світлотерапії з Частотним Скануванням Ділянки Опромінення	198
Олексій Яненко, Василь Кузь, Олександр Крушець	
Інтелектуальна Інформаційна Система Аналізу Вмісту Кошика Покупця Супермаркету	203
Іванна Сенейко, Наталія Кунанець, Антоній Ржеуський	
Мультимедійні Технології Навчання Студентів з Особливими Освітніми Потребами	205
Валерій Семенець, Петро Подпружніков, Аліна Овченко, Сергій Позняков	
Reddit Sentiment Analysis to Improve Election Predictions Bryant Hwang.....	207

СЕКЦІЯ 1

Комп'ютерні науки

Computer Science

FPGA for Implementations of Data Encryption and Cryptanalysis

Maryna Miroshnyk
*dept. Specialized Computer Systems
Ukrainian State University of Railway
Transport
Kharkiv, Ukraine
marinagmiro@gmail.com*

Dmytro Fedorin
*dept. Specialized Computer Systems
Ukrainian State University of Railway
Transport
Kharkiv, Ukraine
marinagmiro@gmail.com*

Anatoliy Miroshnyk
*dept. Automation and control in
technical systems
National Technical University "Kharkiv
Polytechnic Institute"
Kharkiv, Ukraine
marinagmiro@gmail.com*

Abstract— This paper presents two original programmable logic integrated circuits (FPGA) implementations of a DES encryption/decryption core. This implementations are the fastest ones known nowadays. In design, the plaintext, the key, and the mode (encryption/decryption) can be changed with no dead cycles. The resulting design is deployed on eight FPGAs and allows us to find key bits.

Keywords— Cryptography, Data Encryption Standard (DES), linear cryptanalysis, FPGA, efficient implementations

I. INTRODUCTION

The rapid growth of secure transmission is a critical point nowadays. We have to exchange data securely at very high data rates. Efficient solutions have to be hardware implemented and flexible in order to evolve with the permanent changes in norms. FPGA (Field Programmable Gate Array) implementations of the triple-Data Encryption Standard (triple-DES) efficiently meet these constraints. Triple-DES is based on three consecutive DES (without intermediate IP and IP-1 permutations). DES is very well suited for FPGA solutions. Some high-speed DES hardware implementations have been published in the literature. These designs unroll the 16 DES rounds and pipeline them. Patterson made a key-dependent data path for encryption in an FPGA which produces a bitstream of about 12 Gbps. Nevertheless, the latency to change keys is tenth of milliseconds. A DES implementation is also downloadable from FreeIP [12] and encrypts at 3.05 Gbps. Last known implementations were announced by Xilinx company, including FPGA implementations of a complete unrolled and pipelined DES encryptor/decryptor. The 16-stage and 48-stage pipelined ones could achieve data rates of, respectively, 8.4 Gbps and 2 Gbps. It also allowed changing the plaintext, the key, and the encryption/decryption mode on a cycle-by-cycle basis [1-3].

In this paper, author proposes new mathematical descriptions to implement and optimize DES in an FPGA. Author obtains two original designs. Both permit different pipeline levels and encrypt with data rates of 14.5 Gbps and 21.3 Gbps with, respectively, 21 and 37 cycles of latency.

In the second part, this paper deals with linear cryptanalysis. Linear cryptanalysis is a cryptanalytic technique that takes advantage of possible input-output correlations over a cipher. By evaluating the linear approximation for a sufficient number of plaintext / ciphertext pairs, it is possible to recover some bits of the key faster than an exhaustive search. Matsui [4], [5] proposed a known-plaintext linear attack against a full DES. It typically requires 243 known plaintext/ciphertext pairs to recover secret key bits. Recently, Knudsen and Mathiassen proposed three chosen-plaintext attacks [1], the third one becoming the

best chosen-plaintext attack against DES. Their first attack, which turns out to be less efficient from a theoretical point of view, gives birth to a very fast hardware implementation [6]. In fact, this attack requires only 212 chosen-plaintext/ciphertext pairs, but recovers only seven key bits. This attack allowed us [6] to recover the full key in less than two hours with eight FPGAs used in parallel. We denote it Knudsen's attack.

Although Matsui's linear cryptanalysis is the best known-plaintext attack known against DES nowadays, this attack still had a "theoretical" flavor, in the sense that very few experimental applications have actually been performed: A single known-plaintext experimentation for a full DES cipher has been performed in [5] and, until recently, remained the only practical test, to our knowledge. However, recent technological advances have made the required computing power reachable, as is witnessed by a set of 21 experiments for Matsui's approximation [2], [3], using the idle time of 18 Intel Pentium III MMX, capable of performing an attack in 4.32 days. Based on our fast DES implementation, we propose an FPGA implementation of Matsui's attack. It recovers 12+1 key bits in about 2.3 hours working with eight FPGAs. In terms of computation time, Knudsen's attack is better than Matsui's. Nevertheless, according to the number of plaintext / ciphertext pairs needed and the number of secret key bits found, Matsui's attack gives better results. In addition, Matsui's is a more realistic attack compared to Knudsen's attack due to the known-plaintext context. Our solution is very useful to perform practical tests, allowing a comparison with theoretical estimations. We believe that our implementations are the fastest implementations of Matsui's linear cryptanalysis known so far.

II. THE DES ALGORITHM

The Data Encryption Standard (DES) algorithm was adopted as a Federal Information Processing Standard for unclassified government communication. It is still largely in use. DES [10-11] encrypts 64-bit blocks with a 64-bit key, only 56 bits of which are used. The other 8 bits are parity bits for each byte. The algorithm has 16 rounds.

For the enciphering calculation, the plaintext is first permuted by a fixed permutation IP. The result is next split into the 32 left bits and the 32 right bits, respectively, L and R. The R part is expanded to 48 bits with the E box by doubling some R bits. Then, it performs a bitwise modulo 2 sum of the expanded R part and the 48-bit subkey K_i . The result of the XOR function is sent to eight nonlinear S-boxes (S). Each of them has six inputs bits and four outputs. The result is then permuted in the box P. Finally, to obtain the R part of the next round, a new modulo 2 sum is performed between the P output and the R part of previous round. In the

last round, no interchange of the 16-round R and L is performed; the ciphertext is calculated by applying the inverse of the initial permutation IP to the result of the 16th round. The secret key is expanded by the key schedule. The key schedule calculation is first based on the 56-bit permutation PC-1 whose output is split into 28-bit blocks C and D. Then, C and D are left shifted once or twice, depending on the index of the round. The 48-bit subkey is obtained by a second permutation, denoted PC-2. The DES algorithm is detailed in Fig. 1.

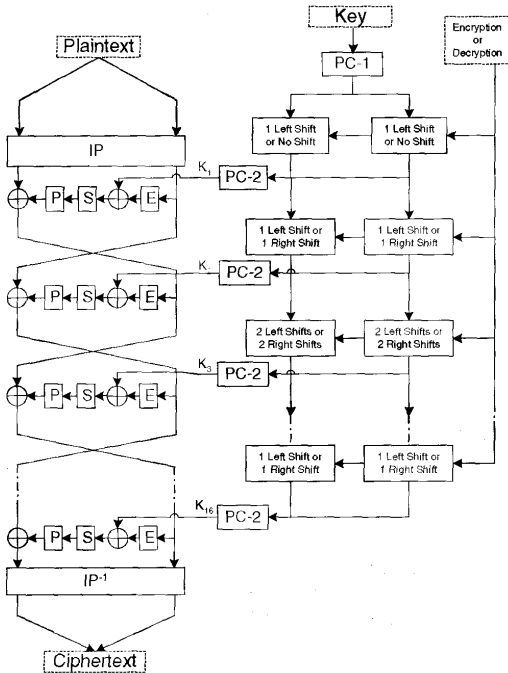


Fig. 1. the DES algorithm

III. XILINX IMPLEMENTATIONS

The first proposed solution is a full unrolled and pipelined DES implementation. It pipelines the data through 16 stages, putting registers after each enciphering/key round. This increases the data rate hugely, but also the logic requirement compared to a sequential design.

According to Fig. 4 and [14], the critical path through the round is quite long. First, a multiplexer selects the correct key bits depending on the encryptor/decryptor mode. The selected key bits are XORed with the R part. The resulting 6-bit fields are used to address the S-boxes whose critical path is one LUT followed by two multiplexer functions. Finally, output bits from the S-boxes are XORed with the L part. Fig. 2 details the critical path.

The first proposed way to reduce this critical path is to combine the F6 function with the final XOR operator. The resulting 4-bit input logic function that fits in an LUT and eliminates the F6 delay. Another improvement is to decouple the key from the enciphering calculation. This is done with a precomputation of the key schedule. So, the multiplexer selecting the key can be removed from the critical path, putting registers after this multiplexer.

Xilinx also proposes a second implementation. To reach higher data rates, one inserts a pipelined stage, respectively, after the key XOR and after F5 functions. It results in a 3-stage pipeline per round and a 48-stage pipeline over cipher.

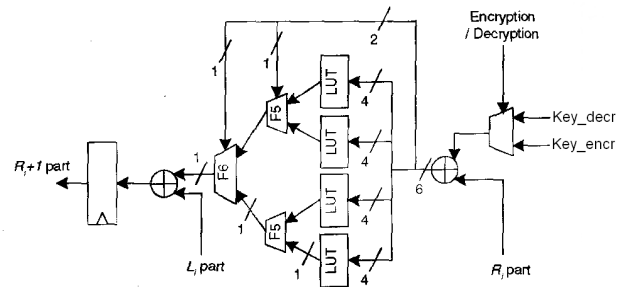


Fig. 2. Critical part of the DES design

Nevertheless, after checking and simulating their available source code on the web, we found two errors, are they forgot to put a 1-stage pipeline after the XOR before the key and R part. Actually, Xilinx implemented in 1-stage pipeline, but sent the XOR directly between that and R part into S-boxes, in place of the correspond registered value. They also forgot to register the key before the XOR function. Therefore, their critical path is quite a bit longer. Finally, their solutions do not imply a correct DES that can encrypt every cycle.

IV. PROPOSED FPGA DESIGNS

To be speed efficient, we propose designs that unroll the 16 DES rounds and pipeline them. In addition, we implemented solutions that allow us to change the plaintext, the key, and the encryption/decryption mode on a cycle-by-cycle basis, with no dead cycle. As a result, we can achieve very high data rates of encryption/decryption with exactly the same interface as Xilinx.

All of our implementations are first based on new mathematical representations of the DES algorithm. Indeed, the original description of DES is not optimized for FPGA implementation regarding the speed performance and the number of LUTs used. An FPGA is based on slice composed of two 4-bit LUTs (Look Up Tables) and two 1-bit registers. Therefore, an optimal way to reduce the LUTs used is to regroup all the logical operations in order to obtain a minimum number of blocks that take 4-bit inputs and give 1-bit outputs. In addition, we have to note that permutation and expansion operations (typically, P, E, IP, IP-1, PC-1, and PC-2) do not require additional LUTs, but only wire crossings.

A. First Solution

In [15], equivalent mathematical descriptions for DES are proposed. Based on these transformations, we propose new representations. First, we transform the round function of the enciphering computation. This transformation has no impact on the computed result of the round. Fig. 3. Modified representation of one DES-round. Fig. 3 shows a modified round representation, where we move the E box and the XOR operation. This involves the definition of a new function denoted R (like reduction):

$$R = E^{-1}; \forall x, R(E(x)) = x; \exists y[E(R(y)) \neq y. \quad (1)$$

Now, if we change all the enciphering parts of DES (see Fig. 1) with this modified round function and if we combine the E and XOR block with XOR block of the previous round, we get the architecture detailed in Fig. 4.

In this new arrangement of the DES structure, the first and last rounds are quite different from intermediate ones. Therefore, we obtain an irregular architecture. In addition,

we increase the number of E and R blocks, which does not alter the number of LUTs consumed. We also keep exactly the same number of S-boxes, which is the expensive part of the architecture. Finally, the number of modulo two sum operators is slightly increased by 32 additional 2-bit XOR operators. We can directly conclude that this design consumes more logic than Xilinx implementations.

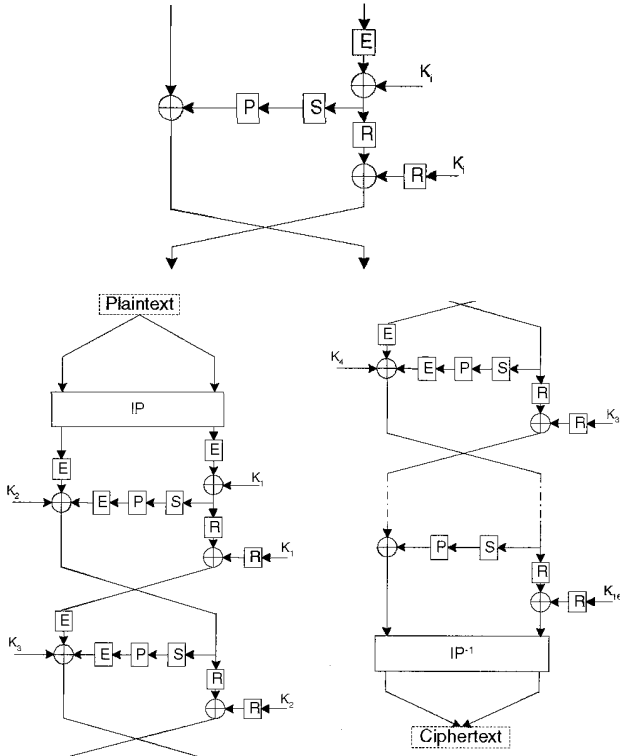


Fig. 4. First modified representation of the DES algorithm

The left part of Fig. 8 illustrates how the critical path, in our solution, is hugely decreased. We only keep one S-box operator and one XOR function. With this solution, we obtain a 1-stage pipeline per round. Due to the irregular structure of our design, we have to add an additional stage in the first round. To be speed efficient for implementation constraints, we also put a 2-stage pipeline, respectively, in the input and in the output. As mentioned in the figure, first and last registers are packed into IOBs. Therefore, we obtain a 21-stage pipeline. In the right part of Fig. 5, we put an extra pipelined stage in each round in order to limit the critical path to only one S-box. As a consequence, we get a 37-stage pipelined design.

B. Second Solution

Another solution is to move the R and XOR of the right part of the round into the left XOR operator of the previous round. As a result, we obtain the architecture shown in Fig. 6. As Fig. 6 underlines, we again obtain an irregular architecture. First and last rounds are quite different from intermediate rounds. We also keep exactly the same number of S-boxes as our precedent design. But, we really decrease the number of modulo two sum operators. We spare 15 x 32 2-bit XOR and can directly conclude that this design consumes less logic than Xilinx implementations.

Fig. 6 gives more details about the initial round of our design. To conclude, we propose efficient and different solutions in terms of space and data rate for the hardware implementation of DES. Depending on environment

constraints, we really believe that one of our designs should be well appropriate. Especially, our second proposal is very interesting in terms of speed, LUTs used, and registers. For high ratio Throughput/Area, the second 37-stage pipelined solution is very efficient.

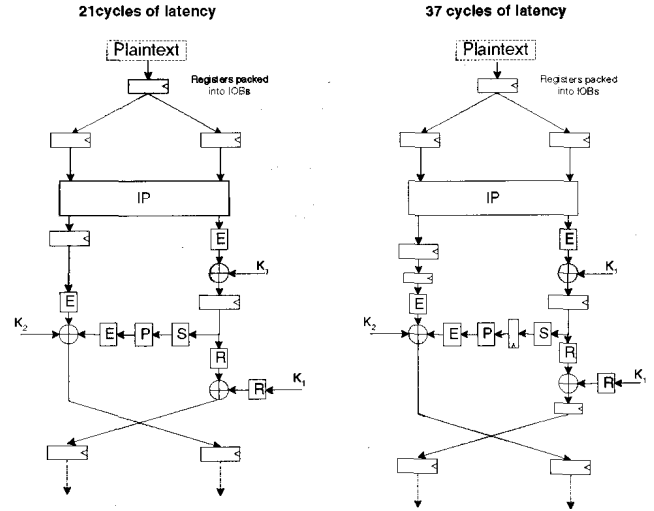


Fig. 5. Pipelining our first solutions.

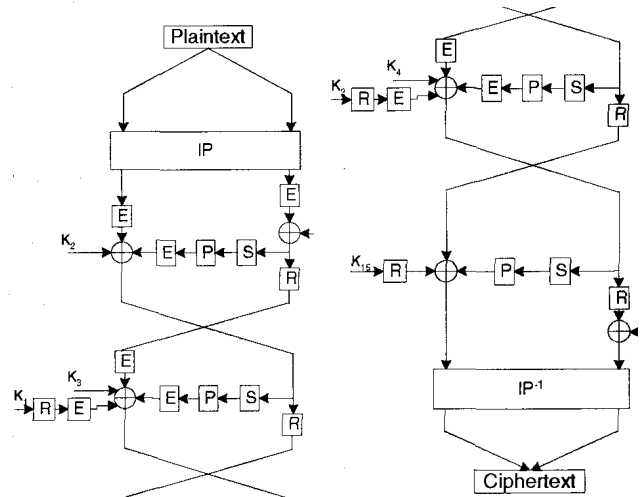


Fig. 6. Second modified representation of the DES algorithm.

V. LINEAR CRYPTANALYSIS

This section is a brief reminder of Matsui's linear cryptanalysis [4], [5], [6] before explaining the resulting VHDL design. Linear cryptanalysis is an attack based on the existence of some unbalanced linear relationship between inputs and outputs of a reduced-round version of the target encryption scheme. In the case of DES, Matsui used the relationship: $PL[15] \oplus PH[7,18,24,29] \oplus CL[7,18,24] = K1[22] \oplus K3[22] \oplus K4[44] \oplus K5[22] \oplus K7[22] \oplus K8[44] \oplus K9[22] \oplus K11[22] \oplus K12[44] \oplus K13[22]$, where $X[7,18,24,29] := X[7] \oplus X[18] \oplus X[24] \oplus X[29]$

Basically, this relationship means that the exclusive-or of some well-chosen bits of the plaintext (namely, the seventh, 18th, 24th, 29th bits of its high-order part) and some well-chosen bits of the ciphertext are equal to the exclusive-or of some well-chosen secret bits of the key with probability different from $\frac{1}{2}$.

VI. FPGA IMPLEMENTATION OF MATSUI'S ATTACK

As previously described, Matsui's linear cryptanalysis allows us to find 26 key bits with about 243 known-plaintexts. We propose an FPGA implementation of Matsui's attack that permits recovering $12 + 1$ key bits with about 243 known-plaintexts. We did not use the second relation to spare hardware resources and we decided to use our second 21-stage pipelined DES, which is the fewer resources consuming design. In order to increase speed performances, we parallelized two of them so that we got a data rate of two encryptions per cycle. We also modified them in order to gain resources space: The key schedule was simplified and the input and output registers were removed.

Nevertheless, for a hardware implementation, the main problem of this attack is the 212 counters needed to perform the key guess. Knowing that about 24,000 LUTs available on our FPGA, the implementation of 212 parallelized counters is much too expensive to be realistic (at 65,000 LUTs). (We have to keep a sufficient bits size, say 16 bits, for the counters to have an efficient and feasible implementation.)

This section will briefly introduce how we implement Matsui's linear cryptanalysis without 4,096 parallel counters in one FPGA board, keeping our very fast data throughput. We do it with 4,096 RAM-based counters. (We configure all the RAMs to have 8-bit address and 16-bit data.)

In practice, we need to implement 4,096 RAM based counter values, with only 32 parallel access (with read and write operations; we use dual access RAMs). Therefore, this operation can be performed in 128 clock cycles.

This is practically performed using a large serial/parallel converter making the ciphertext bits involved in Matsui's linear approximation (4) available during 128 cycles.

By choosing the plaintext bits involved in the linear approximation (4) such that they are fixed during the same 128 clock cycles, we avoid the need of a serial/parallel converter for the plaintext bits. We also avoid the use of I/O operators between plaintext and ciphertext parts. Therefore, we spare a lot of hardware resources. We just need an n -delay shift register (SR block) to synchronize the design.

To generate plaintext bits, we use an LFSR of 57 bits and a 6-bit counter (the remaining bit is used for the two DES parallelization). This counter controls the PL part used to calculate $F1[15]$, varying every 128 cycles. Therefore, we obtain 128 successive cycles where the PL part of $F1(PL, K1)$ is constant. Knowing 256 parallelized results of $PL[7,18,24,29] \oplus PH[15] \oplus CH[7,18,24] \oplus F16(CL, K16)$, we have to count the number of bits equal to 0 and subtract 128, thanks to the previous comment (we only store the bias). We obtain 9-bit result, called bias in Fig. 11. Depending on the 32 parallelized values of $F1(PL, K1+i)$ [15], we have to carry out a subtraction or an addition between the 32 RAM values stored (in the correct address) and the bias value. (We have i = from 0 to 31 and $K1$ equal to 0 or 1.) Therefore, we get one Matsui's attack implementation that allows us to recover $12 + 1$ secret key bits. Our cryptanalysis design is based on a sequentialized access of 4,096 counters, without altering the encryption rate of two DES per cycle. To analyze our experiments, the 4,096 RAMs stored results are sent to the PC when one of them exceeds the 16-bit RAM data size. In

addition, the PC can send the secret key to the FPGA board. This allows us to perform very practical tests.

VII. EXPERIMENTAL RESULTS

In this section, we give the results we got running Matsui's attack on eight Xilinx FPGAs (VIRTEX1000 bg560-4). We carried out the experiments at a work frequency = 66.6 MHz (=226) (Because of the FPGA heat running at 66 MHz, we do not carry experiments at higher frequency. It is why we use our second 21-stage solution, which is the less resource consuming design.). Therefore, we are able to compute 2×226 equations per second. Using eight FPGA boards, 243 evaluations take less than 2.3 hours. We performed tests with 71 different keys. Table 7 summarizes the experimental success rate of the attack for various amounts N of chosen-plaintext/ciphertext pairs. These experimental results suggest that Matsui's theoretical analysis is quite good (see Tables 5 and 6). Indeed, our results are very close to mathematical estimations.

ACKNOWLEDGMENT

This paper deals with two new ideas for FPGA implementations of DES leading to four improved practical appropriate implementations. All of them are very efficient in terms of speed and/or resources needed. Then, this paper presents the first known FPGA implementation of Matsui's linear cryptanalysis. The resulting attack is capable of finding a 13-bit key in less than 2.3 hours, using eight FPGA boards. In addition, it is worth noting that, with the new Xilinx FPGA (Xilinx VIRTEX-II XC2V8000), we would be able to carry out the same attack in about 1 hour, using only one FPGA board. Therefore, in some applications, FPGAs can be used as powerful cryptographic calculation tools.

REFERENCES

- [1] Miroshnik M.A., Kovalenko M.A., "Uses of programmable logic integrated circuits for implementations of data encryption standard and its experimental linear cryptanalysis" *Інформаційно-керуючі системи на залізничному транспорті*, pp. 36-44, 2013/6.
- [2] Miroshnik M.A., "Implementation of cryptographic algorithms on FPGA-based digital distributed systems," *Інформаційно-керуючі системи на залізничному транспорті*, ті: науково-технічний журнал, Харків: УкрДУЗТ № 2 (111) 2015. - с. 25-30
- [3] M.A. Miroshnik, "Методи захисти інформації в розподілених комп'ютерних сегях," *Інформаційно-керуючі системи на залізничному транспорті*, №5, с. 66-70, 2014.
- [4] Miroshnik M.A., Synchronizing Sequences For Verification Of Finite State Machines. / 9th IEEE International Conference on Dependable Systems, Services and Technologies (DESSERT'2018) // May 24 - 27, 2018.
- [5] Miroshnyk M.A. Practical Methods for de Bruijn sequences Generation using Non-Linear Feedback Shift Registers / 14th IEEE International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering // Lviv-Slavske, Ukraine. - 2018/2, p.35.
- [6] Miroshnik M.A. Design logical control units with finite state machine patterns. . 16th IEEE EAST-WEST DESIGN & TEST SYMPOSIUM (EWDTS-2018).14-17.09.2018 .
- [7] M.A. Miroshnik, Methods For Designing Self-Checking Digital Machines. Telecommunications and Radio Engineering. 2017. Begel House Inc. – Vol.76., Num.15, p.1367-1377
- [8] Miroshnyk Marina Communication channel statistical characteristics research methods. 2016. - IEEE Modern Problems of Radio Engineering, Telecommunications and Computer Science, Proceedings of the 13th International Conference on TCSET 2016, p.562-565..

Time Series Forecasting for 3D Printing Data

Oleh Kaida
Ivan Franko National University of
Lviv
Lviv, Ukraine
olkokayda@gmail.com

Abstract — The thesis is devoted develop time series forecasting for 3d printing data. The work consists of an introduction, tree chapters, containing conclusion and list of references. The introduction discusses the goal, object and subject of the thesis.

Keywords — 3D printing, Time series forecasting, CUDA

I. INTRODUCTION

As the number of uses for 3D printing grows, so too will it's reliance on data analysis. Forgoing the more technical details, 3D printing is complex technology that now requires more data to perform it's many capabilities. This can probably best be seen in the manufacturing sector where 3D printing already has a firm foothold. Data analysis plays a major role as a contributor to computer-aided design.

The major challenge from this change in strategy comes from storing the information for each customized product. 3D objects require a tremendous amount of data to print, and the more complex the item, the more data is needed. But even this challenge is one that can be overcome with 3D printing as organizations have the potential to print out their own electronics, including storage devices.

For the moment, it's cheaper for businesses to simply purchase the device instead of creating it for themselves, but the cost of 3D printing continues to go down, especially in manufacturing.

Data analysis also has a role to play in the quality control aspect of 3D printing product creation. The act of 3D printing a lego block, for example, is relatively simply with a high degree margin of error, but what about items that are far more complex, like car engine parts. An error in the process could not only lead to a defective item but a health and safety hazard depending on the item and what it's used for.

In the case of the most complex creations, an alignment error or a fraction of a millimeter can be dangerous. That doesn't even include factors such as temperature, expansion, structural integrity, and more.

II. BASICS AND DATA PREPARATION FOR TIME SERIES FORECASTING

A. Time series forecasting is an important area of machine learning that is often neglected.

A normal machine learning dataset is a collection of observations. For example:

Observation #1,
Observation #2,
Observation #3.

Time does play a role in normal machine learning datasets.

Predictions are made for new data when the actual outcome may not be known until some future date. The future is being predicted, but all prior observations are almost always treated equally. Perhaps with some very minor temporal dynamics to overcome the idea of "concept drift" such as only using the last year of observations rather than all data available.

A time series dataset is different.

Time series adds an explicit order dependence between observations: a time dimension. Time series analysis involves developing models that best capture or describe an observed time series in order to understand the underlying causes. This field of study seeks the "why" behind a time series dataset.

This often involves making assumptions about the form of the data and decomposing the time series into constitution components.

The quality of a descriptive model is determined by how well it describes all available data and the interpretation it provides to better inform the problem domain.

The primary objective of time series analysis is to develop mathematical models that provide plausible descriptions from sample data

Making predictions about the future is called extrapolation in the classical statistical handling of time series data.

More modern fields focus on the topic and refer to it as time series forecasting. Forecasting involves taking models fit on historical data and using them to predict future observations.

Descriptive models can borrow for the future (i.e. to smooth or remove noise), they only seek to best describe the data.

An important distinction in forecasting is that the future is completely unavailable and must only be estimated from what has already happened.

The purpose of time series analysis is generally twofold: to understand or model the stochastic mechanisms that gives rise to an observed series and to predict or forecast the future values of a series based on the history of that series. Time series analysis provides a body of techniques to better understand a dataset.

Perhaps the most useful of these is the decomposition of a time series into 4 constituent parts:

- Level. The baseline value for the series if it were a straight line.
- Trend. The optional and often linear increasing or decreasing behavior of the series over time.

- Seasonality. The optional repeating patterns or cycles of behavior overtime.
- Noise. The optional variability in the observations that cannot be explained by the model.

All time series have a level, most have noise, and the trend and seasonality are optional. There are few problems here:

- LSTMs expect 3D input, and it can be challenging to get head around this the first time.
- LSTMs don't like sequences of more than 200-400 time steps, so the data will need to be split into samples.

B. Preparing data for the LSTM network in Keras

- Load the Data
- Drop Time
- Split Into Samples
- Reshape Subsequences

III. OVERVIEW AND CONFIGURATION TOOLS FOR TIME SERIES FOR FORECASTING

A. TensorFlow

TensorFlow is a fast, flexible, and scalable open source machine learning library for research and production.

TensorFlow is an open source software library for numerical computation using data flow graphs. Nodes in the graph represent mathematical operations, while the graph edges represent the multidimensional data arrays (tensors) communicated between them. The flexible architecture allows to deploy computation to one or more CPUs or GPUs in a desktop, server, or mobile device with a single API.

B. Keras

TensorFlow is a fast, flexible, and scalable open source machine learning library for research and production. Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano. It was developed with a focus on enabling fast experimentation. Being able to go from idea to result with the least possible delay is a key to doing good research.

Use Keras if there need a deep learning library that:

- Allows for easy and fast prototyping (through user friendliness, modularity, and extensibility).
- Supports convolution networks and recurrent networks, as well as combinations of the two.
- Runs seamlessly on CPU and GPU.

C. CUDA. TensorFlow performance test: CPU VS GPU

CUDA is a parallel computing platform and application programming interface (API) model created by Nvidia. It allows software developers and software engineers to use a CUDA-enabled graphics processing unit (GPU) for general purpose processing – an approach termed GPGPU (General-Purpose computing on Graphics Processing Units). The CUDA platform is a software layer that gives direct access to

the GPU's virtual instruction set and parallel computational elements, for the execution of compute kernels. Test conditions and setup:

In order to test every piece of equipment fairly, decided to focus on a common and reproducible deep learning task, such as training CNN on Cifar-10 dataset using tensorflow/models, which download on PC using git clone <https://github.com/tensorflow/models.git>

To reproduce the test, require internet connection and a python environment with installed tensorflow on top. Simply go to directory tutorials/image/cifar10 and run the following code from terminal `python cifar10_train.py`

Test metric:

A single comparison metric is number of examples processed per second (the more the better).

Test notes

1. i5-7300HQ

2018-05-20 14:19:03.298426: step 8010, loss = 1.68 (583.4 examples/sec; 0.219 sec/batch)

2018-05-20 14:19:06.423052:tep 8020, loss = 1.63 (409.1 examples/sec; 0.312 sec/batch)

2018-05-20 14:19:09.493130: step 8030, loss = 1.39 (416.9 examples/sec; 0.307 sec/batch)

2018-05-20 14:19:12.563831: step 8040, loss = 1.47 (416.8 examples/sec; 0.307 sec/batch)

2018-05-20 14:19:15.653078: step 8050, loss = 1.31 (414.3 examples/sec; 0.309 sec/batch)

2. i3-7100

2018-05-30 19:12:03.298426: step 8010, loss = 1.68 (585.4 examples/sec; 0.221 sec/batch)

2018-05-30 19:12:06.425052:tep 8020, loss = 1.63 (404.6 examples/sec; 0.314 sec/batch)

2018-05-30 19:12:09.497130: step 8030, loss = 1.39 (416.9 examples/sec; 0.307 sec/batch)

2018-05-30 19:12:12.567831: step 8040, loss = 1.47 (416.8 examples/sec; 0.307 sec/batch)

2018-05-30 19:12:15.657078: step 8050, loss = 1.31 (414.3 examples/sec; 0.309 sec/batch)

3. GTX1050

2018-05-20 15:04:50.073763: step 8110, loss = 1.08 (4789.9 examples/sec; 0.026 sec/batch)

2018-05-20 15:04:50.329753: step 8120, loss = 0.90 (4807.1 examples/sec; 0.026 sec/batch)

2018-05-20 15:04:50.586215: step 8130, loss = 0.76 (4804.2 examples/sec; 0.026 sec/batch)

2018-05-20 15:04:50.842610: step 8140, loss = 0.90 (4804.6 examples/sec; 0.026 sec/batch)

2018-05-20 15:04:51.098361: step 8150, loss = 0.91 (4808.6 examples/sec; 0.026 sec/batch)

GPUs have almost 200 times more processors per chip than a CPU. For example, an Intel Xeon Platinum 8180 Processor has 28 Cores, while an NVIDIA Tesla

K80 has 4,992 CUDA cores. While a CPU core is more powerful than a GPU core, the vast majority of this power goes unused by applications. A CPU core is designed to support an extremely broad variety of tasks (e.g., render a webpage, drive word processors and enterprise software, manage peripherals) in addition to performing computations, whereas a GPU core is optimized exclusively for data computations. Because of this singular focus, a GPU core is simpler and has a smaller die area than a CPU, allowing many more GPU cores to be crammed onto a single chip. Consequently, applications, which perform large numbers of computations on a vast amount of data, can see huge (i.e., 5 to 10 times) performance improvements when running on a GPU versus a CPU.

IV. DEVELOPMENT TIME SERIES FORECASTING FOR 3D PRINTING DATA

It is common to use persistence or naive forecasts as a first-cut forecast on time series problems.

A better first-cut forecast on time series data with a seasonal component is to persist the observation for the same time in the previous season. This is called seasonal persistence.

It is important to evaluate time series forecasting models consistently.

First, hold the last two years (2016 and 2017) of data back and evaluate forecasts on this data. This works for both monthly and daily data look at.

Use a walk-forward validation to evaluate model performance. This means that each time step in the test dataset will be enumerated, a model constructed on historical data, and the forecast compared to the expected value. The observation will then be added to the training dataset and the process repeated.

Walk-forward validation is a realistic way to evaluate time series forecast models as one would expect models to be updated as new observations are made available.

Finally, forecasts will be evaluated using root mean squared error, or RMSE. The benefit of RMSE is that it penalizes large errors and the scores are in the same units as the forecast values (car model sales per month).

Root mean squared error used to report model skill.

Case Study 1: Monthly Car Model Sales Dataset

The Monthly Car Model Sales Dataset describes the number of car model sales between 2010 and 2017.

The units are a count of the number of sales and there are 108 observations.

A line plot of the data is also provided. Both are a yearly seasonal component and an increasing trend.

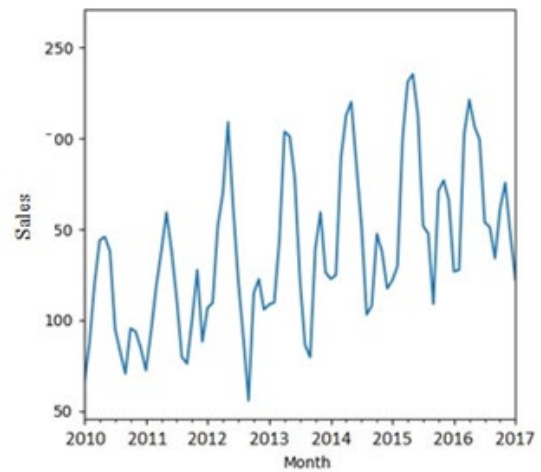


Fig. 1. Line Plot of Monthly Car Model Sales Dataset

The prior 24 months of data will be held back as test data. Investigate seasonal persistence with a sliding window from 1 to 5 years.

The results suggest that taking the average from the last three years is a good starting model with an RMSE of 18.630 car sales.

A plot of the relationship of sliding window size to model error is created.

The plot nicely shows the improvement with the sliding window size to 3 years, then the rapid increase in error from that point.

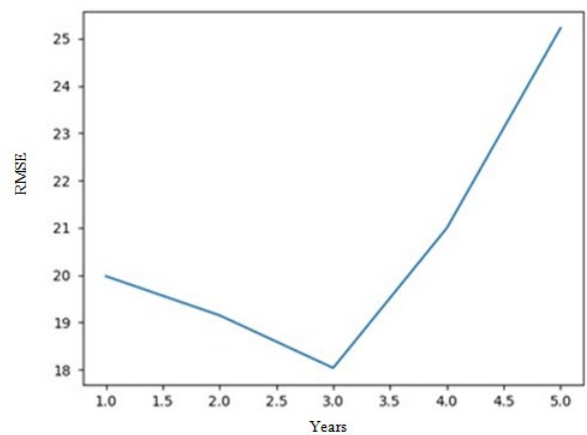


Fig. 2. Sliding Window Size to RMSE for Monthly Car Sales

A. Time Series Forecasting with the Long Short-Term Memory Network in Python

The Long Short-Term Memory recurrent neural network has the promise of learning long sequences of observations. It seems a perfect match for time series forecasting. Dataset describes the monthly number of sales of model over a 3-year period. The units are a sales count and there are 36 observations. A line plot of the series is then created showing a clear increasing trend.

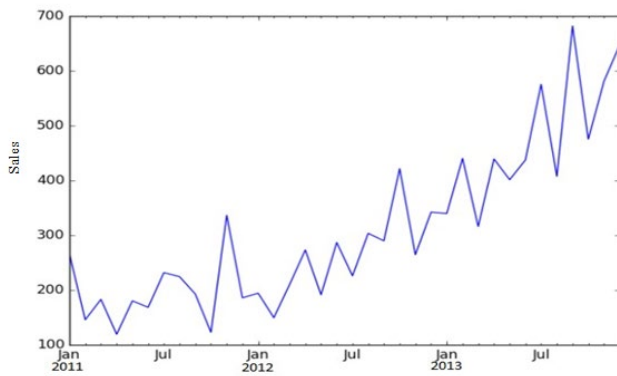


Fig. 3. Line Plot of Monthly Model Sales Dataset

Split the Model Sales dataset into two parts: a training and a test set. The first two years of data will be taken for the training dataset and the remaining one year of data will be used for the test set.

A rolling forecast scenario will be used, also called walk-forward model validation.

Each time step of the test dataset will be walked one at a time. A model will be used to make a forecast for the time step, then the actual expected value from the test set will be taken and made available to the model for the forecast on the next time step.

This mimics a real-world scenario where new Model Sales observations would be available each month and used in the forecasting of the following month.

Finally, all forecasts on the test dataset will be collected and an error score calculated to summarize the skill of the model. The root mean squared error (RMSE) will be used as it punishes large errors and results in a score that is in the same units as the forecast data, namely monthly Model sales.

A good baseline forecast for a time series with a linear increasing trend is a persistence forecast.

The persistence forecast is where the observation from the prior time step ($t-1$) is used to predict the observation at the current time step (t).

Implement this by taking the last observation from the training data and history accumulated by walk-forward validation and using that to predict the current time step.

Accumulate all predictions in an array so that they can be directly compared to the test dataset.

Running the example prints the RMSE of about 136 monthly Model sales for the forecasts on the test dataset.

A line plot of the test dataset (blue) compared to the predicted values (orange) is also created showing the persistence model forecast in context.

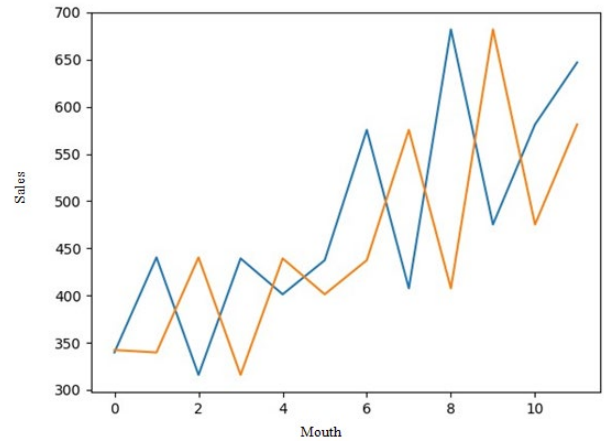


Fig. 4. Persistence Forecast of Observed vs Predicted for Model Sales Dataset

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Before fit an LSTM model to the dataset, must transform the data. This section is broken down into three steps:

- Transform the time series into a supervised learning problem
- Transform the time series data so that it is stationary.
- Transform the observations to have a specific scale.

The Long Short-Term Memory network (LSTM) is a type of Recurrent Neural Network (RNN).

A benefit of this type of network is that it can learn and remember over long sequences and does not rely on a pre-specified window lagged observation as input.

In Keras, this is referred to as stateful, and involves setting the “stateful” argument to “True” when defining an LSTM layer. Define a function called `fit_lstm()` that trains and returns an LSTM model. As arguments, it takes the training dataset in a supervised learning format, a batch size, a number of epochs, and a number of neurons.

The `batch_size` must be set to 1. This is because it must be a factor of the size of the training and test datasets.

`Predict()` function on the model is also constrained by the batch size; there it must be set to 1 because are interested in making one-step forecasts on the test data.

Batch Size: 1

Neurons: 4

Once the LSTM model is fit to the training data, it can be used to make forecasts.

In result prints the expected and predicted values for each of the 12 months in the test dataset.

The example also prints the RMSE of all forecasts. The model shows an RMSE of 71.721 monthly model sales, which is better than the persistence model that achieved an RMSE of 136.761 model sales.

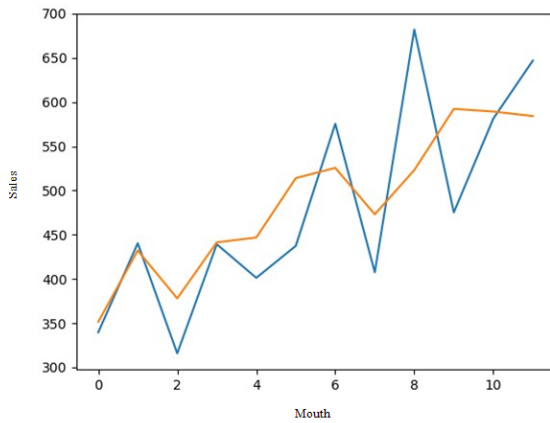


Fig. 5. Line Plot of LSTM Forecast vs Expected Values

A line plot of the test data (blue) vs the predicted values (orange) is also created, providing context for the model skill.

A difficulty with neural networks is that they give different results with different starting conditions.

One approach might be to fix the random number seed used by Keras to ensure the results are reproducible. Another approach would be to control for the random initial conditions using a different experimental setup.

Repeat the experiment from the previous section multiple times, then take the average RMSE as an indication of how well the configuration would be expected to perform on unseen data on average.

This is often called multiple repeats or multiple restarts.

In result prints the RMSE score each repeat. The end of the run provides summary statistics of the collected RMSE scores.

Mean and standard deviation RMSE scores are 138.491905 and 46.313783 monthly model sales respectively.

A box and whisker plot is created from the distribution shown below. This captures the middle of the data as well as the extents and outlier results.

This is a very useful result as it suggests the result reported above was probably a statistical fluke. The experiment suggests that the model is probably about as good as the persistence model on average (136.761), if not slightly worse.

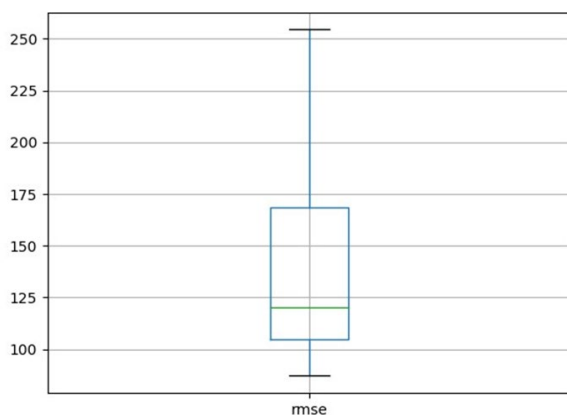


Fig. 6. LSTM Repeated Experiment Box and Whisker Plot

CONCLUSION

Was implemented time series forecasting for printing data and consists of seasonal persistence forecasting and time series forecasting with the LSTM network in Python with Keras on GPU.

The Long Short-Term Memory network (LSTM) is a type of Recurrent Neural Network (RNN).

A benefit of this type of network is that it can learn and remember over long sequences and does not rely on a pre-specified window lagged observation as input.

- Loaded the dataset from CSV file.
- Transformed the dataset to make it suitable for the LSTM model, including:
 - Transforming the data to a supervised learning problem.
 - Transforming the data to be stationary.
 - Transforming the data so that it has the scale -1 to 1.
- Fitting a stateful LSTM network model to the training data.
- Evaluating the static LSTM model on the test data.
- Report the performance of the forecasts.

REFERENCES JITEPATYPA

- [1] Brownlee J. Introduction to Time Series Forecasting With Python / Jason Brownlee. – 367 p.
- [2] Brownlee J. Long Short-Term Memory Networks With Python / Jason Brownlee. – 246 p.
- [3] Delgado R. Big Data's Role in 3D Printing [Electronic resource] / Rick Delgado. -2016. - Resource access mode: <https://techvibes.com/2016/07/29/big-data-3d-printing>.
- [4] Installing TensorFlow [Electronic Resource] - Resource Access Mode: https://www.tensorflow.org/install/install_windows.K. Elissa, "Title of paper if known," unpublished.
- [5] Keras: The Python Deep Learning library [electronic resource] - Access to the resource: <https://keras.io/#keras-the-python-deep-learning-library>.
- [6] CUDA Toolkit Documentation [electronic resource] - Access to the resource: <https://docs.nvidia.com/cuda/>.
- [7] Space-efficient online approximation of time series data: Streams, amnesia, and out-of-order / Gandhi, Sorabh, Foschini, Suri // IEEE. – 2010.
- [8] Processors [electronic resource] - Access to the resource: <https://ark.intel.com/products/>.
- [9] 3D Printing: Data, Data Everywhere [Electronic resource]. - 2013. - Resource access mode: <http://deloitte.wsj.com/cio/2013/06/25/3d-printing-data-data-everywhere/>.
- [10] Redwood B. The additive manufacturing process [electronic resource] / Ben Redwood. - 2015. - Access to the resource: <https://www.3dhubs.com/knowledge-base/additive-manufacturing-process>.

Приладові Панелі та їх Застосування у Бізнес-Аналітиці

Кука Данило Олександрович
Національний технічний
університет «Харківський
політехнічний інститут»
Харків, Україна
danbkabest805@gmail.com

Орловський Дмитро Леонідович
Національний технічний
університет «Харківський
політехнічний інститут»
Харків, Україна
ordm@kpi.kharkov.ua

Копп Андрій Михайлович
Національний технічний
університет «Харківський
політехнічний інститут»
Харків, Україна
kopp93@gmail.com

Indicators Dashboards and their Usage in Business Analytics

Kuka Danylo Oleksandrovych
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine

Orlovskyi Dmytro Leonidovych
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
ordm@kpi.kharkov.ua

Kopp Andrii Mykhailovych
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
kopp93@gmail.com

Анотація—У даній роботі розглянуті питання щодо застосування приладових панелей для вирішення задач бізнес-аналізу, що дозволяє поліпшити якість вирішення задач управління та аналізу діяльності підприємства.

Abstract—This paper considers the use of indicators dashboards in order to solve business analysis problems. It allows improving the quality of solving the management and analysis problems related to enterprise activities. The relevance of this research is defined by the need for continuous improvement of business performance of Ukrainian and foreign organizations due to the constant calls of the competitive environment. Therefore, this paper discusses the foundations of using dashboards in business analysis, their purpose, and role in organizational management. According to the considered survey, modern organizations tend to be interested in applying dashboards systems. This growing interest depends on huge analytical abilities of operational, tactical, and strategic dashboards that provide information for management decisions on each level of organizational governance. As the example, this work considers the prototype of the operational dashboard intended to visualize the selected indicators of the supply process. The common procedure of building the dashboard system based on indicators is outlined.

Ключові слова—панелі індикаторів, ключові показники ефективності, тактичні панелі, стратегічні панелі.

Keywords—indicators dashboards, key performance indicators, tactical dashboards, strategic dashboards.

I. ВСТУП

В умовах високої мінливості світових ринків промислових товарів і послуг, посилення міжнародної конкуренції і низького економічного зростання значно зростають вимоги до процедур оцінки результатів економічної діяльності українських підприємств. Системна криза економіки країни, падіння споживчих доходів і відсутність можливостей для зовнішнього запозичення капіталу привели до того, що для керівництва підприємств ціна помилкових управлінських рішень стає украй високою. Економічна і

фінансова невизначеність обумовлює переведення підприємств фактично на ручне управління. Це означає, що найважливішою умовою здобуття об'єктивної картини справ на підприємстві стає впровадження систем моніторингу показників ефективності, максимально наближеного до реального часу. Завдання управлінського моніторингу в широкому сенсі полягає в тому, щоб реєструвати певні параметри процесу, які називають ключовими показниками або індикаторами ефективності (Key Performance Indicators, KPI), і в наочній формі надавати їх користувачеві. Це дає можливість в оперативному режимі не лише отримувати сигнали про можливі відхилення, але і встановлювати їх причини і намагатися виправити ситуацію. Також повинна існувати можливість передачі результатів [1].

Як відомо, одними з основних завдань управління сучасним підприємством є регулярний аналіз і оцінка ефективності управлінської діяльності, що передбачає виділення ключових бізнес-процесів підприємства і концентрацію на них зусиль виконавчого менеджменту. Для реалізації всіх цих завдань співробітникам необхідні сучасні інструменти контролю і аналізу ефективності компанії, представлені у вигляді набору панелей індикаторів.

II. МЕТОДИКА І РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Поняття «панель індикаторів (indicator dashboard)» присутнє в дослідженнях багатьох вітчизняних і зарубіжних фахівців в області управління, при цьому існують різні підходи до його трактування. Згідно з визначенням американського дослідника У. Еккерсона [2], панеллю індикаторів називається багатошарове застосування на базі інфраструктури бізнесу-аналізу і інтеграції даних, яке дозволяє організації здійснювати вимір, моніторинг і управління бізнесом з більшою ефективністю. Тим самим йдеться не про набір графічної і табличної інформації, а про корпоративну інформаційну систему, що дозволяє менеджменту

організації проводити оперативне управління і добиватися досягнення поставлених цілей.

Панелі індикаторів вирішують наступні завдання інформаційного забезпечення підприємства:

- Моніторинг поточної діяльності, а також зіставлення значень отриманих показників з величинами, визначеними в корпоративній стратегії на поточний період.

- Аналіз отриманих результатів шляхом занурення користувача системи на будь-який рівень показників і даних. Крім того, панелі індикаторів надають користувачам інструмент для проведення різного роду розрахунків, включаючи створення запитів, аналітичну обробку даних, формування повідомлень і звітів і т. д.

- Забезпечення користувача інструментами оперативного управління, оцінки рішень, складання сценаріїв розвитку бізнесу, що дозволяє керівництву організації вибирати правильний напрям подальшого руху бізнесу.

Інформація на панелях індикаторів відбивається в трьох розрізах, або шарах. Перший шар – узагальнене графічне представлення, зазвичай є списком показників, які істотні для організації в поточний період часу [2]. Значенням показників додають спеціальний кольоровий фон, залежний від величини відхилення актуального значення показника від планових (або нормативних) значень [3].

Багатовимірне представлення даних є другим шаром панелей індикаторів і дає детальнішу інформацію про причини виникнення відхилень. В разі багатовимірного аналізу користувачі переглядають дані по різних розрізах, наприклад, по географічному місцю розташування, по клієнтах і т. д., що дозволяє проаналізувати причини виникнення відхилень, проглянувши рахунки, документи і повідомлення.

Третій шар дозволяє користувачам переглядати детальні звіти та оперативні повідомлення, надає можливість виконувати запити до сховища даних [2].

Кожна панель індикаторів призначена для виконання трьох комплексів функціональних завдань, а саме:

- Відстежування критично важливих бізнес-процесів і види діяльності, використовуючи показники ефективності, що забезпечують подачу застережливих сигналів при виникненні потенційних проблем.

- Виявлення причин виникнення проблем через аналіз релевантної і актуальної інформації під різними кутами і з різними рівнями деталізації.

- Управління людьми і процесами так, щоб знаходити оптимальні рішення, максимізувати ефект і вести організацію у напрямі досягнення цілей.

Необхідно відзначити, що багато компаній, що використовують панелі індикаторів в своїй діяльності, при їх створенні часто спираються на запропоновану на початку 90-х рр. XX ст. Р. Капланом і Д. Нортоном [4] збалансовану систему показників, що є механізмом послідовного доведення до персоналу стратегічних цілей

компанії і контролем їх досягнення через ключові показники ефективності.

Згідно з результатами дослідження [5], починаючи з 2013 року знову починає спостерігатися зростання зацікавленості організацій у впровадженні та використанні систем панелей індикаторів (рис. 1).

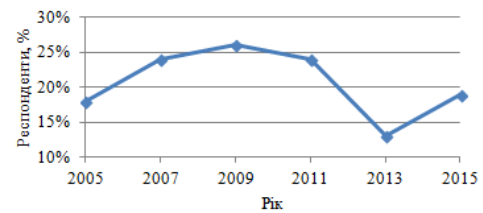


Рис. 1. Зацікавленість організацій у впровадженні та використанні панелей індикаторів.

Зі 116 організацій, які прийняли участь в опитуванні у 2016 році, 20 організацій вже використовують панелі індикаторів у своїй діяльності, а також ще 15 організацій планують використовувати панелі індикаторів для управління та аналізу діяльності підприємства.

III. УПРАВЛІННЯ ПІДПРИЄМСТВОМ НА ОСНОВІ ВИКОРИСТАННЯ ПАНЕЛЕЙ ІНДИКАТОРІВ

Важливою умовою успішного функціонування панелей індикаторів є єдиний формат інформації, що вводиться. Для забезпечення єдиного підходу у використанні і представленні інформації необхідно реалізувати систему нормативно-довідкової інформації підприємства (у складі єдиної бази знань). Здійснюючи впровадження панелей індикаторів, компанії отримують для себе наступні переваги:

- **Інформування співробітників про стратегічні цілі.** Панелі індикаторів представляють корпоративну стратегію через вимірювані показники, цільові значення і різні ініціативи у вигляді, адаптованому для кожної групи (або рівня) персоналу. У будь-який час співробітник підприємства, відкривши відповідну панель індикаторів, отримує ясне уявлення про стратегічні цілі компанії і про те, що він повинен зробити в своїй області для досягнення цих цілей.

- **Коректування стратегії.** Керівники використовують панелі індикаторів як засіб «тонкого підстроювання» для відображення послідовних коректувань курсу на шляху до стратегічної мети, використовуючи стратегічну карту підприємства.

- **Розширення зони видимості.** Панелі індикаторів розширюють «зону видимості» керівників і менеджерів середньої ланки з приводу якості виконання рутинних операцій і оперативних показників, своєчасно надаючи достовірні дані і прогнозуючи тенденції, що складаються, виходячи з результатів минулої діяльності. Це дозволяє компаніям заздалегідь передбачати проблеми, які могли б вплинути на кінцеві результати.

- **Поліпшення взаємодії персоналу.** Оскільки дані про ефективність праці окремих співробітників при використанні панелей індикаторів стають фактично загальнодоступними, це спонукає співробітників різних відділів до тіснішої співпраці. Це також сприяє розвитку діалогу між менеджерами і рядовими співробітниками

відносно результатів роботи, прогнозів, що дозволяє менеджерам частіше і конструктивно аналізувати роботу компанії.

- **Посилення мотивації співробітників.** Візуалізуючи оцінні показники і результати роботи персоналу, панелі індикаторів підсилюють його мотивацію виконувати завдання більш оптимальним чином, принаймні, в тих областях, де результати піддаються виміру. Якщо оплата праці буде прив'язана до результатів роботи, то відбивані панеллю індикаторів дані спонукають співробітників працювати краще для здобуття додаткового прибутку.

- **Створення єдиної інформаційної картини функціонування бізнесу.** Панелі індикаторів консолідують і інтегрують корпоративну інформацію, використовуючи загальні регламенти, нормативи і показники, що піддаються виміру. В результаті формується актуальний набір корпоративних даних, якими користуються всі співробітники компанії, тому конфлікти серед менеджерів і аналітиків зводяться до мінімуму.

Таким чином, панелі індикаторів надають потрібну інформацію їх користувачам в потрібний час, забезпечуючи певну оптимізацію ухвалення управлінських рішень, підвищуючи їх ефективність і сприяючи досягненню цільових результатів [1, 2].

У практиці побудови інформаційної інфраструктури підприємств розрізняють панелі індикаторів трьох основних типів: оперативні, тактичні і стратегічні.

A. Оперативні панелі індикаторів

Використовуються співробітниками нижчої ланки і їх супервайзерами для моніторингу і оптимізації основних робочих процесів. Співробітники організацій

використовують операційні панелі індикаторів, щоб стежити за процесами і операціями, інформація про яких, залежно від характеру процесу, виводиться на інформаційні панелі через декілька секунд, хвилин або годин. Зазвичай інформація в операційних панелях індикаторів оновлюється автоматично.

Оновлення відбувається або відразу після настання події, тобто практично в режимі реального часу, або ж через деякий фіксований проміжок часу відповідно до встановленого регламенту. Найчастіше інформацію на операційних панелях індикаторів оновлюють, як мінімум, двічі в добу [2].

Приклад розробленої оперативної панелі індикаторів наведено на рис. 2. Дана панель індикаторів призначена для аналізу KPI процесу постачання продукції, згідно з моделлю SCOR (Supply Chain Operations Reference) [6], та прийняття відповідних управлінських рішень:

- Витрати на постачання (Cost to Supply, CtS).
- Час циклу постачання (Supply Cycle Time, SCT).
- Відсоток замовлень, поставлених у повному обсязі (Supplied in Full, OSF%).
- Відсоток замовлень, поставлених вчасно (Orders Supplied in Time, OST%).

Для демонстрації показників використовуються такі засоби, як стовпчаста діаграма (для CtS), лінійний графік (для SCT), а також спеціальні засоби візуалізації – датчики або вимірювальні прилади (для OSF% та OST%), які були обрані на основі їх поширеності у галузі бізнес-аналітики та інформативності [7].



Рис. 2. Прототип розробленої панелі індикаторів.

B. Тактичні панелі індикаторів

Відстежують процеси і проекти на рівні відділів і представляють інтерес для певного сегменту організації або обмеженої групи співробітників. Менеджери і бізнес-аналітики використовують тактичні панелі індикаторів для порівняння показників в своїх проектах або зонах відповідальності з бюджетними планами,

прогнозами або результатами, отриманими за останній час. Зазвичай тактичні панелі оновлюються щонеділі або щомісячно. Основна увага в них приділяється аналізу і управлінню, а не моніторингу.

Тактичні панелі індикаторів потрібні для того, щоб менеджери і аналітики могли швидко отримати

інформацію, необхідну їм для управління процесами або оцінки міри вирішення завдань підрозділів [2].

Таким чином, тактичні панелі індикаторів є особливою формою представлення звітів. Такі інтерактивні звіти дозволяють менеджерам робити «моментальні знімки» даних будь-якого рівня узагальнення і на їх основі формувати звіти необхідного рівня деталізації, відправляти їх по електронній пошті або представляти на сайті організації. В цьому випадку різниця між інтерактивним звітом і тактичною панеллю індикаторів практично відсутня.

C. Стратегічні панелі індикаторів

Дозволяють відстежувати просування до стратегічних цілей і часто створюються на базі системи збалансованих показників. Мета використання стратегічної панелі індикаторів полягає в тому, щоб зорієнтувати всю організацію на досягнення стратегічних цілей і добитися, щоб всі групи рухалися в одному і тому ж напрямі. Для цього в компанії створюються окремі спеціалізовані збалансовані системи показників для кожної групи.

Такі каскадні збалансовані системи показників, що зазвичай оновлюються щомісячно, надають в розпорядження керівників потужний інструмент, що дозволяє «доводити» стратегію компанії до її співробітників, чітко і глибше уявляти собі виконувани операції і виділяти найважливіші чинники підвищення ефективності і вартості бізнесу. У стратегічних панелях індикаторів більший акцент робиться на управлінні, ніж на моніторингу і аналізі.

Ключовим моментом в процесі реалізації проекту впровадження панелей індикаторів є вибір і затвердження керівництвом підприємства набору основних показників ефективності [2].

D. Ключові показники ефективності

Ключові показники ефективності (KPI) – це показники, які піддаються кількісному вимірюванню та вважаються найбільш важливими для оцінки ефективності діяльності фірми, відділу чи окремого працівника. KPI – це показники, досягнення яких необхідне для того, щоб наблизитися до поставлених цілей [2, 4].

Застосування системи KPI в організації дає змогу оцінювати її стан та формувати стратегію розвитку; довести цілі всієї організації до структурних підрозділів й окремого співробітника; контролювати ділову активність в режимі реального часу; створити ефективну систему винагороди працівників; оптимізувати фонд винагороди персоналу в організації [4].

Основні етапи побудови системи KPI:

- Уточнення цілей організації.
- Підтримка та заохочення керівництва організації.
- Формування робочої групи.
- Формування цілісної стратегії впровадження і розвитку KPI.
- Розробка цілей і задач підрозділів та показників їх ефективності.

- Опис процесу збору даних для розрахунку преміальних виплат.
- Опис процесу збору даних для розрахунку преміальних виплат.
- Узгодження показників ефективності роботи та їх важливість (вага – розподіляється у відсотках між показниками групи).
- Планування та розрахунок преміального фонду організації та її підрозділів.
- Формування єдиної бази даних та внесення в неї виробничих показників.
- Відбір ключових показників ефективності для організації в цілому.
- Розробка відповідних документів.
- Презентація розробленої системи керівникам та іншому персоналу.
- Впровадження KPI.
- Контроль за результатами впровадження системи KPI.
- Уточнення показників ефективності для підтримки їх актуальності.

Головною задачею при впровадженні даної системи є інформаційна робота з персоналом організації. Правильне формування KPI набагато покращить професійну діяльність окремих працівників та організації в цілому.

IV. ВИСНОВКИ

Панелі індикаторів є ефективним інструментом управління, який у повній мірі задовольняє потребам менеджменту підприємства, що змінилися в умовах економічної кризи. Залежно від поставлених керівництвом підприємства завдань можуть використовуватися операційні, тактичні і стратегічні типи панелей індикаторів. Перед ухваленням рішення про впровадження систем панелей індикаторів керівництво має бути упевнене в тому, що організація готова технічно, технологічно і організаційно до таких змін. Ключову увагу при впровадженні панелей індикаторів слід приділити відбору показників. Рекомендується, щоб в панелі індикаторів були присутні показники, що відображають різні перспективи діяльності підприємства.

ЛІТЕРАТУРА REFERENCES

- [1] B. Wetzstein et al., "Monitoring and analyzing influential factors of business process performance," *Enterprise Distributed Object Computing Conference*, EDOC'09, pp. 141-150, 2009.
- [2] W. Eckerson, *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*. John Wiley & Sons, 2010.
- [3] L. Papas and L. Whitman, "Riding the technology wave: Effective dashboard data visualization," *Human Interface and the Management of Information. Interacting with Information*, pp. 249-258, 2011.
- [4] R. S. Kaplan and D. P. Norton, *The Balanced Scorecard: Translating Strategy into Action*. Harvard Business Review Press, 1996.
- [5] P. Harmon, *The State of Business Process Management. Technical Report*. BPTrends, March 2016.
- [6] S. Council, *Supply Chain Operations Reference Model Revision 11.0*. Washington, DC: APICS, 2012.
- [7] W. Echerson and M. Hammond, *Visual Reporting and Analysis: Seeing Is Knowing: Asset Page*. TDWI-The Data Warehousing Institute, 2011.

Розробка Підходу до Оцінювання Якості Веб-Сайтів

Лютенко І.В.

Кафедра програмної інженерії та інформаційних технологій управління
Національний технічний університет "Харківський політехнічний інститут"
Харків, Україна
liv@kpi.kharkov.ua

Курасов О.І.

Кафедра програмної інженерії та інформаційних технологій управління
Національний технічний університет "Харківський політехнічний інститут"
Харків, Україна
kurasov.oleksii@gmail.com

Website Quality Evaluation Approach Development

Kurasov O.I.

Department of software engineering and management
information technologies
National Technical University "Kharkiv Polytechnic Institute"
Харків, Україна
kurasov.oleksii@gmail.com

Liutenko I.V.

Department of software engineering and management
information technologies
National Technical University "Kharkiv Polytechnic Institute"
Харків, Україна
liv@kpi.kharkov.ua

Анотація—Визначено основні проблеми оцінювання веб-сайтів. Відібрані критерії оцінювання. Обрано метод оцінювання та шкали критеріїв. **The main problem of website evaluation was determined. Evaluation criteria were selected. The method of evaluation and the scale of the criteria were chosen.**

Ключові слова—оцінювання якості, веб-сайт, критерій, агрегування критеріїв, багатокритеріальний об'єкт, quality evaluation, website, criterion, criterion aggregation, multicriterion object

I. ВСТУП

У наш час веб-сайти є одним із найефективніших маркетингових інструментів бізнесу, який допомагає йому витримувати конкуренцію на ринку, розширювати власний сегмент у даному ринку, збільшувати об'єм продажу товарів та формувати власний імідж. Крім того, сайт дозволяє взаємодіяти клієнтам та постачальникам товарів та послуг без зайвих витрат часу та ресурсів.

Актуальність роботи зумовлена важливістю оцінювання якості веб-сайтів, їх порівняння, тому що в наш час збільшується кількість веб-ресурсів, для яких критично важливим є утримання користувачів та ріст їх кількості, що може приносити прибуток для підприємства, що обслуговує сайт.

Головною проблемою усіх ресурсів для оцінювання сайтів є те, що вони оцінюють ресурси без врахування їх типу за такою системою оцінювання, що є адекватною лише для одного типу сайту.

II. ПОСТАНОВКА ЗАДАЧІ

Зрозуміло, що неможливо створити універсальний метод, який дозволив би у повному обсязі оцінити веб-сайти різного призначення, наприклад, сайти-візитівки та інтернет-магазини. Тому необхідно зосередитися на формуванні такої множини критеріїв оцінювання, яка була би спільною для більшості сайтів, що існують у мережі. Ці критерії входять до трьох основних складових якості сайту:

1) технічна реалізація – критерій, що визначає якісний рівень розробки сайту та узагальнює прості критерії технічного характеру.

2) придатність до індексування – критерій, що визначає ступінь пристосованості сайту до його якісного індексування пошуковими системами.

3) зручність користування – складова, що визначає ступінь позитивності досвіду користувачів сайту.

Важливою складовою якості сайту також є інформаційна захищеність, але

оцінка інформаційної безпеки для сайтів багатьох категорій за одним лише методом дасть неточний результат з причини узагальненості самої процедури, що не є критичним для оцінювання вищенаведених складових якості. Визначені складові можуть бути представлені наступним переліком показників, причому показник може впливати на декілька складових якості одразу:

1) «GZIP-архівация» - показник, що дає знання про те, чи використовує сервер методи стиснення даних для пришвидшення роботи користувача.

2) «Обработка помилки 404» – відображає наявність коректної процедури обробки помилки 404. Є важливою як для індексації сайту, так і для зручності використання сайту.

3) «Переадресация» – показує, чи коректно сервер оброблює дубльовані запити користувача. Переадресация є важливим чинником оптимальної індексації сайту.

4) «Оптимізація зображень» – показує обсяг зображень, що оптимізовані для обробки браузером.

5) «Обсяг оптимізованих URL-адрес» – показує обсяг сторінок, що мають зручний для сприйняття користувачем формат.

6) «Конфігурування пошукових систем» - показує наявність файлів robots.txt/sitemap.xml, що дозволяють вказувати пошуковій системі лише ту частину сайту, що необхідно індексувати. Це дозволяє заощадити обмежений час на індексацію та використати його з більшою користю для отримання високої позиції у видачі пошукових систем.

7) «Обсяг деталізованих посилань» - обсяг посилань, опис яких відповідає вимогам лаконічності та інформативності.

8) «Адаптованість» – відображає обсяг сторінок, що є налаштованими для показу на дисплеях різних розмірів за допомогою медіа-запитів.

9) «Реалізація мета-тегів» - обсяг протестованих сторінок сайту, що мають мета-теги description та keywords, що сприяють кращій індексації сайту.

10) «Реалізація заголовків на сторінках» – виражається через обсяг сторінок із правильною структурою заголовків, яка поліпшує результати індексації сайту.

Відношення між показниками та складовими якості описані у таблиці 1.

TABLE 1. ПОКАЗНИКИ ТА СКЛАДОВІ ЯКОСТІ САЙТУ

Назва показника	Складові, на які впливає
GZIP-архівація	Технічна реалізація
Обробка помилки 404	Технічна реалізація; Придатність до індексування
Переадресація	Технічна реалізація
Оптимізація зображень	Технічна реалізація; Зручність користування
Обсяг оптимізованих URL-адрес	Технічна реалізація; Придатність до індексування; Зручність користування
Конфігурування пошукових систем	Придатність до індексування
Обсяг деталізованих посилань	Зручність користування
Адаптованість	Зручність користування
Реалізація мета-тегів	Придатність до індексування
Реалізація заголовків на сторінках	Придатність до індексування

III. ВИРІШЕННЯ ЗАДАЧІ БАГАТОКРИТЕРІАЛЬНОГО ВИБОРУ

Задача багатокритеріального вибору формується наступним чином. Дається множина варіантів A_1, \dots, A_p , кожен з яких характеризується певними критеріями K_1, \dots, K_m . Кожен критерій K_i має шкалу

$$X_i = \{x_i^1, \dots, x_i^{g_i}\}, i = 1, \dots, m$$

дискретні числові або вербальні градації в якій у більшості випадків є впорядкованими. Необхідно, на

основі вподобань особи, яка приймає рішення, обрати один або декілька найкращих варіантів із множини представлених.

Основною складністю є те, що необхідно одночасно аналізувати як кількісні, так і якісні показники різного ступеня важливості, велика кількість яких сильно ускладнює оцінювання веб-сайтів. Додатковою обставиною є те, що для оцінки якості веб-сайту не існує єдиної моделі оцінки якості.

Вирішення подібної задачі можна побудувати на основі задачі пошуку екстремуму однієї або багатьох функцій корисності(цінності). Для виконання задачі необхідно вивести узагальнений критерій із числа багатьох числових критеріїв шляхом їх згортання та знаходження зваженої суми. При великій кількості критеріїв такий метод є занадто трудомістким, тому що потребує від аналітика предметної області великих витрат часу на виведення приблизної функції корисності, а також коефіцієнтів важливості(ваги), які необхідно призначити кожній взятій до уваги властивості, що само по собі є задачею з неоднозначним рішенням [1].

Використання огрублених множин у класифікації багатокритеріальних об'єктів полягає у використанні наборів правил, що визначаються ОПР для віднесення альтернатив до певного класу з різним ступенем точності. Метод є достатньо складним, тому що велике число правил класифікації ускладнює їх аналіз. Крім того, метод потребує попереднє налагодження на підготовлених масивах даних [2].

Враховуючи слабку структурованість задачі, можна звернутися до методології вербального аналізу рішень. Згідно цієї методології, властивості варіантів описуються за допомогою якісних критеріїв, які мають словесні формулювання градацій на шкалах оцінок. В такому випадку числові коефіцієнти важливості критеріїв та цінності варіантів не розраховуються та не використовуються, а вербальні оцінки ознак не перетворюються у числові показники. Таким чином, використовуючи тільки якісні виміри, на множині кортежів багатокритеріальних оцінок задаються відношення верховенства та еквівалентності варіантів рішення, за допомогою яких можна здійснити вибір кращого варіанту.

Оцінити багатокритеріальний об'єкт, яким є веб-сайт, що має велику кількість якісних показників, можливо за допомогою методу ПАКС («Последовательное Агрегирование Классифицируемых Состояний») [3].

Процедура оцінювання одного сайту включає такі етапи:

1) Скорочення розмірності вихідної задачі, наприклад за допомогою введення нових критеріїв, що агрегують критерії початкової групи.

2) Розробка нових шкал оцінювання для критеріїв вищого рівня. Кожній градації вищого критерію відповідає власний кортеж градацій ієрархічно підлеглих показників.

ОПР визначає також смисловий зміст критеріїв та градацій шкал оцінок. Критерії повинні мати такі шкали оцінок, які, з одного боку, будуть відображати

агреговані властивості об'єктів, а з іншого боку, будуть зрозумілі ОПР при остаточному впорядкуванні та класифікації об'єктів.

Після виконання цих кроків ми отримуємо можливість оцінювання окремого сайту. З самого початку метод ПАКС був запропонований як метод агрегації думки багатьох експертів (оцінок, що вони запропонували) при оцінюванні та ранжуванні багатьох альтернатив. Усі альтернативи впорядковувалися за допомогою методу АРАМИС («Агрегирование и Ранжирование Альтернатив Многопризнаковых Идеальных Ситуаций») [2] за цим агрегованим показником від теоретично найгіршої до теоретично найліпшої з альтернатив. Подібна логіка може бути використана, якщо для оцінювання сайту буде запропоновано кілька варіантів ієрархії показників або шкал оцінювання критеріїв.

А. Приклад агрегування критеріїв

Розглянемо невеликий ілюстративний приклад. Нехай ОПР необхідно побудувати шкалу складеного критерію D з градацій оцінок базових показників А, В і С. Припустимо, що всі критерії А, В, С і D мають шкали з трьома вербальними градаціями порядкових оцінок $A = \{a^0, a^1, a^2\}$, $B = \{b^0, b^1, b^2\}$, $C = \{c^0, c^1, c^2\}$ і $D = \{d^0, d^1, d^2\}$, де e^0 - відмінно, e^1 - середньо, e^2 - погано.

Застосовуючи для побудови порядкової шкали складеного критерію метод стратифікації кортежів [2], ОПР може, наприклад, об'єднати градації оцінок вихідних критеріїв в узагальнені градації складеного критерію за таким принципом: все кращі оцінки за базовими показниками утворюють одну кращу оцінку по складеному критерію, все середні оцінки - одну середню оцінку, все гірші оцінки - одну гіршу оцінку. Приклад побудови шкали методом стратифікації кортежів зображений на рис. 1.

$a^0 b^0 c^1$	$a^0 b^0 c^2$	$a^0 b^2 c^0$	$a^2 b^0 c^0$	$a^1 b^2 c^2$
$a^0 b^0 c^0$...	$a^1 b^1 c^1$...	$a^2 b^1 c^2$
$a^1 b^0 c^0$	$a^1 b^1 c^2$	$a^1 b^2 c^1$	$a^2 b^1 c^1$	$a^2 b^2 c^1$
d^0	d^1	d^2		

Fig. 1. Приклад побудови шкали методом стратифікації кортежів

У таблиці 2 наведено шкали вимірювання критеріїв, що використовуються.

TABLE II. КРИТЕРІЇ ОЦІНЮВАННЯ ТА ЇХ ШКАЛИ

Назва критерію	Рівень	Градації
GZIP-архівація	1	1) виконується; 2) відсутня;
Обробка помилки 404	1	1) виконується; 2) відсутня;
Переадресація	1	1) виконується; 2) відсутня;
Оптимізація зображень		1) великий – 80-100% зображень налаштовані; 2) середній – 30-80% зображень налаштовані; 3) малий – 0-30% зображень налаштовані.
Обсяг оптимізованих URL-адрес	1	1) повний – 100% сторінок мають такі адреси; 2) частковий – мінімум половина сторінок мають такі адреси;

		3) малий – менше половини сторінок мають оптимізовані адреси.
Конфігурування пошукових систем	1	1) виконується; 2) відсутнє;
Обсяг деталізованих посилань	1	1) великий – 70-100% посилань деталізовані; 2) середній – 30-70% посилань деталізовані; 3) малий – 0-30% посилань деталізовані.
Адаптованість	1	1) повна – адаптація під смартфони та планшети; 2) задовільна – адаптованість хоча б під один вид пристрою дисплею; 3) відсутня – немає жодного медіа-запиту під планшети або смартфони
Реалізація мета-тегів	1	1) повна – усі сторінки мають коректні заголовки; 2) неповна – хоча б 30% сторінок мають заголовки, але є сторінки без них; 3) незадовільна – 0-30% сторінок мають заголовки
Реалізація заголовків на сторінках	1	1) повна – усі сторінки мають коректні заголовки; 2) задовільна – 30% сторінок мають заголовки, але є сторінки без них; 3) незадовільна – 0-30% сторінок мають заголовки
Технічна реалізація	2	1) відмінна; 2) добра; 3) задовільна; 4) слабка
Придатність до індексування	2	1) висока придатність; 2) задовільна придатність; 3) незадовільна придатність
Зручність користування	2	1) висока; 2) середня; 3) низька
Загальний показник якості	3	1) відмінна якість; 2) висока якість; 3) добра якість; 4) задовільна якість; 5) низька якість

ВИСНОВКИ

При оцінюванні якості веб-сайту він може бути представлений як багатокритеріальний об'єкт, що має великий обсяг показників якісного характеру. Введення агрегованих критеріїв спростить процес оцінювання альтернативи у вигляді веб-ресурсу.

Вибір методу ПАКС дозволить отримати змістовну та зрозумілу оцінку на великій кількості критеріїв з найменшими витратами часу на побудову системи оцінювання для особи що приймає рішення.

ЛІТЕРАТУРА REFERENCES

- [1] Петровский А. Б. Теория принятия решений / А. Б. Петровский Москва : Издательский центр «Академия». – 2009. 402 с.
- [2] Doumpos M., Zopounidis C. Multicriteria Decision Aid Classification Methods. – Dordrecht: Kluwer Academic Publishers, 2002.
- [3] Петровский А. Б. Многокритериальный выбор с уменьшением размерности пространства признаков: многоэтапная технология ПАКС / А. Б. Петровский, Г. В. Ройзензон // Искусственный интеллект и принятие решений. – 2012. - No 4. – С. 88 – 103.

Аналіз Сучасного Стану Реалізації Методу Акустичної Емісії

Надія Івасишин
кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет імені Василя Стефаника
Івано-Франківськ, Україна
nadia14061996@gmail.com

Analysis of the Modern Realization of the Acute Emission Method

Nadiia Ivasyshyn
dept. of Computer Science
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine
nadia14061996@gmail.com

Анотація – В роботі розглянуто сучасний стан методу акустичної емісії, проаналізована оцінка та завдання цього методу, здійснено огляд існуючої акустично-емісійної апаратури. Оскільки метод АЕ-контролю може успішно застосовуватись для виявлення та локалізації місць протікання промислових та технологічних трубопроводів, така робота є особливо актуальною задачею в наш час.

Abstract – In this work the present state of the acoustic emission method is considered, the estimation and tasks of this method are analyzed, an analysis of existing of acoustic emission equipment is carried out. Because, the AE-control method can be successfully used to identify and locate the leaks locations of industrial and technological pipelines, which is a particularly urgent task in our time.

Ключові слова – акустична емісія, неруйнуючий контроль, метод акустичної емісії, дефекти, тріщини.

Keyword s– acoustic emission, non-destructive control, acoustic emission method, leaks.

I. ВСТУП

Метод акустичної емісії (АЕ) заснований на реєстрації і аналізі акустичних хвиль, що виникають в процесі пластичної деформації і руйнування (зростання тріщин) контрольованих об'єктів. Щодо діагностування трубопроводів, то однією з причин зміни технічного стану трубопроводів є поява витоків, які зумовлені виникненням і розвитком дефектів, метеорологічними умовами, пошкодженням або неякісним проведенням ремонтних робіт, просіданням ґрунту чи несанкціонованими врізаннями. В зв'язку з цим метод АЕ-контролю може успішно застосовуватись для виявлення та локалізації місць протікання промислових та технологічних трубопроводів, що є особливо актуальною задачею в наш час.

II. ЗАВДАННЯ МЕТОДУ КОНТРОЛЮ ЗА АКУСТИЧНОЮ ЕМІСІЄЮ

Завданням контролю за акустичної емісією є виявлення, визначення координат і моніторинг за джерелами акустичної емісії, пов'язаними з несучільностями на поверхні або в об'ємі стінки посудини, зварного з'єднання і виготовлених частин і компонентів [4].

Неруйнуючий контроль (НК) – контроль властивостей і параметрів об'єкту, при якому не повинна бути порушена придатність об'єкту до використання і експлуатації.

При виявленні сигналу на 1-му і 2-му приймачах, реєструється час приходу сигналу t_1 і t_2 відповідно. По зареєстрованих t_1 і t_2 визначається різниця часу приходу сигналу $\Delta t = t_2 - t_1$. Потім по відомій швидкості хвилі в матеріалі і відомим координатам приймачів обчислюються координати джерела акустичної емісії [1].

III. ПОРЯДОК ЗАСТОСУВАННЯ МЕТОДУ АКУСТИЧНОЇ ЕМІСІЇ:

- АЕ контроль проводять у всіх випадках, коли він передбачений нормативно-технічними документами або технічною документацією на об'єкт.
- АЕ контроль проводять у всіх випадках, коли нормативно-технічною документацією на об'єкт передбачено проведення неруйнуючого контролю одним з методів, що регламентуються, але по технічних або інших причинам проведення такого контролю неможливе.

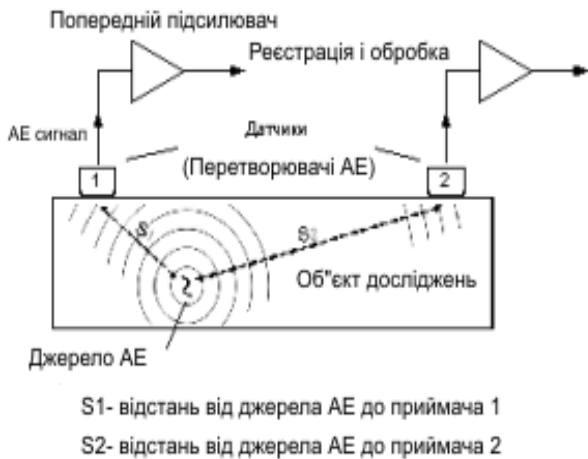


Рис. 1 – Схема методу акустичної емісії

IV. ОЦІНКА РЕЗУЛЬТАТІВ АЕ КОНТРОЛЮ

Після обробки прийнятих сигналів результати контролю представляють у вигляді ідентифікованих і класифікованих джерел АЕ [4]. Дії персоналу, що рекомендуються, виконуючого АЕ контроль при виявленні джерел АЕ того або іншого класу, наведені в табл. 1.1:

ТАБЛИЦЯ 1. Порядок дій персоналу при виявленні джерел АЕ

Джерело	Клас	Дії, що рекомендуються
Пасивний	I	реєструють для аналізу динаміки його подальшого розвитку
Активний	II	1. реєструють і стежать за розвитком ситуації в процесі виконання даного контролю; 2. відзначають в звіті і записують рекомендації по проведенню додаткового контролю з використанням інших методів.
Критично активний	III	1. реєструють і стежать за розвитком ситуації в класу процесі виконання даного контролю; 2. роблять заходи по підготовці можливого скидання навантаження.
Катастрофічно активний	IV	1. проводять негайне зменшення навантаження до 0, або до величини, при якій клас джерела АЕ знизиться до рівня II або III класу; 2. після скидання навантаження проводять огляд об'єкту і при необхідності контроль іншими методами.

V. ОСОБЛИВОСТІ МЕТОДУ КОНТРОЛЮ ЗА АКУСТИЧНОЮ ЕМІСІЮ

Основними перевагами методу акустичної емісії перед традиційними методами неруйнуючого контролю є наступні:

Інтегральність методу, яка полягає в тому, що, використовуючи один або декілька датчиків, встановлених нерухомо на поверхні об'єкту, можна проконтролювати весь об'єкт цілком (100% контроль).

На відміну від скануючих методів неруйнуючого контролю, метод АЕ не вимагає ретельної підготовки поверхні об'єкту контролю.

Виявлення і реєстрація дефектів, що дозволяє класифікувати дефекти не по розмірах, а по ступеню їх небезпеки для контрольованого об'єкту.

Висока продуктивність, що у багато разів перевершує продуктивність традиційних методів неруйнуючого контролю, таких як ультразвуковий, радіографічний, вихреструммовий, магнітний і ін.

Дистанційність методу – можливість проведення контролю при значному віддаленні оператора від досліджуваного об'єкту.

Можливість відстежування різних технологічних процесів і оцінка технічного стану об'єкту в режимі реального часу.

VI. ІСНУЮЧІ АКУСТИЧНО-ЕМІСІЙНІ СИСТЕМИ

Малахіт АС-15А

Основні переваги у використанні системи [3] :

- Реєстрація і відображення форми хвилі або її спектру в кожному каналі в реальному часі
- Спільна робота приладових блоків в будь-якому поєднанні з довільною кількістю каналів
- Амплітудний діапазон 100 дБ (динамічний амплітудний діапазон 86дБ з дискретністю вимірювання 0,1 дБ у всьому діапазоні)
- Фіксовані або слідкуючі рівні порогів у всьому діапазоні амплітуд
- Режим автоматичного тестування датчиків
- Невеликі габарити і маса
- Зовнішній вигляд системи наведений на рис .



Рис. 2 – Система «Малахіт AC-15A»

Програма «Малахіт»

Програма «Малахіт» входить до складу базового пакету програм AE Studio. «Малахіт» є основною програмою збору і обробки даних акустико-емісійних систем. Системи «Малахіт AC-12A» і «Малахіт AC-15A» працюють під управлінням операційних систем Windows XP-Windows 10 (32-розрядні і 64-розрядні)[3].

Переваги :

- Робота в автономному режимі (без комп'ютера) або в режимі акустико-емісійної системи
- Вбудована пам'ять для зберігання до 63 результатів випробувань
- Зв'язок з комп'ютером по USB інтерфейсу
- Звукова і світлова індикація перевищення сигналом порогового рівня
- Інтегральні датчики з режимом випромінювача і автотестом

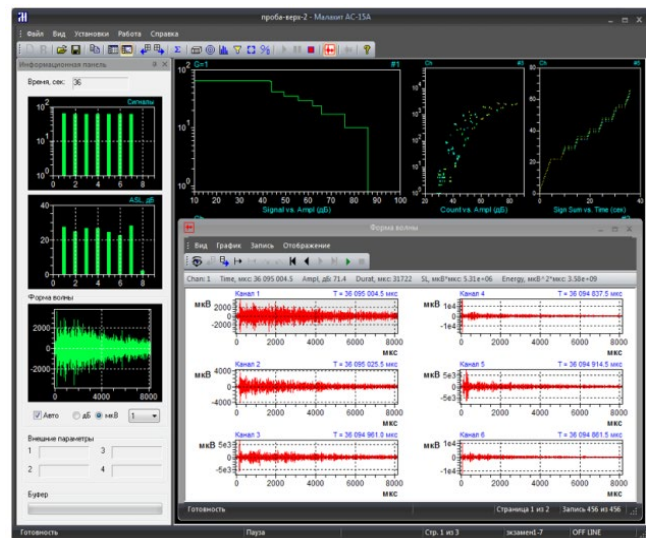


Рис. 3– Програма «Малахіт». Головне вікно.

ВИСНОВОК:

У даній роботі проаналізовано сучасний стан систем акустико-емісійного методу виявлення тріщин.

Широкі можливості методу контролю акустичної емісії надає і при дослідженні різних властивостей матеріалів, речовин, конструкцій. На сьогоднішній день без застосування акустичного контролю і моніторингу вже неможливі створення і надійна експлуатація багатьох відповідальних технічних об'єктів.

ЛІТЕРАТУРА REFERENCES

- [1] Грешников В.А., Дробот Ю.Б. Акустическая эмиссия. – М.: Машиностроение, 1976. – 128с.
- [2] Ермолов И.Н. Неразрушающий контроль. – Кн.2. Акустические методы контроля. – М.: Высшая школа, 1991. – 283 с.
- [3] Євчук О.В. Пристрій для дослідження структурних змін в металевих конструкціях за відносною приведеною частотою гармонік сенсора-струни / О.В. Євчук, В.А. Ровінський, Ю.Й. Стрілецький // Quality Control Tools and Techniques. – 2010 – №1(40). – С. 29-36
- [4] Євчук О.В. Швидкий алгоритм фільтрації для систем діагностування об'єктів нафтогазового комплексу/ О.В. Євчук, В.А. Ровінський // Івано-Франківськ: ІФНТУНГ, 2006 – 208 с
- [5] Євчук О.В., Ровінський В.А., Стрілецький Ю.Й., Фуфалько О.Ю., Інформаційна технологія моделювання біполярних транзисторів хвиловими цифровими фільтрами, Вінниця: ВНТУ, 2014
- [6] Заміховський Л.М. Експериментальне визначення місць витоку повітря на технологічному трубопроводі / Л.М. Заміховський, Л.О., Штаєр, В.А. Ровінський., // Quality Control Tools and Techniques. – 2010 – №1(24). – С. 73-77
- [7] Неразрушающий контроль и диагностика / под ред. В.В. Клюева. – М.:Машиностроение, 2005. – 656с.
- [8] Описание метода акустической эмиссии // <http://www.avek.ru/pages/23/>
- [9] ПБ 03-593-03. Правила организации и проведения акустико-эмиссионного контроля сосудов, аппаратов, котлов и технологических трубопроводов. – М. – 2003.

Моделювання Топологічних Структур Мереж Логістики з Реверсивними Потокami

Світлана Мальцева
кафедра системотехніки
Харківський національний
університет радіоелектроніки
Харків, Україна
svitlana.maltseva@nure.ua

Аліна Зубенко
кафедра системотехніки
Харківський національний
університет радіоелектроніки
Харків, Україна
zubenkoalina.az@gmail.com

Володимир Безкоровайний
кафедра системотехніки
Харківський національний
університет радіоелектроніки
Харків, Україна
vladimir.beskorovainyi@nure.ua

Modelling of Topological Structures for Logistics Networks with Reversible Flows

Svitlana Maltseva
Department of System Engineering
Kharkiv National University of
Radioelectronics
Kharkiv, Ukraine
svitlana.maltseva@nure.ua

Alina Zubenko
Department of System Engineering
Kharkiv National University of
Radioelectronics
Kharkiv, Ukraine
zubenkoalina.az@gmail.com

Volodymyr Beskorovainyi
Department of System Engineering
Kharkiv National University of
Radioelectronics
Kharkiv, Ukraine
vladimir.beskorovainyi@nure.ua

Анотація— Розглянуто задачу оптимізації топологічних структур тривірневих мереж логістики з прямими і реверсивними потоками та рознесеними центрами виробництва і переробки. Запропоновано модифікації моделей для оцінки показників витрат, часу транспортування потоків та аналізу динаміки перехідних процесів.

Abstract—The problem of optimization of topological structures of three-level networks of logistics with direct and reverse flows and dilapidated production and processing centres is considered. Modifications of the models for estimation of expenses, time of transport of flows and analysis of dynamics of transients are offered.

Ключові слова—логістична мережа, структура, топологія, оптимізація, модель.

Keywords—logistics network, structure, topology, optimization, model.

I. ВСТУП

Зростання обсягів споживання в усіх сферах людської діяльності призводить до відповідного зростання обсягів тари та відходів, що вимагають повернення до компаній-виробників чи пунктів утилізації, відновлення або рециклінгу. В наш час задачі оптимізації схем прямих і зворотних потоків є предметом дослідження екологічної (зеленої) та реверсивної логістики [1]. Метою створення

мереж реверсивної логістики є забезпечення умов екологічно та економічно ефективного повторного використання відходів продукції і тари (упаковки) у виробничому циклі [2]. Характерною особливістю задач оптимізації таких мереж є неповна визначеність даних щодо кількості та якості товарів, продуктів, тари тощо, що повертаються від споживача до центрів переробки [3]. На практиці це може призводити до суттєвої нестаціонарності прямих і зворотних потоків та, як наслідок, до дефіцитів або затоварювань в місцях виробництва, зберігання, реалізації й утилізації. Для забезпечення ефективного функціонування логістичних мереж зі зворотними потоками необхідна їх структурна, топологічна, параметрична та технологічна оптимізація, що є складною науково-прикладною проблемою.

II. АНАЛІЗ СУЧАСНОГО СТАНУ ПРОБЛЕМИ

Логістичні мережі являють собою територіально розподілені об'єкти, ефективність яких багато в чому визначається їх топологією (територіальним розміщенням елементів і зв'язків між ними) [4]. Традиційно ефективність систем логістики оцінюється за показниками витрат та часу на доставку вантажів заданого обсягу. Задачі їх оптимізації у найбільш загальному випадку передбачають визначення найкращих місць розташування центрів виробництва (ЦВ) і переробки (ЦП), кількості терміналів (складів, хабів) Т, місць їх розташування, підмножин кінцевих споживачів (КС), що обслуговуються

кожним з терміналів, типів терміналів та транспортних засобів.

Ефективність топологічної структури логістичної мережі залежить від топології КС, розташування ЦВ і ЦП, обсягів прямих і реверсивних потоків (рис. 1) [5]. За результатами системологічного аналізу та декомпозиції проблеми оптимізації таких об'єктів встановлено: тісний взаємозв'язок виділених підзадач; комбінаторний характер більшості з них; присутність суттєвої невизначеності вхідних даних; практичну непридатність для їх аналізу класичних аналітичних методів дослідження [6].

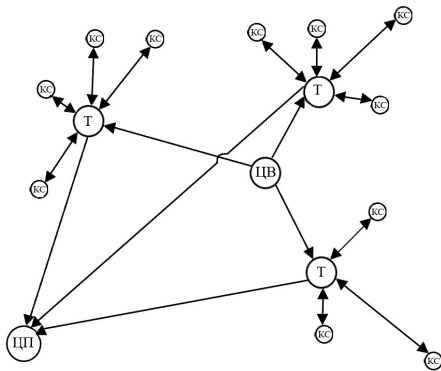


Рис. 1. Структура мережі з прямими та реверсивними потоками.

Традиційно задачі оптимізації логістичних мереж розв'язуються в два етапи окремо для прямих і зворотних потоків [2, 3, 5]. Одним з найбільш ефективних за комплексним показником «точність-складність» методом розв'язання задачі структурно-топологічної оптимізації трирівневих однонаправлених мереж за критерієм витрат вважається метод спрямованого перебору [7]. Для аналізу часових характеристик мереж з великою кількістю КС та їх динамічних властивостей ефективними виявляються методи системної динаміки [8].

Незважаючи на численні публікації, присвячені розв'язанню задач проектування, оптимізації логістичних мереж та управління потоками, потребують подальшого дослідження задачі оптимізації комбінованих мереж з прямими і зворотними потоками, включаючи аналіз їх динамічних властивостей.

III. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

За результатами аналізу особливостей комплексу проблемно пов'язаних задач визначено вимоги, яким повинні задовольняти ефективні процедури їх розв'язання: тісний взаємозв'язок задач синтезу й аналізу, їх неповна інформаційна визначеність обумовлює ітераційний характер методів і процедур їх розв'язання; висока часова й ємнісна складність методів, обумовлена комбінаторним характером задач і широкий діапазон їх розмірності передбачають використання множини методів з різною складністю і точністю; з метою використання досвіду проектувальників, менеджерів і врахування чинників, що важко піддаються формалізації процеси оптимізації мереж доцільно будувати на основі

інтерактивних (людино-машинних) процедур; на всіх етапах доцільно використовувати прийоми, що знижують трудомісткість розв'язання задач системного синтезу й аналізу.

Витрати на централізовану трирівневу логістичну мережу з прямими і зворотними потоками (рис. 1) можна подати у вигляді суми наведених витрат на доставку прямого потоку від центру виробництва до терміналів C_{BT} , на обробку в терміналах C_T , на доставку прямого потоку від терміналів до кінцевих споживачів C_{TC} , на доставку зворотного потоку від кінцевих споживачів до терміналів C_{CT} , на доставку зворотного потоку від терміналів до центру переробки C_{TP} у вигляді:

$$C = C_{BT} + C_T + C_{TC} + C_{CT} + C_{TP}. \quad (1)$$

Оптимальній за критерієм витрат топологічній структурі мережі відповідає мінімум функції (1). При цьому схеми зв'язків споживачів, терміналів і центрів встановлюються по мінімуму вартості транспортування потоків. З урахуванням різних варіантів можливого розміщення n_T терміналів залежність функції витрат (1) від кількості терміналів мережі $C(n_T)$ є багатоекстремальною. Локальні мінімуми функції (1) відповідають оптимальному розв'язку задачі розміщення фіксованої кількості терміналів $n_T = 1, 2, 3, \dots$. Кількість можливих варіантів розміщення n_T терміналів на множині з n_C місць дорівнює числу сполучень з n_C по n_T .

Витрати на створення й (або) експлуатацію терміналів $C_T(n_T)$ зростають зі збільшенням їх кількості n_T . Зі збільшенням кількості терміналів в мережі n_T витрати на транспортування потоків між ними та центрами $C_{BT}(n_T) + C_{TP}(n_T)$ збільшуються, а витрати на транспортування потоків між кінцевими споживачами і терміналами $C_{TC}(n_T) + C_{CT}(n_T)$ знижуються. Їх сума являє собою однокоекстремальну функцію $C_{BTCTP}(n_T)$ [9]. З урахуванням цього обвідна локальних мінімумів функції (1) при фіксованих значеннях кількості терміналів n_T у діапазоні від 1 до n_C буде також однокоекстремальною залежністю (рис. 2). Це дозволяє використати для структурно-топологічної оптимізації логістичних мереж такого виду ефективні модифікації методу спрямованого перебору варіантів за кількістю вузлів у мережі $n_T = 1, 2, \dots, n_T^o, n_T^o + 1$ [10].

Час транспортування потоків для s -го варіанту побудови мережі $\tau_i'(s)$ від центру виробництва до кожного з кінцевих споживачів $\tau_i(s)$, $i = \overline{1, n_C}$ можна подати у вигляді суми відрізків часу перебування прямого потоку на кожному з етапів ланцюжка: «центр виробництва - термінал»; «термінал»; «термінал - кінцевий споживач»:

$$\tau_i^1(s) = \tau_i^{BT}(s) + \tau_i^T(s) + \tau_i^{TC}(s), \quad i = \overline{1, n_c}, \quad (2)$$

де $\tau_i^{BT}(s)$, $\tau_i^T(s)$, $\tau_i^{TC}(s)$ - відповідно час перебування прямого потоку до i -го споживача на етапах «центр виробництва - термінал»; «термінал»; «термінал - кінцевий споживач».

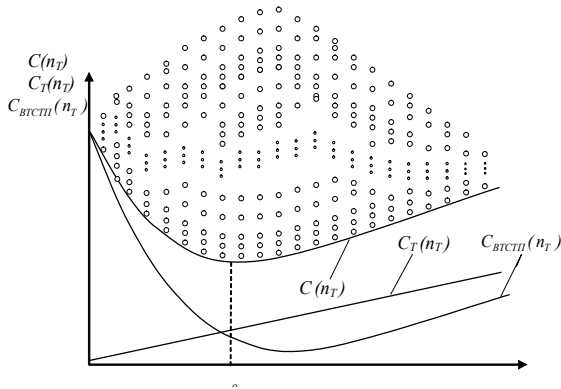


Рис. 2. Обвідна локальних мінімумів функції витрат $C(n_T)$ (1).

Час транспортування потоків для s -го варіанту побудови мережі $\tau_i^2(s)$ від кожного з кінцевих споживачів до центрів переробки $\tau_i(s)$, $i = \overline{1, n_c}$ подамо у вигляді суми відрізків часу перебування прямого потоку на кожному з етапів ланцюжка: «кінцевий споживач - термінал»; «термінал»; «термінал - центр переробки»:

$$\tau_i^2(s) = \tau_i^{CT}(s) + \tau_i^T(s) + \tau_i^{TP}(s), \quad i = \overline{1, n_c}, \quad (3)$$

де $\tau_i^{BT}(s)$, $\tau_i^T(s)$, $\tau_i^{TC}(s)$ - відповідно час перебування зворотного потоку від i -го споживача на етапах «кінцевий споживач - термінал»; «термінал»; «термінал - центр переробки».

Визначення динаміки перехідних процесів здійснюється з використанням модифікованої моделі виробничо-збутової системи [8]. Суть модифікації полягає у відображенні множини зворотних потоків на шляху «кінцевий споживач - термінал»; «термінал»; «термінал - центр переробки». З метою спрощення моделі: відносно невеликі за значенням запізнювання пересилання замовлень з усіх ланок мережі не враховуються; лінійні ланцюжки запізнювань у процесах формування замовлень в усіх ланках мережі подаються як одне загальне запізнювання; запізнювання у паралельних каналах руху замовлень, товарів, тари тощо об'єднуються в одному каналі та подаються одним запізнювання у загальному каналі.

Для отримання залежності темпів вихідних потоків від вхідних на всіх етапах логістичної мережі для моделі динаміки необхідно за методом найменших квадратів визначити порядки та постійні відповідних запізнень (рис. 3):

$$K = \sum_{j=1}^m \{y(t_j) - F_M[u_M(t_j), q_M]\}^2 \rightarrow \min_{q_M}, \quad (4)$$

де j - номер часової точки оцінювання потоку; m - кількість точок оцінювання; $y(t_j)$ - значення темпу потоку в момент часу t_j , $j = \overline{1, m}$; F_M - оператор моделі показникового запізнювання; $u_M(t_j)$ - темп вхідного потоку для моделі; q_M - вектор параметрів моделі (порядок та стала запізнювання).

ВИСНОВКИ

Організація логістичних процесів сучасних виробничо-збутових комплексів все частіше передбачає спільне розв'язання задач оптимізації мереж потоків товарів і мереж потоків тари та відходів, що потребують переробки. Для систем з рознесеними центрами виробництва і переробки запропоновано модифікації моделей для оцінки показників витрат, часу транспортування потоків та аналізу динаміки перехідних процесів, що дозволять підвищити функціонально-вартісні показники комбінованих логістичних мереж.

REFERENCES ЛІТЕРАТУРА

- [1] Л.М. Гурч, Л.С. Хмара, «Розвиток «зеленої логістики» в Україні», *Вісник Національного університету «Львівська політехніка». Логістика*, № 811, с. 86–91, 2014.
- [2] Э.М. Букринская, *Реверсивная логистика*, Санкт-Петербург: СПбГУЭФ, 79 с., 2010.
- [3] М.В. Нагай, А.О. Овсянникова, А.И. Сальникова, «Особенности управления возвратными потоками», *Логистические системы в глобальной экономике*, №1, с. 315–319, 2011.
- [4] Э.Г. Петров, В.П. Пискалова, В.В. Бескорвайный, *Территориально распределенные системы обслуживания*, Киев: Техника, 208 с., 1992.
- [5] Д.А. Карх, С.В. Потапова, «Некоторые теоретические и прикладные аспекты возвратной логистики», *Известия УрГЭУ*, №2, с. 118–122, 2012.
- [6] В.В. Бескорвайный, «Системологический анализ проблемы структурного синтеза территориально распределенных систем», *Автоматизированные системы управления и приборы автоматики*, Вып. 120, с. 29–37, 2002.
- [7] В.В. Бескорвайный, З.А. Имангулова, А.И. Петрова, «Оптимизация количества и местоположения распределительных центров транспортно-складской системы», *Восточно-Европейский журнал передовых технологий*, № 5/3 (59), с. 24–28, 2012.
- [8] Дж. Форрестер, *Основы кибернетики предприятия (Индустриальная динамика)*, М.: Прогресс, 340 с., 1971.
- [9] Э.Г. Петров, А.Б. Болотов, В.В. Бескорвайный, «Алгоритм структурно-топологической оптимизации централизованных сетевых систем», *Механизация и автоматизация управления*, №1, с. 28–31, 1986.
- [10] V. Beskorovainyi, K. Podoliaka, «Reengineering the topological structure of large-scale monitoring systems», *ECONTECHMOD*, Vol. 4 (3), pp. 13–18, 2015.

Спосіб Формування Доменів Маршрутизації в SDN

Ю.О. Кулаков
каф. ОТ, ф-тет інформатики та
обчислювальної техніки
НТУУ «КПІ ім. І. Сікорського»
м.Київ, Україна
ya.kulakov@gmail.com

А.В. Коган
каф. АСОІУ, ф-тет інформатики
та обчислювальної техніки
НТУУ «КПІ ім. І. Сікорського»
м.Київ, Україна

Є.О. Павленкова
каф. АСОІУ, ф-тет інформатики
та обчислювальної техніки,
НТУУ «КПІ ім. І. Сікорського»
м.Київ, Україна
pavlenkovayevheniia@gmail.com

The Method of Forming Routing Domains in SDN

Yurii Kulakov,
Department of Computing Engineering,
Kyiv, Ukraine
ya.kulakov@gmail.com

Alla Kohan,
Department of Technical Systems of
Automation and Control of Production,
Kyiv, Ukraine
a.v.kohan433@gmail.com

Yvheniia Pavlenkova
Department of Technical Systems of
Automation and Control of Production,
Kyiv, Ukraine
pavlenkovayevheniia@gmail.com

Анотація—У даній роботі розглянуто задачу оптимізації маршрутизації в мережі шляхом розбиття її на домени маршрутизації з визначенням місця розташування контролера SDN. Представлені основні критерії, за якими проводиться розбиття графа на домени маршрутизації.

Ключові слова—SDN, контролер SDN, маршрутизація, домени

I. ВСТУП

Сучасна тенденція збільшення розмірності комп'ютерних мереж призводить до збільшення обсягу трафіку, який передається (для кожного типу трафіку висуваються свої вимоги до якості обслуговування (QoS)), а процес управління мережею стає все складнішим і менш контрольованим. Для вирішення такої задачі необхідно використовувати новий підхід, що дозволить контролювати всі процеси в мережі одночасно. В вирішенні даної задачі може допомогти SDN (Software-Defined Networking), такий підхід полягає у відділенні функцій передачі трафіку від функцій управління. У традиційних комутаторах і маршрутизаторах ці процеси невіддільні один від одного і реалізовані в одному пристрої. Згідно з концепцією SDN, вся логіка управління вноситься в так звані контролери, які здатні відслідковувати роботу всієї мережі.

При організації мережі великої розмірності, контроль буде здійснюватися декількома контролерами, у зв'язку з цим виникає необхідність в розподілі SDN-мережі на оптимальну кількість доменів маршрутизації, що дозволить контролерам SDN здійснювати миттєве реагування на зміни в топології мережі та досягти рівномірного завантаження в кожному домені та мережі в цілому.

Враховуючи новизну концепції SDN, її стрімкий розвиток і впровадження вимагають удосконалення способів та методів маршрутизації і адаптацію їх під нові умови. Задача маршрутизації в SDN вирішується централізовано на спеціальних серверах, які є більш продуктивними, ніж традиційні IP-маршрутизатори. Це дозволяє використовувати більш ефективні і одночасно більш складні з обчислювальної точки зору протоколи маршрутизації. На сьогоднішній день протоколи маршрутизації з метою задоволення QoS розробляються

для балансування навантаження в мережі з організацією багатопрохідної маршрутизації [1, 2]. SDN-контролер дозволить підвищити ефективність використання протоколів маршрутизації, а саме дистанційно-векторних та по стану каналу в комутаторах MPLS.

Забезпечення балансування навантаження дозволяє оптимізувати рішення завдання маршрутизації і ефективно використовувати ресурси мережі, в результаті чого поліпшуються значення ключових показників QoS

II. ПОСТАНОВКА ЗАДАЧІ

Ключовим принципом SDN мережі є розподіл процесів передачі та управління даними. На відміну від традиційного уявлення мережі, SDN-контролер забезпечує моніторинг, доступ, управління ресурсами всієї мережі, а не конкретного вузла. Також важливою задачею залишається питання про місцезнаходження контролера в домені, який володіє інформацією про топологію мережі в будь-який момент часу і містить інформацію про розміщення користувачів і серверів, інших елементів і сервісів мережі, а крім того, прив'язку між іменами та адресами.

III. ФОРМУВАННЯ ДОМЕНІВ МАРШРУТИЗАЦІЇ

В процесі функціонування комп'ютерної мережі реконфігурація доменів здійснюється з метою скорочення часу передачі інформації та оптимізації мережевого трафіку. Як критерій вибору структури і розміру доменів служить ефективність передачі даних, яка дорівнює [3]:

$$K_{ef} = \frac{Q_n}{Q_n + Q_{n\bar{e}}}$$

де K_{ef} – коефіцієнт ефективності передачі даних.

Основні фактори, що впливають на збільшення обсягу службового трафіку в процесі реконфігурації домену:

- частота реконфігурації домену;
- кількість вузлів в домені;
- інформація, передана одним вузлом в момент реконфігурації.

Отже, для зменшення обсягу службового трафіку в домені, частота реконфігурації на заданому проміжку часу і кількість вузлів домену має прагнути до мінімуму:

$$F_r \rightarrow \min, V_0 \rightarrow \min$$

де F_r – частота реконфігурації домену; V_0 – інформація, передана одним вузлом в момент реконфігурації.

Таким чином, основними параметрами, що впливають на вибір розміру і структури домену, є:

- кількість вузлів в домені;
- кількість реконфігурацій домену;
- діаметр домену;
- ступінь зв'язності вузлів домену.

Отже, завдання розбиття графа на підграфи зводиться до задачі представлення вихідного графа $G=(V,E)$ у вигляді множини підмножин вершин V_1, V_2, \dots, V_r ($r>1$), таких, що $V_1 \cup V_2 \cup \dots \cup V_r = V$, $V_k \cap V_l = \emptyset$, $k, l = 1, \dots, k \neq l$.

Представимо основні критерії, за якими проводиться розбиття графа на домени маршрутизації [4]:

- щільність домену

$$h(G_k) = \frac{2|E_k|}{|V_k|(|V_k|-1)}$$

$$h(G_k) = \max\{s_{ij}\}, i \in V_k$$

де s_i – степінь вершин $i \in V_k$ домена $G_k=(V_k, E_k)$; діаметр $d(G_k)$ домена $G_k=(V_k, E_k)$, сума ваг вершин $g(V_k)=\sum_{i \in V_k} w_i$ в кожному домені повинні бути близькі один до одного:

$$\max_{1 \leq k \leq r} \{g(V_k)\} - \min_{1 \leq k \leq r} \{g(V_k)\} \rightarrow \min$$

Маршрутизація від однієї вершини до іншої повинна складатися з маршрутизації всередині домену SDN і маршрутизації між ними.

Відомі методи [5] формування доменів і вибору місця розташування їх контролерів малоефективні для мереж, структура яких постійно змінюється. В першу чергу це пов'язано з тим, що відомі методи не забезпечують вибору оптимального місця розташування контролера домену. Це, як правило, призводить до необхідності більш частішої реконфігурації. Зі збільшенням числа реконфігурацій домену, обсяг службового трафіку в домені збільшується по нелінійному закону, за рахунок чого різко падає ефективність передачі даних.

Після розбиття мережі на домени маршрутизації, важливо визначити місце розташування SDN-контролера в ньому. Також визначити ступінь зв'язності всіх вершин графа, так як контролер повинен мати максимальну зв'язність з усіма вузлами для повного контролю.

Крім того, при виборі місця розташування SDN-контролера, необхідно враховувати ймовірність вибору вершини SDN-контролером, яку представимо у вигляді:

$$K_l = K_s \cdot \delta_k(v_m) + K_a \cdot a_i,$$

де K_l – коефіцієнт вибору SDN-контролера;

K_s – ваговий коефіцієнт ступеня;

K_a – ваговий коефіцієнт a_i ;

$\delta_k(v_m)$ – щільність мережевого оточення k -го порядку(1).

На відміну від існуючих методів визначення місця розташування SDN контролера a_i , для вершини $v_m \in V$ визначимо також щільність мережевого оточення k -го порядку $\delta_k(v_m)$.

$$\delta_k(v_m) = \frac{S(v_m) + \sum_{v_i, v_j \in \Gamma^k(v_m)} e_{ij}}{S(v_m)} \quad (1)$$

де $S(v_m)$ – максимальна ступінь зв'язності вершини v_m ;
 $\Gamma(V_m)$ – відповідність першого порядку, множині всіх вершин, пов'язаних з вершиною v_m , $\Gamma^2(v_m) = \Gamma(\Gamma(v_m))$;
 e_{ij} – зв'язок між суміжними вершинами v_i і v_j з максимальним ступенем;
 k – порядок мережевого оточення вершини v_m .

Оскільки значне навантаження роботи всієї мережі випадає на SDN контролер, необхідно враховувати безпеку та надійність даного вузла в мережі, та додатково використовувати програмні засоби шифрування інформації для захисту від зловмисників. Також виникає необхідність у визначенні не тільки основного SDN-контролера а й резервного, який зможе продовжити підтримувати працездатність мережі у випадку відмови основного контролера.

ВИСНОВОК

В роботі розглянута задача розбиття графа на оптимальну кількість доменів маршрутизації та визначені основні критерії розбиття мережі на зони. Також розглянута задача визначення місця розташування контролера маршрутизації. SDN-контролер дозволить оптимізувати навантаження в комп'ютерних мережах за рахунок розділення процесів передачі та управління даними, а так само забезпечить максимально ефективне використання системних ресурсів.

REFERENCES ЛІТЕРАТУРА

- [1] Zhaogang Shu. Traffic Engineering in Software-Defined Networking: Measurement and Management / Zhaogang Shu, Jiafu Wan, Jiaxiang Lin, // Special section on green communications and networking for 5G wireless. - 2016. - Volume 4. – Pp. 3246- 3256.
- [2] Кулаков Ю. А. Способ формирования множества непересекающихся путей от одной к нескольким вершинам в компьютерной сети большой размерности / Кулаков Ю. А., Коган А. В., Диброва М. А. / Міжнародна наукова конференція імені Т. А.Таран «Інтелектуальний аналіз інформації ІАІ-2016», Київ, 18-20 травня 2016р. – С. 104-107.
- [3] Давиденко И.Н. Способ определения количества и местоположения агентов домена в мобильных сетях / И.Н. Давиденко // Вісник Інженерної академії України. – 2008. – № 1. – С. 53–58.
- [4] Kohan. Innovative solutions in modern science / A. Kohan, Y. Kulakov, M. Sperkach, O. Zhdanova // Technical sciences formalization of the problem of the SDN network partitioning into routing zones. - 2018. - No. 3(22). - Pp 36-45.
- [5] Li Y. Constructing Connected Dominating Sets with Bounded Diameters in Wireless Networks / WASA 2007 : International Conference on Wireless Algorithms, Systems and Applications, 1-3 Aug. 2007. : thesis rep. – Chicago (USA), 2007. – P. 89–94.

Платформа для Локалізації Веб та Мобільних Додатків

Смольц Дмитро Олександрович
Факультет інформаційних технологій та комп'ютерної інженерії
Вінницький національний технічний університет
Вінниця, Україна
dima.smolts@gmail.com

Мельник Жанна Анатоліївна
Факультет інформаційних технологій та комп'ютерної інженерії
Вінницький національний технічний університет
Вінниця, Україна
lki.15bmelnyk@gmail.com

Platform for Localization of the Web And Mobile Applications

Dmytro Smolts
Faculty of Information Technologies and Computer Engineering
Vinnytsia National Technical University
Vinnytsia, Ukraine
dima.smolts@gmail.com

Zhanna Melnyk
Faculty of Information Technologies and Computer Engineering
Vinnytsia National Technical University
Vinnytsia, Ukraine
lki.15bmelnyk@gmail.com

Анотація – У статті описано актуальність локалізації, а також існуючі способи та шляхи її реалізації. Запропоновано універсальну платформу для локалізації веб та мобільних додатків на основі багатосарової клієнт-серверної архітектури.

Ключові слова: локалізація, переклад, глобалізація, сервіс, лінгвістика, програмування, платформа, мережа.

Abstract – The article describes the relevance of localization, as well as known methods and ways of its implementation. A universal platform for localizing web and mobile applications based on multi-layered client-server architecture is proposed.

Keywords: localization, translation, globalization, service, linguistics, programming, platform, network.

I. ВСТУП

Сучасні інформаційні технології увібрали в себе лавиноподібні досягнення електроніки, а також математики, філософії, психології та економіки. В умовах сучасного темпу розвитку ІТ суміжні сфери отримують поштовх до розповсюдження та популяризації:

A. Наукова діяльність

Вчені по всьому світу отримують способи для комунікації, новітні програмні та апаратні інструменти для проведення досліджень, моделювання експериментальних умов, здійснення нових відкриттів.

B. Освітні програми

Студенти та учні мають доступ до навчальних матеріалів із невичерпних джерел інформації з будь якої частини землі, що призводить до підвищення їх кваліфікації в будь-якій сфері.

C. Бізнес

Люди мають змогу розвивати свій бізнес та робити продукт своєї діяльності доступним за межами власної країни.

Таких сфер діяльності як і прикладів використання інформаційних технологій у них можна назвати безліч і очевидним є те, що з ІТ людство прогресує у набагато швидшому темпі.

II. АКТУАЛЬНІСТЬ ЛОКАЛІЗАЦІЇ

На даний момент, кількість нативних носіїв мови у світі не співпадає із кількістю продуктів цією мовою, що видно з діаграм, представлених на рисунках 1 і 2. На це є декілька причин:

- розвиненість країни, де спілкуються цією мовою;
- складність мови;
- освіченість населення і т.д.

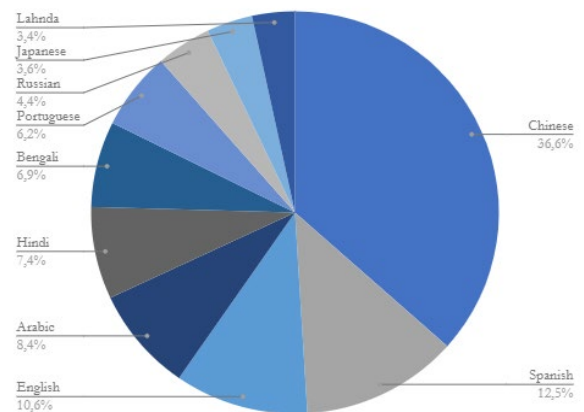


Рис. 1. Популярність мови у світ

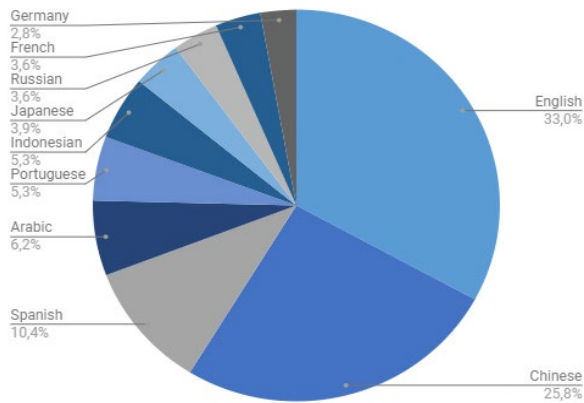


Рис. 2. Популярність мови в інтернеті

На жаль, це призводить до труднощів та незручностей. На дану тему було проведено багато досліджень, з метою визначити на скільки комфортніше сприймати інформацію своєю мовою. Ось деякі висновки з них:

- 80% користувачів, які користуються мобільними додатками не англomовні [1];
- локалізовані версії сайтів та додатків мають на 42% більший коефіцієнт кліків та на 22% більше переходів ніж одномовні [2];
- 72% покупців стверджують, що вони здійснюють покупки лише у тих магазинах, де є опис товару їхньою мовою (Common Sense Advisory) [3];
- 56% покупців вважають наявність перекладу на їхню мову важливішою за ціну продукту (Common Sense Advisory) [3].

Наочно продемонструвати актуальність локалізації певного сайту можна за допомогою формули 1:

$$T = \frac{a_u - a_n}{k_w} \quad (1)$$

де a_u – це загальна кількість користувачів цього сайту,

a_n – це кількість носіїв мови сайту серед користувачів,

k_w – це коефіцієнт популярності мови в світі, який в свою чергу визначається за формулою 2:

$$k_w = \frac{L}{100} \quad (2)$$

де L – популярність мови у світі, яку можна побачити на рисунку 1.

Беручи до уваги формулу 1, можна зрозуміти, що якщо актуальність локалізації напряму залежить від кількості іноземних користувачів цього сайту та обернено від популярності мови у світі.

Зважаючи на це, стає зрозуміло, що локалізація – це невід’ємна частина успіху програмного продукту.

III. СПОСОБИ ТА ШЛЯХИ ЛОКАЛІЗАЦІЇ

Найпростіший варіант – створити декілька HTML сторінок (для кожної мови окремо) та додати можливість перемикати мову переходячи на іншу сторінку. Даний метод дуже простий і не потребує ніяких особливих навичок. Але він зовсім не оптимізований. При переході на іншу мову доведеться повністю перезавантажувати всю HTML сторінку, а отже і всі ресурси. До того ж це неймовірно важко для розробників – вирішивши змінити верстку доведеться змінювати її в кожному HTML файлі.

Наступний спосіб – використання файлів локалізації. Файлом локалізації може виступати файл у форматах JSON, XML, YAML, CSV та інші [4]. Головна вимога до такого файлу наявність пари ключ-значення, де ключ буде ідентифікатором стрічки і буде однаковим для кожної мови, а значення буде перекладом стрічки на певну мову.

Наступним кроком файл локалізації під’єднується до HTML сторінки. Це можна зробити декількома способами (за допомогою PHP, JavaScript та/або Ajax). Використання файлу локалізації суттєво спрощує процес розробки, підтримки та використання продукту: при необхідності змінити структуру документу потрібно зробити мінімум змін, при зміні мови перезавантажується лише файл локалізації (розмір якого в десятки менший ніж розмір цілого HTML файлу).

IV. АРХІТЕКТУРА СИСТЕМИ

В основу розроблюваної системи покладена клієнт-серверна архітектура, яка має ряд переваг відносно відомих програмних засобів локалізації (зниження об’єму даних, що передаються, ізоляція доступу користувача до даних та логіки обробки запитів, можливість додавання клієнтів іншого типу). У свою чергу клієнт та сервер також не монолітні. Для підсилення абстракції, зменшення кількості коду та збільшення продуктивності вони побудовані на основі патернів проектування. Крім того, вони поділені на шари, що підсилює інкапсуляцію, робить систему інтуїтивно зрозумілою та спрощує процес її тестування. Далі наведено запропоновану авторами структуру шарів клієнта і сервера.

Шари клієнта:

- шар представлення – відповідає за зовнішній вигляд додатка;
- проміжний шар – формує запити для сервера, та обробляє відповіді.

Шари сервера:

- проміжний шар – являє собою “ворота”, для доступу до сервера та перевіряє права користувача;
- шар логіки – містить у собі всю обробку даних та взаємодіє із віддаленим сервером для машинного перекладу;

- шар роботи з даними – ізолює стандартні операції роботи з даними (створення, редагування, читання та видалення).

Окремим модулем архітектури представлено сховище, що містить реляційну базу даних для збереження всіх даних у відношеннях, нереляційну документо-орієнтовану базу даних для збереження текстів та перекладів, а також сховище бінарних файлів та пошуковий сервер.

Окрім того особливістю запропонованої системи є використання сторонніх хмарних сервісів, а саме Firebase для аутентифікації користувача в системі та Google Translate API для здійснення машинного перекладу.

V. ПРОГРАМНА РЕАЛІЗАЦІЯ

Для реалізації серверної частини був обраний .NET Core Framework та мова C#. Особливістю даного фреймворку є кросплатформеність та оптимізація порівняно із .NET Framework. У якості ORM використовується Entity Framework Core, який спрощує процес взаємодії із реляційною базою даних (створення таблиць із класів моделей, підтримка міграцій) та MongoDB Driver для NoSQL бази даних. Обмін даними реалізований за допомогою надсилання HTTP запитами серіалізованих даних у форматі JSON, що набагато більш оптимізований ніж його класичний аналог XML.

Для реалізації клієнта використовувався Angular 6. Клієнт запрограмований на TypeScript, який є наступною сходинкою розвитку Javascript і на відміну від свого “батька” має строгу типізацію, краще реалізує концепції ООП та дозволяє підтримувати та масштабувати код значно легше. Для інтерфейсу користувача використовується HTML + SASS.

Для збереження даних використовуються MS SQL (всі реляційні дані), MongoDB (документи з текстом та перекладами), Azure Blob Storage (картинки проєктів та вкладені до тексту файли). Також для швидкого пошуку по документах використовується пошуковий сервер

ElasticSearch, для розгортання якого був використаний Docker.

Завдяки бібліотеці SignalR досягнуто обмін даних між сервером і клієнтом у реальному часі (без оновлення сторінки).

Машинний переклад реалізований за допомогою хмарного сервісу – Google Translate API, дані на який пересилають за допомогою HTTP запитів.

Для користувача було розроблено спеціальну можливість, сенс якої полягає в спостереженні за прогресом виконання перекладу. Для цього було проаналізовано всі можливі варіанти перекладу стрічки на задані мови, а саме:

- виконано всі переклади для однієї стрічки;
- виконано кілька перекладів для однієї стрічки;
- не виконано ні одного перекладу для однієї стрічки.

Використовуючи ці дані було виведено формулу 3, у якій l – це кількість мов у проєкті, n – це кількість неперекладених мов для стрічки, s – це кількість стрічок у проєкті, i – це певна стрічка..

$$P = \frac{\sum(l - n_i)}{s \cdot l} \cdot 100\% \quad (3)$$

ЛІТЕРАТУРА REFERENCES

- [1] Google Play improves all around, still trails iOS in revenue [Електронний ресурс]. – Режим доступу: <https://www.androidauthority.com/q2-2014-app-stats-show-google-play-ios-improve-407106/>.
- [2] Language Localization: Pushing Mobile Ad Conversion Across the Globe [Електронний ресурс]. – Режим доступу: <http://www.adotas.com/2014/04/language-localization-pushing-mobile-ad-conversion-across-the-globe/>.
- [3] Survey of 3,000 Online Shoppers Across 10 Countries Finds that 60% Rarely or Never Buy from English-only Websites [Електронний ресурс]. – Режим доступу: <http://www.commonseadvisory.com/Default.aspx?Aid=21500&Contenttype=ArticleDet&tabID=64>.
- [4] Общие ресурсы локализации [Електронний ресурс]. - Режим доступу: <https://metanit.com/sharp/aspnet5/28.8.php>.

Моделі Надійності Алгоритмів, що Виконуються з Помилками Різних Типів

Олексій Козачко
Кафедра системного аналізу,
комп'ютерного моніторингу та
інженерної графіки
Вінницький національний технічний
університет
Вінниця, Україна
lekoz80@gmail.com

Дмитро Солдатов
Кафедра системного аналізу,
комп'ютерного моніторингу та
інженерної графіки
Вінницький національний технічний
університет
Вінниця, Україна
dmitrsold@gmail.com

Артур Адаменко
Кафедра системного аналізу,
комп'ютерного моніторингу та
інженерної графіки
Вінницький національний технічний
університет
Вінниця, Україна
artur.adamenko@gmail.com

The Reliability Models of Algorithms, Performing with Errors of Diverse Types

Oleksiy Kozachko
Department of system analysis,
computer monitoring and engineering
graphics
Vinnitsa National Technical University
Vinnitsa, Ukraine
lekoz80@gmail.com

Dmitriy Soldatov
Department of system analysis,
computer monitoring and engineering
graphics
Vinnitsa National Technical University
Vinnitsa, Ukraine
dmitrsold@gmail.com

Artur Adamenko
Department of system analysis,
computer monitoring and engineering
graphics
Vinnitsa National Technical University
Vinnitsa, Ukraine
artur.adamenko@gmail.com

Анотація—Робота присвячена розв'язанню задачі моделювання надійності багатовимірних алгоритмів, при виконанні яких вносяться, виявляються та усуваються одночасно помилки різних типів. Наведені приклади типових представників багатовимірних алгоритмічних процесів, проаналізовано методи оцінювання та забезпечення їх надійності, поставлені задачі дослідження.

Abstract— The article is devoted to the solution of reliability modeling task of multidimensional algorithm i.e. processes, when being performed, errors of different types are noted, detected and deleted. The analyze of multidimensional algorithmic process patterns was carried out, estimation methods and providing of their reliability were analyzed, research tasks were set.

Ключові слова—алгоритми, алгоритмічні процеси, надійність алгоритмів, помилка різних типів.

Keywords—algorithms, reliability, errors of different types

I. ВСТУП

В роботі аналізуються алгоритмічні процеси (АП), при виконанні яких можуть вноситися, виявлятися та усуватися помилки різних типів, аналізуються методи оцінювання та забезпечення їх надійності, пропонуються принципи моделювання надійності таких АП та ставляться задачі дослідження [1].

Аналіз типових представників багатовимірних АП показав, що при моделюванні надійності використовують такі показники: ймовірність безпомилкового виконання АП; ймовірність наявності помилок різних типів на виході АП; середній час або вартість виконання АП. Основні труднощі оцінювання та забезпечення надійності АП обумовлені

необхідністю врахування помилок різних типів. Обмеженням відомих методів забезпечення надійності АП є те, що вони враховують помилки функціонування за бінарним принципом: помилки є, або немає. Вони не враховують, що для різних типів помилок різняться ймовірності їх внесення, виявлення, виправлення та витрати на ці дії. Аналіз теорії надійності алгоритмів свідчить, що необхідно розробити новий метод для синтезу багатовимірних АП, оптимальних за критеріями надійності, та для прогнозування надійності таких АП при нечітких початкових даних. Розв'язки цих задач забезпечить підвищення якості проектування АП.

II. ПОСТАНОВКА ЗАДАЧІ

Оптимізація АП полягає у визначенні структури процесу, що забезпечує необхідні рівні ймовірнісно-часових показників надійності функціонування. На практиці найбільш поширеними задачами оптимізації надійності АП є розподіл контрольних точок та вибір кратностей контролів. В цих задачах АП представляє собою послідовне виконання n робочих операторів A_1, A_2, \dots, A_n . При виконанні робочих операторів в АП можуть бути внесені помилки різних типів. Якість виконання робочого оператора A_i перевіряється контролем ω_i x_i -раз. Для задачі розподілу контрольних точок $x_i \in \{0, 1\}$, а для задачі вибору кратностей контролів $x_i \in \{0, 1, 2, \dots\}$, $i = \overline{1, n}$. Операція доробки U_i виправляє помилки, які виявлені контролем ω_i

знаходженні такого вектора $\mathbf{X} = (x_1, x_2, \dots, x_n)$, що забезпечує [2]:

пряма постановка:

$$\left. \begin{aligned} C(\mathbf{X}) \rightarrow \min, \\ p^1(\mathbf{X}) \geq P^* \text{ та } p^{0_j}(\mathbf{X}) \geq q_j, j = \overline{1, m}, \end{aligned} \right\} \quad (1)$$

зворотна постановка:

$$\left. \begin{aligned} p^1(\mathbf{X}) \rightarrow \max, \\ C(\mathbf{X}) \leq C^* \text{ та } p^{0_j}(\mathbf{X}) \leq q_j, j = \overline{1, m}, \end{aligned} \right\} \quad (2)$$

де $C(\mathbf{X})$ – вартість виконання АП, який задано вектором \mathbf{X} ; $p^1(\mathbf{X})$ – ймовірність правильного виконання АП; $p^{0_j}(\mathbf{X})$ – ймовірність помилки j -го типу на виході АП; P^* – мінімально допустима ймовірність безпомилкового виконання АП; $q_j(\mathbf{X})$ – максимально допустима ймовірність наявності помилки j -го типу на виході АП; C^* – максимально допустимі витрати на виконання АП; m – кількість різних типів помилок.

III. МОДЕЛІ НАДІЙНОСТІ

Пропонуються такі моделі цільової функції та обмеження задач оптимізацій [1-3]:

$$\mathbf{P}(\mathbf{X}) = \prod_{i=1, n} \mathbf{P}_{A_i} \cdot (\mathbf{K}_{\omega_i}^1 + \mathbf{K}_{\omega_i}^0 \cdot \mathbf{P}_{U_i})^{x_i}, \quad (3)$$

$$C(\mathbf{X}) = \sum_{i=1, n} \left[c_{A_i} + x_i \cdot c_{\omega_i} + c_{U_i} \cdot \sum_{j=1, m} p_{A_i}^{0_j} \cdot \left(1 - (k_{\omega_i}^{01_j})^{x_i} \right) \right], \quad (4)$$

де $\mathbf{P}_{A_i}(x_i) = \begin{pmatrix} p_{A_i}^1 & p_{A_i}^{0_1} & \dots & p_{A_i}^{0_m} \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ і

$$\mathbf{P}_{U_i} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ v_{U_i}^1 & v_{U_i}^{0_1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ v_{U_i}^m & 0 & \dots & v_{U_i}^{0_m} \end{pmatrix} \text{ – характеристики надійності}$$

операторів A_i і U_i :

$p_{A_i}^{0_j}$ – ймовірність внесення оператором A_i помилки j -го типу;

$p_{A_i}^1$ – ймовірність безпомилкового виконання оператора A_i ; $v_{U_i}^1(v_{U_i}^{0_j})$ – ймовірність усунення (не усунення) помилки j -го типу, $j = \overline{1, m}$;

$$\mathbf{K}_{\omega_i}^1 = \begin{pmatrix} k_{\omega_i}^{11} & 0 & \dots & 0 \\ 0 & k_{\omega_i}^{01_1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & k_{\omega_i}^{01_m} \end{pmatrix} \text{ і } \mathbf{K}_{\omega_i}^0 = \begin{pmatrix} k_{\omega_i}^{10} & 0 & \dots & 0 \\ 0 & k_{\omega_i}^{00_1} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & k_{\omega_i}^{00_m} \end{pmatrix} \text{ –}$$

характеристики надійності контролю ω_i : $k_{\omega_i}^{11}(k_{\omega_i}^{10})$ – ймовірність того, що відсутність помилок ідентифіковано правильно (неправильно); $k_{\omega_i}^{01_j}(k_{\omega_i}^{00_j})$ – ймовірність пропуску (виявлення) помилки j -го типу, $j = \overline{1, m}$; c_{A_i} , c_{ω_i} та c_{U_i} – вартість виконання робочого оператора A_i , контролю ω_i і доробки U_i , відповідно, $i = \overline{1, n}$.

Формули (3)-(4) отримані x_i – кратним застосуванням моделі надійності структури “робота-контроль-доробка”

ВИСНОВКИ

В роботі наведено моделі надійності багатовимірних АП, які на відміну від існуючих, враховують помилки функціонування різних типів, що дозволяють розв’язувати задачі забезпечення надійності, наприклад, задача оптимізації розподілу контрольних точок та вибору кратностей контролів в таких процесах. Практична цінність результату полягає в тому, що на його основі може здійснюватися оперативне управління надійністю багатовимірного АП.

ЛІТЕРАТУРА REFERENCES

- [1] Штовба С.Д., Козачко О.М. Генетична мінімізація вартості контролів в технологічному процесі з урахуванням дефектів різних типів // Вісник Вінницького політехнічного інституту. – 2005. – №3. – С.74-79
- [2] Штовба С.Д., Козачко О.М. Генетична оптимізація кратностей контрольно-доробчих операцій в технологічних процесах з урахуванням дефектів багатьох типів // Вісник Житомирського державного технологічного університету. – 2004. – №4. – Том 2. – С.180-187.
- [3] Shtovba S., Kozachko O., Dounias G. A fast genetic algorithm for optimising the checking – retrofit procedures in multidimensional technological processes // Штучний інтелект. – 2004. – №2. – С.225-230.

Integro-Differential Model of Pricing on the Electricity Market

Tatyana Zaika
Faculty of Computer Science and Software Engineering
National Technical University "Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
zaika.tats@gmail.com

Інтегро-Диференційна Модель Ціноутворення на Енергоринку

Тетяна Заїка
Факультет комп'ютерних наук і програмної інженерії
Національного Технічного Університету «Харківський Політехнічний Інститут»,
Харків, Україна
zaika.tats@gmail.com

Abstract—A model for determining the equilibrium price of electricity markets is proposed. The model includes hedging risks using a portfolio of derivatives. The control of model parameters makes it possible to restrict price changes by bifurcations of the type of cycle birth.

Keywords: *Integro-differential equation system, pricing model, electricity market, derivative financial instruments.*

Анотація — У роботі запропонована модель визначення рівноважної ціни на енергоринках. У модель включено хеджування ризиків з використанням портфеля деривативів. Управління параметрами моделі дозволяє обмежити зміну цін біфуркації типу народження циклу.

Keywords: *Інтегро-диференціальна система рівнянь, моделі ціноутворення, електроенергетичні ринки, похідні фінансові інструменти.*

I. INTRODUCTION

The modern electricity market assumes that the supplier is fighting for its buyer. Companies generating electricity corresponding to the price range for their products. The buyer freely chooses the generator company and enters into a supply contract with him in the appropriate Internet platform. The power industry market is constantly in dynamics.

The general model of competitive electricity markets can have the following structure: the wholesale electricity market, retail electricity markets, the capacity market, the system services market, the market for financial transmission rights, the derivatives market. Prices in all these markets are interrelated [1].

Electricity production should respond quickly to real-time changes in electricity consumption. Electricity can't be efficiently stored in large volumes. The process of electricity consumption is nonlinear. Energy efficiency of electric energy supply in modern electric networks should take into account asymmetric, nonlinear and fast-moving nature of the consumption of electric energy by modern consumers. The main factors are human physiological cycles, changing the time of day and seasonality, shift work. The increasing complexity of the topology of electrical

networks, the increase in the number of renewable energy sources and the peculiarities of the electricity market leads to the search for new methods and approaches for modeling and forecasting the parameters of the power system.

Nowadays, the methods of calculating intelligence, methods of game theory, fractal models, dynamical systems, multiagents, etc. are used to rebuild the energy consumption model [2-5].

Methods of computing intelligence are based on heuristic algorithms that are used, for example, in fuzzy logic, artificial neural networks and evolutionary simulation.

Creation of fractal models with a set of electric power and heat networks capable of describing real consumers with uneven loads and power points is a promising direction. On the basis of fractal theory, methods of short-term load forecasting are proposed.

To protect the integrity of the grid, it is important to take into account the physical limitations and limitations for network security. According to consumer demands for electricity demand, the forecast of annual and monthly balance of electric energy is formed [7]. Therefore, the creation of an overload management scheme that addresses the problem of limiting flows in power grids is one of the most difficult tasks in the energy market structure.

Most existing energy market models either stationary or dynamic linear feedbacks. They all apply the hypothesis of the global asymptotic constancy of a closed energy system. But in modern conditions, the use of this hypothesis does not always allow the construction of sufficiently adequate models of energy consumption.

Creating a model, it is necessary to compare the form and complexity of its mathematical description with the capabilities and specifics of the method that will be used to analyze the dynamics of the new model.

II. THE MAIN RESULT

Taking into account the above, the purpose of the work was to develop a non-linear model of the energy market, which:

1. The model assumes that the cost of electricity includes the cost of services and financial rights;
2. The model assumes that the rate of price change depends on weighted average prices in the past;
3. The model assumes that the weighted average price in the capacity market is described by a well-known function;
4. It is possible to hedge the risks of abrupt price changes in the portfolio of futures and options;
5. There are control parameters that are determined empirically.

Then, following the ideas developed in [6], we obtain the following model:

$$\begin{aligned} \dot{q}(t) &= \lambda q(t) \left[1 - \int_{-\infty}^0 q(t+\tau) w(-\tau) d\tau \right] - \mu p(t) q(t) \\ \dot{p}(t) &= -\alpha p(t) + \beta p(t) q(t) - f(p(t), t) \end{aligned} \quad (1)$$

where $\lambda, \mu, \alpha, \beta$ are positive parameters, $\alpha > \beta$;

μ is the bifurcation parameter;

$q(t)$ is the weighted average supply price of the wholesale electricity market;

$p(t)$ is the weighted average supply price of the retail electricity market;

$w(\tau)$ is the weighted average supply price of the wholesale capacity market;

$f(p(t), t)$ is the price of the option premium.

The integral term in equation (1) is associated with the need to take into account the dependence of the instantaneous values of the characteristics on their respective previous values, i.e., the impact on the current state of the system of its prehistory is taken into account. In congestion management, price is a deterrent to energy consumption for consumers.

Connections in power systems are convenient to represent in the form of graphs. Suppose that in modern electrosystems, they consist of microgrid, which subsequently unite into a common system. Restrictions are imposed on the power generated, on the power line, on the number of participants in the energy market. The construction of a dynamic graph taking into account the constraints and the balance of generation and consumption of electricity is a promising and difficult task. In the

dynamic graph, changes are constantly occurring, such as the ribs appear and disappear (we will denote generators / consumers), their weights change (we will denote the amount of electricity that is transmitted). Changes need to be processed quickly so that power supplies and their prices are effective.

CONCLUSIONS

Model (1) has the only equilibrium price value. For the kernel $w(t)$ of a general form, the system is not finite-dimensional. However, in special cases, the model can be reduced to a system of four ordinary equations. In these cases, the cycles are asymptotically orbitally stable, being born with the bifurcation value λ_0 of the parameter λ . Classical methods for evaluating financial derivatives do not work well on the energy markets. Therefore, the actual task is to build models for evaluating "energy" futures, options and hedge portfolios.

Integro-differential models have advantages over differential models with numerical implementation. But problems of application and computer realization of integro-differential models have not been solved yet. Researches of such models are an important and relevant direction, and the development of methods and tools for their computer modeling is necessary for a wide range of application development problems.

REFERENCES ЛІТЕРАТУРА

- [1] Wolak, F.A., Patrick, R.H. "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market", POWER Working Paper PWP-047. University of California Energy Institute, April 1997.
- [2] Pierluigi Mancarella, Chin Kim Gan, Goran Strbac, "Fractal models for electro-thermal network studies", 17th Power Systems Computation Conference (PSCC), Stockholm, Sweden, 22-26 August 2011
- [3] Mordjaoui M., Boudjema B., Bouabaz M., Daira R. "Short-Term Electric Load Forecasting Using Neuro-Fuzzy Modeling for Nonlinear System Identification". Proceedings of International 3rd Conference on Nonlinear Science and Complexity (NSC'10), No. 64, Ankara, Turkey, 2010.
- [4] Romero I. "Dynamic Load Models for Power Systems - Estimation of Time-Varying Parameters During Normal Operation Department of Industrial Electrical Engineering and Automation", Lund Institute of Technology, Lund, 2002.
- [5] 4. Zhang Qing, Zhang Li, "Short Term Load Forecasting Using Chaotic Analysis Algorithm for Neural Network". Mechanical and Electrical Engineering Institute, Agricultural University of Hebei, P.R.China, 2007.
- [6] May R. Time-delay versus stability in population models with two and three trophic levels. – Ecology, 1973, 54, p. 315-325
- [7] <http://zakon.rada.gov.ua/laws/show/z1224-16>

Non-proportional Digital Image Scaling

Ekaterina Tryfonova

department of computer science and management information
systems protection
Odessa National Polytechnic University
Odessa, Ukraine
katikkatik@gmail.com

Oleksandr Ogoliuk

department of computer science and management information
systems protection
Odessa National Polytechnic University
Odessa, Ukraine
aleksandr.ogoliuk@gmail.com

Непропорційне масштабування цифрового зображення

Катерина Трифонова

кафедра інформатики та управління захистом
інформаційних систем
Одеський Національний Політехнічний Університет
Одеса, Україна
katikkatik@gmail.com

Олександр Оголюк

кафедра інформатики та управління захистом
інформаційних систем
Одеський Національний Політехнічний Університет
Одеса, Україна
aleksandr.ogoliuk@gmail.com

Abstract—As a result of the spatial affine transformation implementation, proportional scaling of the digital image, the image size is resized with the proportional preservation of its content. To account for the content of a digital image when performing scaling, apply a non-proportionate or content-dependent scaling of the digital image. In the work presented, the implementation of a disproportionate scaling of a digital image is presented. The functions of determining the importance of the image pixels constituting the basis of the presented algorithm are studied.

Keywords: *digital image, proportional scaling, non-proportional scaling, gradient, filter.*

I. INTRODUCTION

With the massive distribution various means of reproduction and information transmission for various formats (smart phones, mobile phones, tablets, laptops), scaling of visual content became more widely used. In this case, in addition to implementing the adaptive layout of the corresponding interactive visual content, digital images that cannot be displayed in full scale must be generated in the appropriate format.

In addition, an image magnification for a more detailed examination of the entire image or its fragment, as the task is used in various interactive visual content. At the same time, a very important task of any modern application should be fulfilled - saving of memory: there is a digital image of the size specified by the user at the moment not only opens from some archive, but should be generated automatically.

Implementation of effective and fast interactive interaction of the user with visual content constantly requires qualitatively new methods of processing digital images. That is why the development of advanced scaling methods is very important, which makes the topic of this work extremely relevant.

II. STATEMENT OF THE MAIN MATERIAL

Different algorithms are used for scaling [1-3]. The simplest and fastest, but simultaneously with the low quality

of the resulting image, is the algorithm for scaling the closest neighbor. This algorithm deletes rows and columns when the image is diminished and duplicates them when enlarged. The algorithm of bilinear interpolation gives the best result due to the fact that when the image is enlarged, the voids formed are not filled with duplicate adjacent pixels, but by their averaged value. Similar to the algorithm bicubic interpolation, but in it to create a new pixel not only neighbors, but also neighbors of the neighbors are used. And finally, the most complex algorithm is fractal. Before scaling, the picture is vectorized, and in the process of choosing the color for new points, vectors whose edges make the contrast more involved.

But all these methods have application in scaling the image only if its dimensions vary proportionally both horizontally and vertically. In the case of non-proportionate scaling, using the above-mentioned algorithms, the context of the digital image loses meaning. Therefore, scaling algorithms, based on preliminary analysis of the image, are of particular importance.

As a mathematical model of the image, regardless of the registration method, as a result of sampling and quantization of the continuous signal, the digital image is considered as a set of pixels.

As a result of the sampling procedure, the digital image is considered as a set of points located in a rectangular coordinate system. Adhere to the adopted coordinate system with the beginning in the upper left corner of the image and with positive directions from this point downwards and to the right [1,2].

As a result of the quantization procedure, depending on the selected color model, the set of pixels forms a matrix or several matrices of the brightness of the digital image. In the color model of the grayscale $F = (f_{y,x}) - R \times C$ the brightness matrix of the digital image with the elements $f_{y,x} \in [0, 2^n - 1]$, $y = \overline{1, R}$, $x = \overline{1, C}$, n - the number of bits is determined to hold the brightness value of one pixel. Figure 1 shows a gray-scale digital image matrix,

which, as an example, will be used for further analysis and a description of the algorithm for non-proportionate scaling.

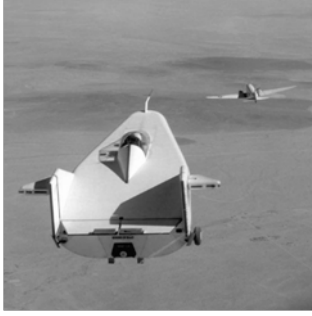


Fig.1. Gray-scale digital image matrix

The geometric transformation of a digital image consists of two main operations:

- the spatial transformation of a digital image, which results in a change in the location of the set of image points on a plane from one two-dimensional coordinate system to another;
- restoring a digital image, that is, interpolating the brightness values, which results in assigning the brightness values of the image pixels that were subjected to spatial transformation.

One of the most common classes of geometric transformations is affine.

The affine transformation is the most general one-to-one mapping of the plane onto the plane, as a result of which: the straight line goes straight; parallel straight lines become parallel; the ratio of the lengths of segments lying on one line or parallel straight lines is preserved; the ratio of the areas of the figures is preserved [3].

The affine transformation on the plane is described by the equations [3]:

$$\begin{cases} x' = Ax + Cy + E \\ y' = Bx + Dy + F \end{cases}, \quad (1)$$

where A, B, C, D, E, F are constants;

$(x, y), (x', y')$ – initial and coordinate after the conversion.

It is convenient to write the affine transformation in the matrix form. The constants A, B, C, D, E, F form the transformation matrix. To take into account the constants E, F goes to the homogeneous coordinates, due to the introduction of an additional line in the transformation matrix, and the points of the plane are described by three-dimensional vectors of the form $(x, y, 1)$. Then, in homogeneous coordinates, the affine transformation is described [3]:

$$(x', y', 1) = (x, y, 1) \cdot \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ E & F & 1 \end{pmatrix}. \quad (2)$$

The proportionate scaling of the digital image is determined as follows:

$$\begin{cases} x' = Ax \\ y' = Dy \end{cases}. \quad (3)$$

Proportionate scaling matrix in homogeneous coordinates:

$$T = \begin{pmatrix} A & 0 & 0 \\ 0 & D & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4)$$

The first geometric transformation operation, the spatial transformation of a digital image for non-proportionate scaling, consists of the following steps [4,5]:

- determining the digital image gradient;
- constructing a matrix of minimal sums of importance for pixels;
- determining the sequence of coordinates.

Let's consider each step in more detail.

Since the main purpose of the non-proportionate scaling is to resize the image by changing the size of the insignificant portions of the image without distorting the content of the digital image, then the first step of non-proportional scaling is the assignment to each pixel of the digital image of its importance. A gradient is used to determine the degree of importance of each pixel. The digital image gradient represents the direction and norm of the maximum brightness change rate in each pixel of the digital image.

$$G_g = (g_{g_{y,x}}) \text{ and}$$

$G_v = (g_{v_{y,x}})$, $y = \overline{1, R}$, $x = \overline{1, C}$, according to the spatial coordinates, are calculated as follows [1,2]:

$$G_g = F * S_g, g_{g_{y,x}} = \sum_{w=-n}^n \sum_{u=-n}^n f^{y+w, x+u} \cdot s_{g_{w,u}}, \quad (5)$$

$$G_v = F * S_v, g_{v_{y,x}} = \sum_{w=-n}^n \sum_{u=-n}^n f^{y+w, x+u} \cdot s_{v_{w,u}}, \quad (6)$$

where is F the matrix of the brightness of the digital image – S_g, S_v core of gradient filters;

– $[*]$ operator of a two-dimensional correlation.

Norm $G = (g_{y,x})$ and angle $A = (\alpha_{y,x})$, $y = \overline{1, R}$, $x = \overline{1, C}$, image gradient [1,2]:

$$g_{y,x} = \sqrt{g_{g_{y,x}}^2 + g_{v_{y,x}}^2}, \quad (7)$$

$$\alpha_{y,x} = \arctg \left(\frac{g_{g_{y,x}}}{g_{v_{y,x}}} \right). \quad (8)$$

The easiest way to calculate the gradient of an image using a filter core [6]:

$$S_g = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}, S_v = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}. \quad (9)$$

The vector gradient field for the image and the gradient's norm is represented on the figure 2.

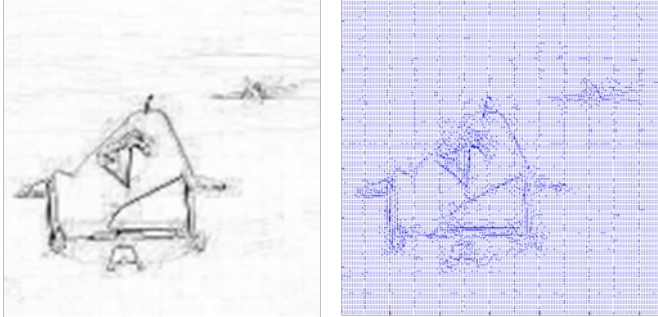


Fig.2. Digital image gradient research with a simplest filter core

Another common way to calculate the gradient of an image using core of Roberts filter [6]:

$$S_g = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, S_v = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}. \quad (10)$$

The vector gradient field for the image and the gradient's norm is represented on the figure 3.

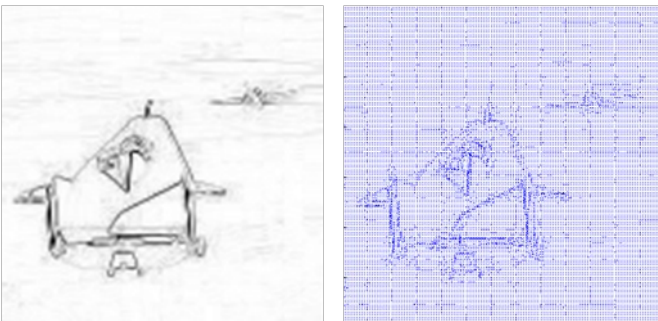


Fig.3. Digital image gradient research with a Roberts filter core

Another way to calculate the gradient of an image using core of Pruitt filter [6]:

$$S_g = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, S_v = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}. \quad (11)$$

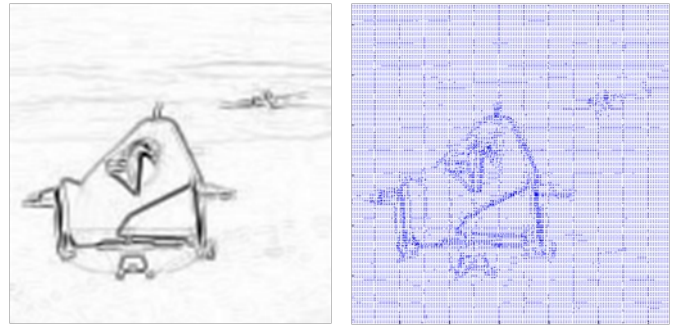


Fig.4. Digital image gradient research with a Pruitt filter core

Another common way to calculate the gradient of an image using core of Sobel filters [6]:

$$S_g = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, S_v = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}. \quad (12)$$

The vector gradient field for the image and the gradient's norm is represented on the figure 5.

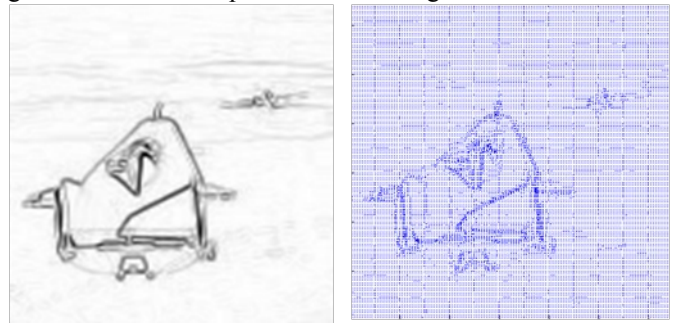


Fig.5. Digital image gradient research with a Sobel filter core

For a non-proportional scaling of a digital image, as a result of increasing or decreasing the digital image, it is necessary to determine the pixels with the least importance, so that in addition to taking into account the contents of the image, each row (column) is changed to the same number of pixels.

Therefore, it is suggested in [4,5] to construct chains of pixels - sequences of points, such that exactly one pixel is selected in each line, and adjacent pixels in it are connected by sides or corners. Therefore, the next step in the algorithm of non-proportionate scaling is to build a matrix of minimum sums of importance of pixels:

$$m_{y,x} = \begin{cases} g_{y,x} & , y = 1 \\ g_{y,x} + \min_{k=-1}^1 (m_{y-1,x+k}), & y \neq 1 \end{cases}. \quad (13)$$

In figure 6 for the digital image matrixes of minimum sums is represented for different gradients, which was calculated by a two-dimensional correlation with using cores of simplest, Roberts, Pruitt and Sobel filters

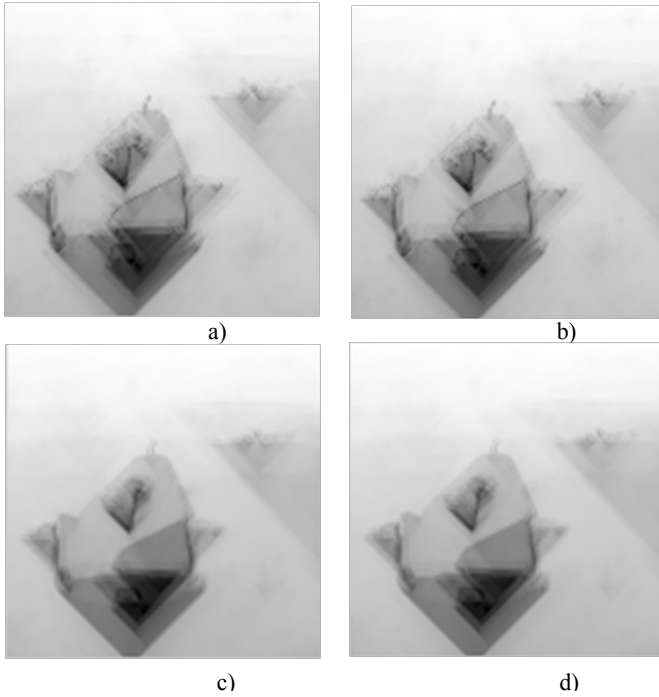


Fig.6. Matrixes of minimum sums research

The final step is to determine the sequence of coordinates for further enlarging or reducing the image:

$$l_y = \begin{cases} \text{num}(\min_{x=1}^c(m_{y,x})) & , y = R \\ \text{num}(\min_{k=1}^1(m_{y,l_{y+1}+k})) & , y = \overline{R-1,1} \end{cases} \quad (14)$$

The result of this step is shown in the figure 7.

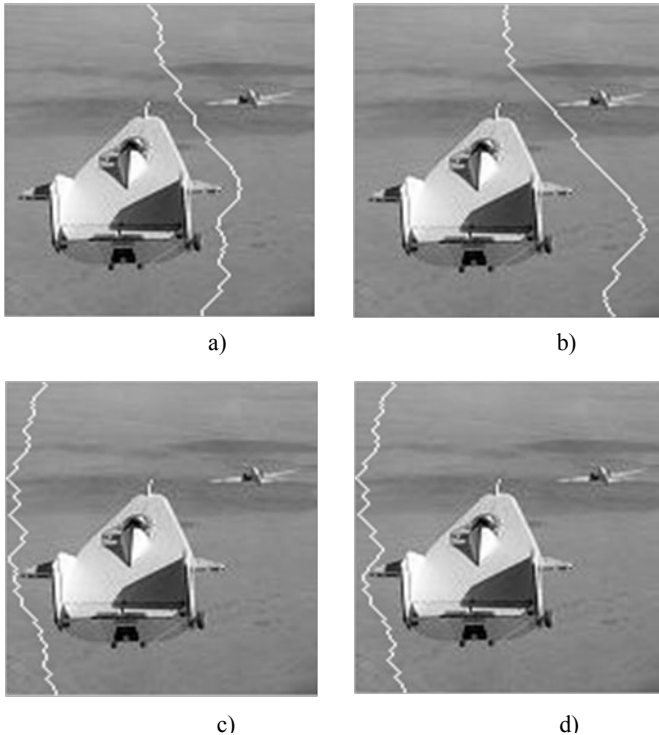


Fig.7. Research of coordinates sequence for further enlarging or reducing digital image

CONCLUSION

In this paper, the first step is taken to achieve the main goal [7] – to determine the violation of the integrity of the digital image as a result of scaling. For this purpose, it was implemented and investigated the various algorithms not only proportional, but also the main non-proportionate scaling.

ACKNOWLEDGMENT

The work was carried out within the framework of state budget scientific research topics of the Department of Informatics and Information Systems Protection of the Odessa National Polytechnic University "Development of Methods for Increasing the Efficiency of a Comprehensive Information Security System" (Governmental Registration No. 0115U000834).

REFERENCES ЛІТЕРАТУРА

- [1] Гонсалес, Р. Цифровая обработка изображений в среде Matlab, Р. Гонсалес, Р. Вудс, С. Эддинс, М.: «Техносфера», 2006, 616 с.
- [2] Гонсалес, Р. Цифровая обработка изображений, Р. Гонсалес, Р. Вудс, М.: «Техносфера», 2005, 1072 с.
- [3] Старовойтов, В.В. Цифровые изображения: от получения до обработки, В.В. Старовойтов, Ю.И. Голуб, Минск: ОИПИ НАН Беларуси, 2014, 202 с.
- [4] Shai, A. Seam Carving for Content-Aware Image Resizing, A. Shai, S. Ariel, ACM Transactions on Graphics, Vol.26, №10, 2007, pp. 82–91.
- [5] Делаем Liquid Resize своими руками: [Электронный ресурс], Хабрахабр. Режим доступа: <https://habrahabr.ru/post/48518/> (10.04.2018).
- [6] Шийка, Ю. Енергетичні функції в задачах масштабування з врахуванням вмісту растрових зображень, Ю. Шийка, Р. Шувар, Теоретична електроніка, Вип.60, 2009р., С. 139–146.
- [7] Кобозева, А.А. Повышение эффективности метода обнаружения фальсификации цифрового изображения, основанного на анализе сингулярных чисел матрицы / А.А. Кобозева, Е.А. Трифонова // Научный и производственно-практический сборник. Труды Одесского политехнического университета. – Вып.1(29). – 2008. – С. 183–190.

Оцінювання та Аналіз Навичок та Компетенцій Студентів ВНЗ

Кузьмін Олексій Олегович
Національний технічний
університет «Харківський
політехнічний інститут»
Харків, Україна
oleksii.kuzmyn@gmail.com

Орловський Дмитро Леонідович
Національний технічний
університет «Харківський
політехнічний інститут»
Харків, Україна
ordm@kpi.kharkov.ua

Копп Андрій Михайлович
Національний технічний
університет «Харківський
політехнічний інститут»
Харків, Україна
kopp93@gmail.com

Assessment and Analysis of Students Skills and Competences

Kuzmin Oleksii Olegovich
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
oleksii.kuzmyn@gmail.com

Orlovskiy Dmytro Leonidovych
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
ordm@kpi.kharkov.ua

Kopp Andrii Mykhailovych
National Technical University
"Kharkiv Polytechnic Institute"
Kharkiv, Ukraine
kopp93@gmail.com

Анотація—У даній роботі розглянута задача нечіткого логічного висновку для прийняття рішень в задачах підбору контингенту студентів виходячи з вимог потенційних роботодавців, що дозволяє зменшити час на обробку інформації про кандидатів на конкретну посаду.

Ключові слова: вибір альтернатив, нечітка логіка, автоматизована інформаційна система, нечіткий висновок, нечіткі судження, агрегування альтернатив.

Abstract—This paper considers the problem of the fuzzy logic inference in order to support the decision making procedure of students' selection based on the requirements of the potential employers. This approach allows reducing the time spent on the processing of the information about candidates for a certain position. The fuzzy characteristics of the input data set are implemented, the set of fuzzy logic inference rules is defined, the algorithm of fuzzy logic inference, which allows selecting students that better suite for a certain competence, is outlined. For the example we have considered the problem of selection of the best candidate for the position "Junior Front-End Developer" among the set of four sample candidates. By using the set of required skills in the field of front-end web development, we have applied the proposed approach to estimate the scores of such skills for each candidate in order to pick the candidates with the best score.

Keywords: choice of alternatives, fuzzy logic, automated information system, fuzzy conclusion, fuzzy judgments, aggregation of alternatives.

I. ВВЕДЕННЯ

Системи, засновані на нечіткій логіці, розроблені і успішно впроваджені в таких областях, як управління технологічними процесами, транспортом, побутовою технікою, медична та технічна діагностика, фінансовий менеджмент, фінансовий аналіз, біржове прогнозування, розпізнавання образів, дослідження ризикових і критичних операцій, прогнозування землетрусів, складання автобусних розкладів, кліматичний контроль в будівлях.

Більшість сучасних задач управління просто не можуть бути вирішені класичними методами через дуже велику складність математичних моделей, що їх описують.

В умовах сьогодення, як ніколи актуальним постає розгляд проблематики працевлаштування студентів ВНЗ. Також постає питання щодо підготовки студентів, яка повинна максимально відповідати тим вимогам, які висувають роботодавці.

Актуальним є розробка механізму прийняття рішень для об'єктивної оцінки знань та навичок студентів. Таке завдання є слабоструктурованим. В умовах неповноти і неточності інформації побудова точної математичної моделі виявляється проблематичною. З іншого боку, створення моделі складних об'єктів, що погано формалізуються, стає важко здійсненним. У таких випадках найбільш ефективними є нечіткі методи моделювання, які в значній мірі засновані на знаннях експертів, на підставі яких, в свою чергу можуть бути отримані позитивні результати в ітеративному процесі уточнення несуперечливої моделі. У зв'язку з цим доцільно використовувати метод нечіткого логічного висновку для вирішення завдання оцінювання та аналізу навичок та компетенцій студентів ВНЗ.

II. АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ

За останні роки збільшилася кількість публікацій з теоретичних та прикладних аспектів систем оброблення нечіткої інформації. Розроблені програмні та апаратно-програмні комплекси для підтримки прийняття рішень на базі нечіткої логіки та знань з різними рівнями інтелектуальності [1, 2].

Предметною областю останніх досліджень є додатки у побутовій техніці, медицині та інших сферах людської життєдіяльності [3-5]. Необхідно зазначити, що останні прикладні розробки об'єднує напрямком на реалізацію однієї з можливих модифікацій алгоритмів нечіткого логічного висновку, а саме – композиційного висновку.

III. МЕТА ДОСЛІДЖЕННЯ

Метою даної роботи є обґрунтування доцільності практичного використання нечіткого логічного висновку з метою прийняття рішень в задачах вибору найбільш підготовлених студентів для працевлаштування в ІТ-компанії. Це також допоможе знизити навантаження на HR-менеджерів та зменшити час на проведення співбесід.

IV. ВИКОРИСТАННЯ НЕЧІТКОЇ ЛОГІКИ ДЛЯ ОЦІНКИ КОМПЕТЕНЦІЙ

У сучасній практиці нечітка логіка застосовується у багатьох різних методах оцінки, у тому числі анкетування, інтерв'ю, тестування, описовий метод, метод класифікації, парне порівняння, рейтинговий метод, ділові ігри моделі компетентності тощо [5, 6]. Кожен з методів має свої переваги та недоліки, але ефективні вони тільки в складі єдиної системи управління персоналом. Необхідно, щоб методологія оцінки, яка забезпечує базу для реалізації функцій управління персоналом:

- була основана на об'єктивній інформації та давала об'єктивні оцінки кандидатів;
- стимулювала тих, хто здійснює відбір, до обґрунтування рішення з урахуванням узгодженої системи критеріїв;
- забезпечила взаєморозуміння сторін при оцінці персоналу [7];
- забезпечила реалізацію обґрунтованих управлінських рішень.

При аналізі результатів оцінки персоналу виникає певна ступінь нечіткості. Як метод для реалізації системного підходу до оцінки персоналу запропоновано використовувати нечітку логіку, математичний апарат, який дозволяє побудувати модель об'єкту, засновану на нечітких розсудах і правилах. Найголовніша умова створення такої моделі полягає в тому, щоб перевести нечіткі, якісні оцінки, що застосовуються людиною, до мови математики, зрозумілої комп'ютеру [8, 9]. Дослідження в області застосування нечіткої логіки в соціоекономічних системах [10] дозволяють говорити про можливість її використання для оцінки компетенцій студентів ВНЗ.

V. МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ ВИРІШЕННЯ ЗАДАЧ

Нехай U – безліч елементів, A – його нечітка підмножина, ступінь приналежності елементів якого є число з одиничного інтервалу $[0,1]$. Підмножин A є значеннями лінгвістичної змінної X .

Нехай безліч рішень характеризується набором критеріїв X_1, X_2, \dots, X_p , тобто лінгвістичних змінних на базових множинах U_1, U_2, \dots, U_p відповідно. Набір з декількох критеріїв з відповідними значеннями характеризують уявлення про задовільність рішення. Мінлива S «задовільно» також є лінгвістичною. Приклад висловлювання:

d_1 : «Якщо X_1 = низька та X_1 = гарне, то S = висока».

У загальному випадку висловлювання має вигляд:

d_i : «Якщо $X_1 = A_{1,i}$ та $X_2 = A_{1,i}$ та ... $X_p = A_{1,p}$, тоді $S = B_i$ ».

Позначимо перетин $X_i = A_{1,i} \cap X_2 = A_{2,i} \cap \dots \cap X_p = A_{p,i}$ через $X = A_i$. Операції перетину нечітких множин відповідає знаходження мінімуму їх функцій приналежності

$$\mu_{A_i}(\vartheta) = \min_{\vartheta \in V} \{ \mu_{A_{i,1}}(u_1), \mu_{A_{i,1}}(u_2), \dots, \mu_{A_{i,p}}(u_p) \},$$

де $V = U_1 \times U_2 \times \dots \times U_{p,i} = (u_1, u_2, \dots, u_p)$, $\mu_A(u_j)$ – значення приналежності елемента u_j нечіткій множині $A_{i,j}$. Тоді висловлювання набуде вигляду:

d_i : «Якщо $X = A_i$, то $S = B_i$ ».

Позначимо базову множину через W . Тоді A^i – нечітка підмножина W , в той час як B^i – нечітка підмножина одиничного інтервалу I .

Імплікація нечітких множин виражається в такий спосіб:

$$\mu_H(w, i) = \min_{w \in W} \{ 1, (1 - \mu_A(w) + \mu_B(i)) \},$$

де H – нечітке підмножина.

Аналогічним чином висловлювання d_1, d_2, \dots, d_q перетворюються у множини H_1, H_2, \dots, H_q .

Їх об'єднанням є множина D :

$$D = H_1 \cup H_2 \cup \dots \cup H_q$$

та для кожного $(w, i) \in W \times I$:

$$\mu_D(w, i) = \min_{w \in W} \{ \mu_H(w, i) \}, i = 1, 2, \dots, q.$$

Розглянемо вибір альтернатив, кожна з яких описується нечіткою підмножиною C з підмножини W . Задовільність альтернативи визначається на основі композиційного правила виведення. Тоді:

$$\mu_C(i) = \max_{w \in W} \{ \min \{ \mu_C(w), \mu_D(w, i) \} \}.$$

Зіставлення альтернатив відбувається на основі точкових оцінок.

$$A_\alpha = \{ x | \mu_A(x) \geq \alpha, x \in I \}.$$

Для кожного A_α можна обчислити середнє число елементів $M(A_\alpha)$:

$$M(A_\alpha) = \frac{\sum_{i=1}^n \frac{a_i + b_i}{2} (b_i + a_i)}{\sum_{i=1}^n (b_i + a_i)}.$$

Тоді точкове значення для множини A :

$$M(A) = \frac{1}{\alpha_{max}} \int_0^{\alpha_{max}} M(A_\alpha) d\alpha,$$

де α_{max} – значення, при якому A має максимум.

При виборі альтернатив, для кожної з них визначається задовільність і обчислюється відповідна точкова оцінка. Кращою вважається альтернатива з найбільшим її значенням [11].

VI. ЗАСТОСУВАННЯ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ДЛЯ ПІДТРИМКИ ВИРІШЕННЯ ЗАДАЧІ ВИЯВЛЕННЯ НАЙКРАЩОГО ПРЕТЕНДЕНТА

Для вирішення задачі виявлення найкращого претендента на певну вакансію пропонується використання клієнт-серверної технології, яка базується на використанні серверу баз даних MySQL, мов

програмування PHP, JS, CSS, HTML. У фрагменті моделі даних представлені наступні сутності.

Структура бази даних ілюструється фрагментом бази даних, яка наведена на рисунку 1. Наведений фрагмент показує основні сутності (таблиці) за допомогою яких зберігаються данні про студентів, їх навички, компетенції, рівні компетенцій.

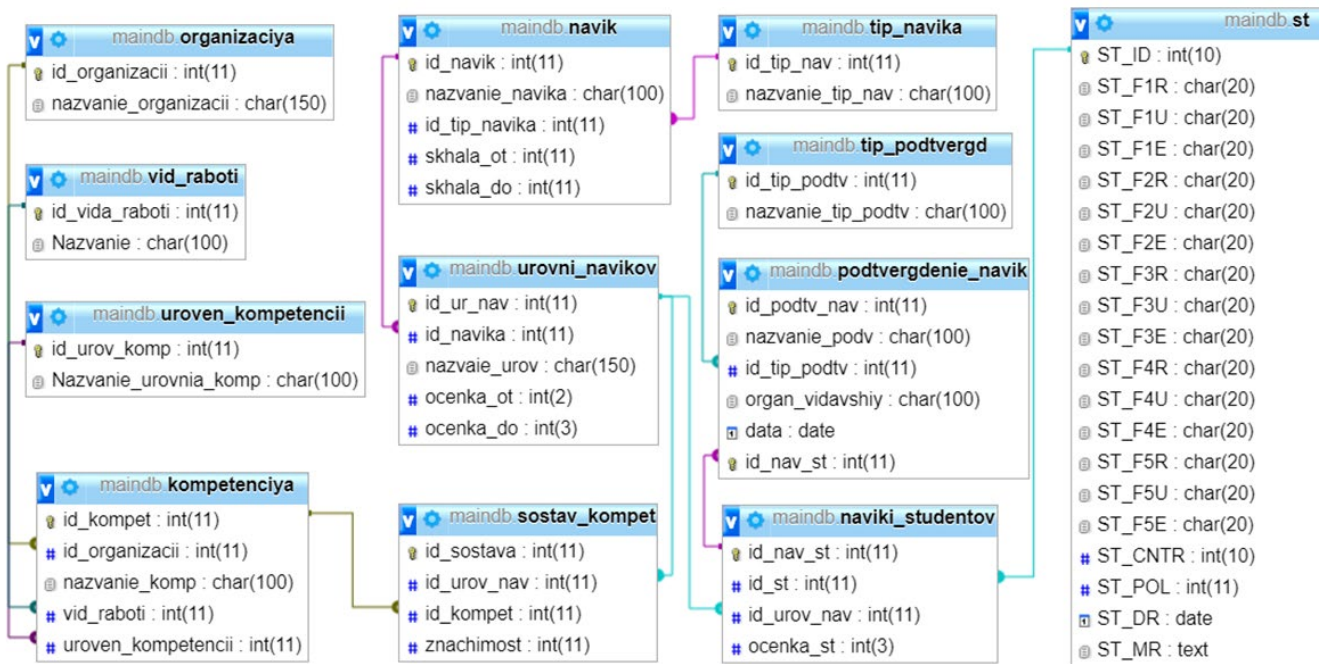


Рис. 1. Фрагмент моделі даних.

Послідовність використання програмного забезпечення може бути наступна. Студенти, які розглядаються як потенційні кандидати, складають тести або заповнюють анкети на підставі яких формується базова інформація щодо їх компетенцій, знань та навичок. Ця інформація потрапляє до бази даних. Потім за допомогою програмного забезпечення реалізується первинна оцінка контингенту студентів, на підставі якої формується список кандидатів, який ранжується за значенням точкової оцінки. Саме ці дані є первинними для відбору найбільш перспективних кандидатів, з якими потім проводиться додаткова співбесіда.

VII. РОЗРАХУНОК КОНТРОЛЬНОГО ПРИКЛАДУ

Завдання полягає в тому, щоб використовуючи метод нечіткого логічного висновку, виявити найкращого претендента. Приклад підбору кандидатів на конкретну посаду «junior front-end developer», основними вимогами якої є:

- D₁: «Якщо кандидат має навички роботи з JS, CSS3, HTML5, то він задовольняє (що відповідає вимогам)»;
- D₂: «Якщо він до d₁ має навички роботи з JQuery, то він більш ніж задовольняє»;
- D₃: «Якщо він до d₁ має навички роботи з Bootstrap, то він більш ніж задовольняє»;

- D₄: «Якщо він має всі обумовлені навички в d₁, d₂, d₃, то він бездоганний»;
- D₅: «Якщо не має навички роботи з JS або CSS3 або HTML5, він не задовольняє».

Аналіз п'яти інформаційних фрагментів дає п'ять критеріїв, які використовуються в ухваленні рішення:

- X₁ – навичка роботи з JS;
- X₂ – навичка роботи з JQuery;
- X₃ – навичка роботи з CSS3;
- X₄ – навичка роботи з HTML5;
- X₅ – навичка роботи з Bootstrap.

Маємо 4 кандидати для оцінювання їх компетенцій.

Інформація про кандидатів u₁ – u₄ є умовною, виходячи з обмежень щодо розголошення персональних даних. Оцінки кожного кандидата представлені у таблиці 1.

ТАБЛИЦЯ 1. Оцінки кожного кандидата по навичкам

Кандидати	Навички				
	JS	JQuery	CSS3	HTML5	Bootstrap
u ₁	0,8	0,5	0,5	0,1	0,3
u ₂	0,6	0,78	0,8	0,77	0,5
u ₃	0,3	0,8	0,33	0,5	0,2
u ₄	0,3	0,2	0,75	0,78	0,8

Таким чином, використовуючи алгоритм нечіткого виводу для підбору студентів ВНЗ на необхідну посаду, отримуємо точкові оцінки кандидатів. Точкова оцінка задовільності для альтернативи u_1 дорівнює 0,3, для u_2 дорівнює 0,7041, u_3 – 0,4257 та u_4 – 0,4957. В якості найкращої альтернативи вибираємо того кандидата, у

якого оцінка найвища. В таблиці 2 показано рейтинг кандидатів.

На рисунку 2 зображена діаграма рівнів навичок кандидатів. Таке представлення може підвищити рівень наочності та поліпшити сприйняття інформації.

ТАБЛИЦЯ П. РЕЙТИНГ КАНДИДАТІВ

Кандидати	Навички					Точкова оцінка
	JS	JQuery	CSS3	HTML5	Bootstrap	
u_2	0,6	0,78	0,8	0,77	0,5	0.7041
u_4	0,3	0,2	0,75	0,78	0,8	0.4957
u_3	0,3	0,8	0,33	0,5	0,2	0.4425
u_1	0,8	0,5	0,5	0,1	0,3	0.3

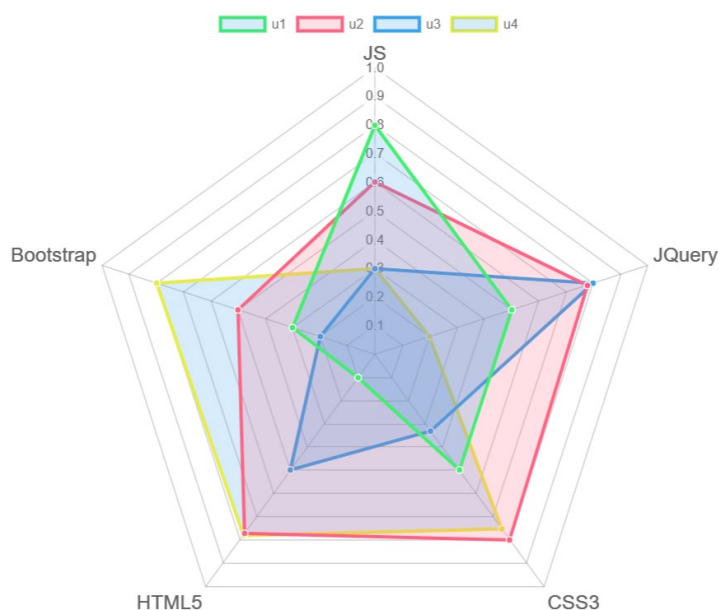


Рис. 2. Діаграма рівнів навичок кандидатів.

ВИСНОВКИ

Результати проведеного дослідження дозволяють зробити висновок, що використання методу нечіткої логіки сприяє підбору найбільш відповідних кандидатів на конкретну посаду і тим самим сприяє скороченню часу на обробку інформації про кандидатів HR-менеджером компанії, який взаємодіє не з десятками, а з сотнями кандидатів. Дана методологія дає інформацію про кандидатів, а остаточне рішення приймає людина, яка має на це повноваження.

Розглянута методологія оцінки компетентності фахівців може бути адаптована і для інших сфер з урахуванням специфіки підприємства та особливостей системи управління персоналом.

ЛІТЕРАТУРА REFERENCES

[1] А. В. Андрейчиков, О. Н. Андрейчикова, *Анализ, синтез, планирование решений в экономике*. Москва: Финансы и статистика, 2000.

[2] И. И. Коваленко, Т. А. Фарионова, С. Б. Приходько, *Методы принятия решений: Учебное пособие*, Николаев, НУК, 2009.
 [3] Н.П. Крап, В.М. Юзевич, *Управление развитием сложных систем*, Москва: Инфра-М, 2012.
 [4] R. E. Bellman, L. A. Zadeh, *Decision making in a fuzzy environment*, Management Science, 1970.
 [5] А. Я. Кибанов, *Управление персоналом организации*, Москва: Инфра-М, 2010.
 [6] С. В. Иванова, *Искусство подбора персонала. Как оценить человека за час*, Москва.: Альпина Паблишерз, 2011.
 [7] М. И. Магура М. Б. Курбатова, *Современные персонал – технологии*, Москва: Издательство «Управление персоналом», 2003.
 [8] А. Пегат, *Нечеткое моделирование и управление*, пер. с англ., Москва: БИНОМ. Лаборатория знаний, 2009.
 [9] В. В. Круглов, М. И. Дли, Р. Ю. Голубов, *Нечеткая логика и искусственные нейронные сети*, Москва: Физматлит, 2001.
 [10] Р. В. Кулик, *Применение принципов нечеткой логики в методологии BSC*, Москва: Экономические науки, 2009.
 [11] А. Н. Борисов, О. А. Крумберг, И. П. Федоров, *Принятие решений на основе нечетких моделей*, Рига: Зинатне, 1990.

Комп'ютерне Моделювання Поверхневих Стоків

Іван Боровець
Прикладна математика
Національний університет водної та екологічної інженерії
Рівне, Україна
borovets_ak15@nuwm.edu.ua

Ольга Мічута
Прикладна математика
Національний університет водної та екологічної інженерії
Рівне, Україна
o.r.michuta@nuwm.edu.ua

Computer Modeling of Surface Streams

Ivan Borovets
Applied Mathematics
National University of Water and Environmental Engineering
Rivne, Ukraine
borovets_ak15@nuwm.edu.ua

Olga Michuta
Applied Mathematics
National University of Water and Environmental Engineering
Rivne, Ukraine
o.r.michuta@nuwm.edu.ua

Анотація - Розглянуте питання про вплив атмосферних опадів на ґрунтовий укрив землі, і наслідки підвищення рівня підтоплення і затоплення на ділянці землі. Запропонований спосіб моделювання поверхневих стоків для передбачення проблемних ситуацій на місцевості пов'язаних з надмірними опадами.

Ключові слова - математичне моделювання, опади, затоплення, рівень води, ландшафт, поверхня, інфраструктура, сільське господарство

I. ВСТУП

Сучасна діяльність людини тісно пов'язана з таким процесом, як атмосферні опади, вони ж в конкретному випадку – дощі. В багатьох випадках надмірні опади призводять до негативних наслідків, таких як затоплення і підтоплення поверхні землі. Існує необхідність прогнозування процесів поверхневих стоків у випадках інтенсивних опадів з метою прийняття правильних рішень щодо запобігання негативних впливів, таких як ерозія ґрунтів і затоплення територій. Як в умовах містобудування так і в розвитку сільського господарства, важливо правильно оцінювати особливості рельєфу і його взаємодію з водними масами атмосферних опадів.

II. ПОСТАНОВКА ПРОБЛЕМИ

Як зрозуміло, внаслідок проходження атмосферних опадів водні маси в деяких своїх кількостях опиняються на суходолі. Після під дією сили тяжіння ці води природно рухаються у водойми та зниження рельєфу. Такий рух по поверхні землі називається поверхневий стік. Розрізняють три фази поверхневого стоку: схиловий, тальвеговий і річковий. Поверхневий стік виражають через такі показники: витрата води, об'єм стоку, модуль стоку, шар стоку.

Поверхневий стік відіграє дуже важливу роль в процесі кругообігу, визначаючи розподіл водних мас по земній поверхні, ґрунтових водах і водоймах. І людська діяльність починає конфліктувати з цим процесом. Поверхневі стоки вражають сільське господарство, порушуючи водний баланс на територіях угідь, викликаючи ерозію ґрунтів, знищуючи наявні посіви, та міську інфраструктуру, розмиваючи фундаменти будівель, затоплюючи вулиці та спиняючи звичну діяльність.

Ми будемо розглядати певну задану у загальному випадку ділянку території. Відомий її рельєф, особливості характеристик поверхні, метеорологічні

умови та кількість атмосферних опадів. Необхідно змодельовати дану територію згідно наявних характеристик, та за заданими умовами атмосферних опадів відобразити процес поверхневого стоку у зручному для аналізу вигляді, що дозволило би визначити критичні точки ділянки, в яких рівень затоплення прийматиме небажані значення.

Важливо відзначити, що мають місце два різних підходи до постановки такої задачі, а саме для території сільськогосподарської діяльності та території міста.

Перший випадок описує природну територію полів, луків, пасовищ та інших видів сільськогосподарських угідь. Дані території мають переважно одноманітний рельєф на одній визначеній ділянці, тому особливими зонами дослідження поверхневих стоків будуть межі різних ділянок, де рельєф змінюється для переходу у нову ділянку. Також важливо врахувати, що на таких природних територіях значною мірою на процеси поверхневого стоку впливає рослинний шар, що буде суттєво змінювати поведінку процесів стоку згідно своїх властивостей. Тут можна відмітити питання про лісові насадження, що завдяки своїй структурі дозволяють в певній мірі спинити процеси поверхневого стоку завдяки великій кількості води, що поглинаються деревами. Із цього ж питання виникає споріднена тема підняття рівнів ґрунтових вод та початок процесу заболочення місцевості внаслідок вирубування лісів, що власне ці води поглинали.

Другий випадок стосується принципів містобудування. Тут поверхневий покрив території являє собою деякі будівельні матеріали, такі як асфальт, бетон та інші. Процеси інфільтрації (просочування води в поверхню) на даній території будуть надзвичайно повільними, сприяючи надмірному накопиченню води на поверхні. Сюди ж додається той факт, що значна частина території має автоматично позбавлятися від накопичення води. Це в основному стосується дахів будинків, звідки системи водосточу справляють воду до підніжжя будинків, а тобто вулиць міста. Отже, на поверхню даної території відбудуться значні навантаження зі сторони впливу поверхневих стоків. Щоб уникнути такої ситуації необхідно ретельно проаналізувати території міста, оцінити вразливі точки та підготувати роботоздатні системи водосточу, що будуть перенаправляти рух водних мас з важливих місць території, такі як вулиці міста, у безпечні для відведення місця, такі як, наприклад, річки або канали, як можуть бути на території міста.

Отож, для усунення негативних наслідків для діяльності людини від природних процесів кругообігу води, необхідно правильно аналізувати можливі випадки та на їх основі приймати рішення про встановлення необхідних систем контролю рівня затоплень.

III. МАТЕМАТИЧНА МОДЕЛЬ ЗАДАЧІ

Для того, щоб аналізувати процес поверхневого стоку і рівень затоплюваності необхідно скласти математичну модель, в яку входить:

Досліджувана територія. Поверхня, на яку потрапляють атмосферні опади, її рельєф та характеристика покриву. На першому етапі розробки проекту використаємо переріз території, який можна представити в двовимірному просторі за допомогою деякої математичної функції кривої. Це досить велике спрощення все ж дозволить якісно оцінити процеси поверхневого стоку і дослідити закономірності, які пізніше можна буде розвинути у повноцінну тривимірну задачу.

Для визначення такої території у двовимірному випадку буде достатньо вхідного масиву значень висоти території. Таке представлення можна математично інтерполювати і отримати необхідну апроксимацію даних для подальшої роботи.

Кількість атмосферних опадів. Оскільки нас цікавить загальна кількість опадів, їх можна задавати стандартними значеннями міліметрів в годину, яке легко вимірюється метеорологічними приладами чи за доступними даними тих самих метеорологічних станцій.

Алгоритм обчислення заповнення. Будемо вважати, що заповнення території починається з заглибин, тобто точок локальних мінімумів. Заповнення продовжується, поки деякий визначений кількістю опадів об'єм води не наповнить заглибину, або не буде досягнутий перший локальний максимум заглибини, тобто місця де вода почне переливатися на сусідню ділянку, або утворювати цим нову, об'єднану і більшу за об'ємом.

A. Додаткові умови

Також варто відзначити, що в перспективах розвитку можна ввести додаткові умови, що покращують розрахунок рівня затоплень. Серед таких:

- **Випаровування.** Якщо передбачається тривалий час розгляду процесу затоплення, то знадобиться розрахунок швидкості випаровування вологи з визначеної ділянки.
- **Інфільтрація.** Значна частина вологи просочується в ґрунт, визначаючи цим процеси підтоплення, підвищення рівня ґрунтових вод, або ж попадання у інші шари процесу поверхневих стоків.

IV. АЛГОРИТМ ЗАПОВНЕННЯ

Розглянемо один із розроблених алгоритмів заповнення заглибини. Нехай в нас вже є задана в двовимірному випадку територія, яка була інтерполювана, і отримана апроксимуюча функція, що описує її рельєф. Також для кожної точки локального мінімуму є своя величина об'єму заповнення, яка описується відстанню між її точками максимуму. Тоді кожна заглибина матиме свій цільовий об'єм, який буде

заповнений внаслідок дії деякої величини v – характеристики кількості опадів.

Тепер можна приступити до моделювання процесу затоплення для однієї визначеної i -тої ділянки:

Описуємо ділянку точкою мінімуму, сусідніх точок максимумів, цільовим об'ємом заповнення, кроком величини заповнення, що дорівнюватиме тисячній долі різниці між значенням висоти меншого максимуму та мінімуму.

Початковий рівень води у заглибині дорівнює нулю, ліва і права межа відповідно дорівнюють координаті точки мінімуму.

Беремо поточний рівень води і додаємо до нього крок заповнення. Знаходимо ліву і праву точку рівня води за принципом: перебираємо аргументи функції вліво від точки мінімуму, поки різниця між поточним рівнем води та значенням функції не досягне заданої точності. В такому випадку маємо ліву межу нового рівня води.

Якщо був досягнутий лівий край, приймаємо ліву межу поверхні води за лівий максимум, знаходимо поточний об'єм наповнення як площу криволінійної трапеції, що утворилася між відрізком-рівнем води та кривою поверхні. Різницю між цільовим об'ємом і тим, що набраний, передаємо на сусідню ліву ділянку, а поточну ділянку вважаємо заповненою.

Аналогічні міркування проводяться щодо правої точки рівня води з відповідними поправками, рухаємось вправо, і умовою переповнення буде права межа.

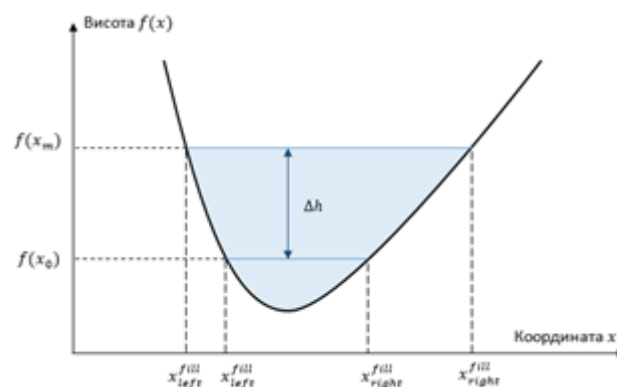


Рис. 1 Наповнення заглибини

ВИСНОВОК

В процесі дослідження було обґрунтоване теоретичне виникнення поверхневих стоків, їх вплив на місце проживання людей у містах і селах, на існуючу інфраструктуру містобудування та сільського господарства. Запропоновані до розгляду одна із математичних моделей процесу поверхневих стоків та її програмна реалізація для моделювання і відображення результату затоплень території внаслідок інтенсивних опадів.

ЛІТЕРАТУРА REFERENCES

- [1] Методи обчислень: конспект лекцій для студентів механіко-математичного факультету / В. В. Попов. – К.: Видавничо-поліграфічний центр "Київський університет", 2012. – 303 с.
- [2] Методи обчислень: Підручник: У 2 ч. / Гаврилук І.П., Макаров В.Л. – К.: Вища шк., 1995, – Ч. 1. – 367 с.
- [3] Методи оптимізації та дослідження операцій: Навч. посібник. / П.М. Мартинюк, О.Р. Мічута – Рівне: НУВГП, 2011. – 283 с.

Засоби псевдонедетермінованого потокового шифрування

Караван Владислав Русланович
факультет інформаційних технологій та комп'ютерної інженерії
Вінницький національний технічний університет
Вінниця, Україна
vlad30.96.12@gmail.com

Баришев Юрій Володимирович
к. т. н., доцент кафедри захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
yuriy.baryshev@gmail.com

Анотація—Розглянуто аналіз відомих потокових шифрів, які використовуються для захисту інформації від несанкціонованого доступу. Було розроблено структуру засобу псевдонедетермінованого потокового шифрування. Використання розробленого засобу дозволяє досягти збільшення швидкості потокового шифрування даних, що підтверджено експериментальними дослідженнями.

Ключові слова — шифрування, потокові шифри, псевдовипадковий, криптографія, алгоритм, генератор.

I. ВСТУП

Сучасні системи зв'язку оперують безперервним потоком даних, а там, де відбувається обмін інформацією, можливий і її витік. Використовуючи вразливості даних систем зловмисники можуть отримати несанкціонований доступ до приватних даних користувачів. Тому є сенс захищати системи зв'язку для усунення витіку інформації з них.

Найчастіше для реалізації поставленої задачі використовують криптографічні засоби. Стійкість таких засобів залежить від обраного криптографічного перетворення, зокрема шифрування. У системах зв'язку найчастіше використовують потокове шифрування, адже дані, що передаються, не є напередвизначеними.

Пристрої, що реалізують дані алгоритми, відповідно втратили свою актуальність, оскільки зараз є багато способів для їх обходу. Тому доцільно розробити спеціальний пристрій для потокового шифрування на базі нового алгоритму. Псевдонедетермінований метод за рахунок своїх особливостей є доволі складним для дослідження зловмисником, тому його використання забезпечить сильну стійкість розробленого пристрою в цілому [1, 2].

Об'єктом даного дослідження є процес потокового шифрування. Предметом – засоби псевдонедетермінованого потокового шифрування.

Метою дослідження є збільшення швидкості потокового шифрування.

Для її досягнення необхідно:

- проаналізувати відомі потокові шифри
- розробити метод псевдонедетермінованого шифрування;
- розробити пристрій псевдонедетермінованого потокового шифрування;

- реалізувати пристрій за допомогою мови опису апаратного забезпечення (HDL).

II. АНАЛІЗ ПРИСТРОЇВ НА ОСНОВІ ПОТОКОВОГО ШИФРУВАННЯ

Потокові шифри – це група симетричних шифрів, які шифрують кожен символ відкритого тексту незалежно від інших символів. В основі таких шифрів є операція XOR. З тією відмінністю, що гама для шифру XOR формується за певним алгоритмом, який є криптостійким генератором псевдовипадкої послідовності символів [3].

На сьогодні існує багато шифрів даного класу. Більш старі та ненадійні шифри – це, наприклад RC4 [4]. Сучаснішим й надійнішим вважають, наприклад, Salsa20 [4].

A. Апаратно-орієнтовані потокові шифри

Існують потокові шифри, які розраховані на їх апаратну реалізацію. До таких шифрів відносять A5, ORYX, RC4 [5, 6].

Шифр A5 використовується в системі мобільного цифрового зв'язку GSM для забезпечення конфіденційності даних між телефоном і базовою станцією [5]. Шифр заснований на побітовому додаванні за модулем два, що генерується псевдовипадковою послідовністю. У A5 псевдовипадкова послідовність реалізується на основі трьох регістрів зсуву з лінійним зворотнім зв'язком (РЗЛЗЗ). Регістри мають довжини 19, 22 і 23 біти відповідно [5].

Потоковий шифр ORYX передбачалося використовувати для засекречування даних, переданих в американських мережах стільникового мобільного зв'язку [6]. До складу генератора псевдовипадкової послідовності (ПВП) ORYX входять три 32-розрядних РЗЛЗЗ А, В і С, а також 8-розрядний S-блок з фіксованою таблицею заміни. Ключем є початкова заповнення трьох РЗЛЗЗ. В експортному виконанні застосовується алгоритм розгортання ключа, що скорочує ключовий простір до розміру, який легко перевіряється при пошуку ключа повним перебором.

B. Програмно-орієнтовані потокові шифри

SEAL (Software-Optimized Encryption Algorithm) – оптимізований під програмну реалізацію потоковий

шифр, автори алгоритму Ф. Роуер і Д. Копперсміт [6]. За допомогою відносно повільної операції SEAL на попередньому етапі перетворює ключову інформацію в набір таблиць загальним обсягом 3 Кб. Потім ці таблиці використовуються для швидкого шифрування інформації, витрачаючи близько п'яти машинних інструкцій на байт даних.

Однією з особливостей цього шифру є те, що він не є потоковим шифром в традиційному сенсі, а являє собою сімейство псевдовипадкових функцій. 160-бітовий ключ k і 32-бітове значення n (індекс) шифр перетворює в L -бітовий рядок $k(n)$. L може приймати будь-яке значення, що менше 64 кілобайт. Передбачається, що якщо k обирається випадково, то $k(n)$ обчислювально не буде відрізнятися від випадкової L -бітової функції від n [6].

Salsa20 – система потокового шифрування, розроблена Даніелем Бернштейном [7]. Алгоритм був представлений на конкурсі «eSTREAM», метою якого було створення європейських стандартів для шифрування даних, що передаються поштовими системами. Алгоритм використовує геш-функцію з двадцятьма циклами. Завдяки тому, що перетворення кожного стовпця і кожного рядка не залежать одне від одного, обчислення, необхідні для шифрування, легко розпаралелити. Це дає суттєвий вигравш в швидкості для більшості сучасних платформ [7].

HC-256 – вперше опублікований в 2004 році. 128-бітний варіант був представлений на конкурсі eSTREAM, метою якого було створення європейських стандартів для систем потокового шифрування [8]. Потоковий шифр HC-256 генерує ключову послідовність (keystream) довжиною до 2128 біт за допомогою 256-бітового ключа (secret key) і 256-бітного вектора ініціалізації (initialization vector).

HC-256 містить дві секретні таблиці, в кожній з яких 1024 32-бітних елемента. При кожному кроці оновлюється один елемент з таблиці за допомогою нелінійної функції зворотного зв'язку (feedback function). Через кожні 2048 кроків всі елементи двох таблиць будуть оновлені.

Отже, аналіз показав, що всі відомі потокові шифри вже були зламані зловмисниками, а тому має сенс розробити власний алгоритм, який буде відкритим для користувачів, за для доведення його стійкості, але його конкретна реалізація буде максимально захищена від зламу зловмисником.

III. РОЗРОБКА СТРУКТУРИ ПРИСТРОЮ

Для реалізації даного пристрою було розроблено його узагальнену схему. В загальному до складу пристрою повинні входити такі блоки:

- блок генерування гами;
- блок накладання гами;
- блок інтерфейсу.

Узагальнену структурну схему пристрою наведено на рисунку 1.

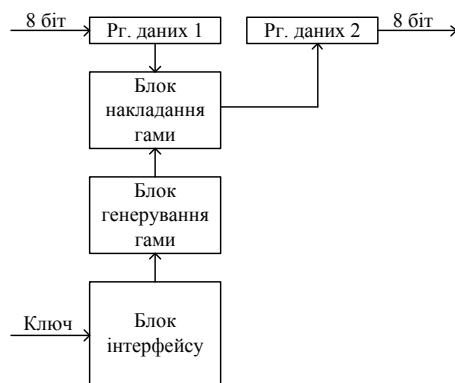


Рис. 1 – Узагальнена структурна схема пристрою.

До блоку інтерфейсу повинен надходити ключ, який потім буде розбиватися на декілька частин. Отримані частини передаються на блок генерування гами. В межах цього блоку відбувається безпосереднє формування гами, яка потім надходить до блоку накладання гами. До даного блоку також надходять дані, які необхідно зашифрувати. В результаті дані накладаються з гамою за допомогою або операції XOR, або операції інверсного XOR. На вибір однієї з цих операцій впливатиме один з регістрів зсуву з лінійним зворотнім зв'язком. Основний функціонал пристрою припадає на блок генерування гами та блок накладання гами.

A. Блок генерування гами

Формування гами відбувається за псевдодетермінованим методом [1, 2]. Блок генерування гами складається з двох трьохбітних та трьох восьмибітних регістрів зсуву з лінійним зворотнім зв'язком та мультиплексора.

До усіх регістрів надходять їх початкові стани. РЗЛЗЗ 1_1 та 1_2 формують число, яке потім надходить на РЗЛЗЗ 2_1 та 2_2 . Дане число задає кількість зсувів, що мають виконати РЗЛЗЗ 2_1 та 2_2 . В результаті на мультиплексор надсилаються дві можливі гами, що будуть використовуватися для накладання. Регістр Select керує входами мультиплексора, тобто займається власне вибором однієї з отриманих гам [2]. Загальна схема блоку генерування гами наведено на рисунку 2.

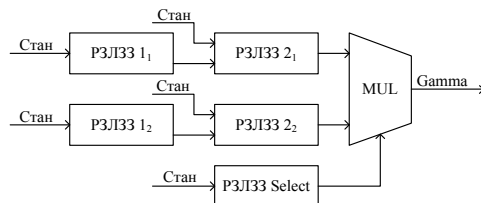


Рис. 2 – Структура блоку генерування гами

Отже, на виході мультиплексора отримується готова гама, що в результаті накладеться на дані.

B. Блок накладання гами

Сам процес шифрування відбувається безпосередньо у блоці накладання гами. Даний блок складається з двох восьмибітних регістрів, двох операційних пристроїв,

одного восьмибітного регістру зсуву з лінійним зворотнім зв'язком та мультиплексора.

До регістрів даних та гама надходять відповідно дані та гама, а до РЗЛЗЗ – початковий стан. Далі гама та дані побітово накладаються за допомогою XOR або інверсного XOR. Виходи операційних пристроїв є входами мультиплексора, якими керує регістр зсуву з лінійним зворотнім зв'язком. Загальна схема блоку накладання гама наведено на рисунку 3.

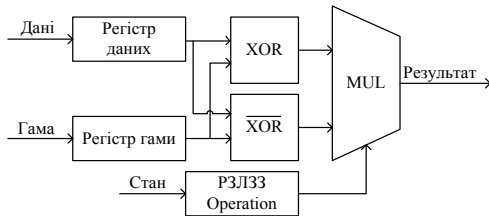


Рис. 3 – Структура блоку накладання гама

Отже, на виході мультиплексора отримується шифротекст, що готовий для передавання по системах зв'язку.

IV. РЕЗУЛЬТАТИ ТЕСТУВАННЯ

Запропоновану структуру пристрою було реалізовано за допомогою ПЛІС з використанням мови VHDL. Структуру пристрою реалізовано за допомогою таких основних сутностей:

- Zagalna;
- Twi;
- Gamma;
- Encrypt.

Для дослідження роботи реалізованого пристрою було розроблено сутність, яка моделює зовнішнє середовище. На рисунку 4 наведено часову діаграму сигналів пристрою.

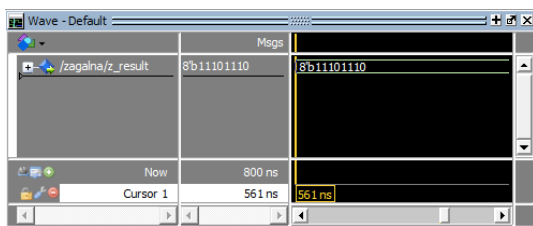


Рис. 4 – Часова діаграма сигналів пристрою

З рис. 4 та проведених під час досліджень блочного та інтеграційного тестування видно, що пристрій працює коректно та дозволяє шифрувати дані, однак потребує більшого часу на генерування гама порівняно з відомими апаратно-орієнтованими шифрами. Даний недолік пропонується усунути за рахунок збільшення частоти синхросигналів блока генерування гама.

Крім реалізації як спеціалізованого пристрою, запропонований метод реалізовано програмно як модуль засобу віддаленої автентифікації користувачів

[9]. Результати тестування даного модуля показали, що при використанні ключів та РЗЛЗЗ з розрядністю, кратною розрядності використовуваної обчислювальної платформи, даний засіб не породжує додаткової часової складності виконання алгоритму при роботі з бітовими масивами, як це виникає з іншими апаратно-орієнтованими потоковими шифрами при програмній реалізації. Відповідно даний шифр може конкурувати з програмно-орієнтованими потоковими шифрами в їх прикладних галузях застосування.

ВИСНОВКИ

Аналіз задач показав необхідність розробки нових засобів потокового шифрування даних у сучасних системах зв'язку. В основу розробки таких засобів взято метод псевдондетермінованого потокового шифрування.

Розроблено структуру пристрою, що реалізує даний метод. Апаратно-орієнтованість даного методу породила необхідність розробки спеціалізованого засобу на основі ПЛІС. Крім того, даний метод було реалізовано як програмний модуль для розв'язання задач, які ставлять перед програмно-орієнтованими потоковими шифрами.

Спроектований та розроблений пристрій можна використовувати в будь-яких системах зв'язку. Це може бути, наприклад, шифрування як GSM зв'язку у телефонах та станціях мобільних операторів, так і даних, що передаються організаціями між віддаленими відділами.

REFERENCES ЛІТЕРАТУРА

- [1] Караван В.Р. Метод псевдондетермінованого шифрування / В. Р. Караван // Тези доп. наук. - практ. конф., 22 березня 2018 р. / М-во освіти і науки України, ДВНЗ "Вінницький національний технічний університет". – Вінниця : ВНТУ, 2018. – Режим доступу: <https://conferences.vntu.edu.ua/index.php/all-fitki/all-fitki-2018/paper/view/5172/4543> (дата звернення 10.11.18 р.)
- [2] Патент України на корисну модель №128611 МПК Н 04 L 9/18. Спосіб псевдондетермінованого поточного шифрування / Барішев Ю. В., Караван В. Р.; заявник та патентовласник Вінницький національний технічний університет. – №u 2018 03873; заявл. 10.04.2018; опубл. 25.09.2018, Бюл. №18/2018
- [3] Горбенко І.Д. Прикладна криптологія. Теорія, практика, застосування / І.Д. Горбенко, Ю.І. Горбенко – Харків : Форт, 2013 – 534 с.
- [4] Петров А. А. Компьютерная безопасность / А. А. Петров – Москва: Лайт Лтд., 2000 – 448 с.
- [5] Шнайер Брюс. Прикладная криптография. Протоколы, алгоритмы и исходные тексты на языке С. 2-е издание. / Брюс Шнайер – Москва: Дело, 2003 – 610 с.
- [6] Мирский А.А. Поточные шифры. Книга 3 / А.А. Мирский, М.А. Иванов, А.В. Асосков – Москва: Кудиц-Образ, 2003 – 336 с.
- [7] Salsa20 [Електронний ресурс]. Режим доступу: URL: <http://www.e.stuipr.eu.org/stream/salsa20pf.html> – Назва з екрану. (дата звернення 10.11.18 р.)
- [8] HC256 [Електронний ресурс]. Режим доступу: URL: <http://www.e.stuipr.eu.org/stream/hc256.html> – Назва з екрану. (дата звернення 11.11.18 р.)
- [9] Барішев Ю. В., Караван В. Р. Комп'ютерна програма "Метод автентифікації на основі псевдондетермінованого шифрування". Свідчення про реєстрацію авторського права на твір № 78849, Україна, - власник свідоцтва Барішев Ю. В., Караван В. Р.. Заявлено 23.04.2018 : опубліковано 05.05.2018.

The Throughout Modeling Of E-UTRA Network with Adaptive Modulation

Anna Artemenko-Didenko

Television and radio broadcasting department

O.S. Popov Odessa National Academy of telecommunications

Odesa, Ukraine

anna.ivanivna.artemenko.didenko@gmail.com

Dmytro Makoveyenko

State enterprise «Ukrainian scientific research institute of radio and television»

O.S. Popov Odessa National Academy of telecommunications

Odesa, Ukraine

dikatama.dm@gmail.com

Abstract — This paper focuses on the analysis of possible throughput in network, that uses E-UTRA.

Previously, scientists considered the effectiveness of the AMC system, it was calculated the throughput according to Shannon's theorem in that research. The purpose of this work is to compare the throughput according to Shannon's theorem and real throughput according to the AMC system.

These results showed that even a smaller signal-to-noise ratio is able to provide signal transmission by the channel with acceptable errors.

Keywords—LTE, 4G, full frequency reuse, FFR, OFDM, orthogonal frequency-division multiplexing, MCS, AMC, BLER, block error rate, adaptive modulation and coding, E-UTRA, throughput.

I. INTRODUCTION

New generations of cellular communications appear every year with a smaller interval between them. Telecommunication companies are introducing new technologies in this field. But as in any other area, the question is to optimize the use of all resources with the satisfaction of as many quality indicators as possible.

Recently, telecom operators in Ukraine began to provide services of the 4G standard (LTE - long-term evolution). This standard uses OFDM (orthogonal frequency-division multiplexing) multiplexing (modulation).

One of the simplest methods of frequency domain planning in OFDM systems is full frequency reuse (FFR). This method has advantages and also disadvantages. The benefits are simplicity and higher throughput compared to other methods. A major disadvantage is the high level of noise, because of the reason that it uses the same frequency and as a result the signal / noise ratio (SNR) decreases.

The use of adaptive modulation and coding (AMC) also helps to use the type of modulation that is optimal for a certain signal / noise ratio. Such modulation should provide the minimum possible value of the BLER (block error rate).

Previously, scientists considered the effectiveness of the AMC system [1], it was calculated the throughput according to Shannon's theorem in that research. The purpose of this work is to compare the throughput according to Shannon's theorem and real throughput according to the AMC system.[2,3]

II. THE IMPACT OF LOSSES ON THE AUTOMATIC CHOICE OF MODULATION TYPE IN THE AMC SYSTEM

The design of LTE networks is to improve its physical layer (E-UTRA Radio Access Network).

The E-UTRA network allows adapting to a subscriber terminal by applying various types of modulation (Table.I). [4] Each type can be used with a predetermined threshold

value of signal / noise (SNR), that depends on the losses in signal propagation.

The adaptive change of the modulation code scheme, as the main part of the AMC procedure, includes the ability to determine the SNR threshold value and, using the feedback circuit, change the transmission parameters and the type of MCS.

№ MCS	Modulation type	Code rate	Spectral efficiency, bit/s/Hz	SNR, dB
1	QPSK	0.076	0.1523	-6,8
2	QPSK	0.12	0.2344	-5,2
3	QPSK	0.19	0.3770	-3,1
4	QPSK	0.3	0.6016	-2,2
5	QPSK	0.44	0.8770	-0,8
6	QPSK	0.59	1.1758	1,4
7	16QAM	0.37	1.4766	3,3
8	16QAM	0.48	1.9141	5,3
9	16QAM	0.6	2.4063	7
10	64QAM	0.45	2.7305	10,5
11	64QAM	0.55	3.3223	11,6
12	64QAM	0.65	3.9023	14,1
13	64QAM	0.75	4.5234	16,8
14	64QAM	0.85	5.1152	18
15	64QAM	0.93	5.5547	21,5

III. THE SIGNAL POWER LOSSES

A. The losses in the main channel signal

This paper considered the changing parameters at 10 points (9 equal segments on the line of movement of the mobile station), as shown in the figure (Fig.1)

The Extended Hata model for the medium-sized cities [5] was used as the basis for calculating the power losses of the signal. The input data are shown in the Table 2.

$$L = 69.55 + 26.16 \log(f) - 13.821 \log(h_r) + [44.9 - 6.55 \log(h_r)] \log(d)^b - \alpha \quad (1)$$

Where:

L :median propagation losses (dB);
 f : frequency (MHz);

h_{tr} : transmitter antenna height above ground (m);
 h_r : receiver antenna height above ground (m);
 d : distance between transmitter and receiver (km);
 $b=1$, if $d \leq 20$ km;

$$\alpha = (1.1 \cdot \log(f) - 0.7)h_r - (1.56 \cdot \log(f) - 0.8) \quad (2)$$

TABLE II. INPUT DATA

Parameter	Value
Frequency	1805-1880 MHz (Band3)
Maximum distance	1 km
Transmitting antenna height	30 m
Receiving antenna height	2 m
Bandwidth	15 MHz
Transmitter power	10 Wt
Antenna gain	16.5 dBi
Feeder Loss	2 dB
Bandwidth	4.5 MHz

B. The summary interfering signal power including losses

There are six signals with the same frequency around the central cell (Fig.1), that they make interferences for every signal in the cell.[6]

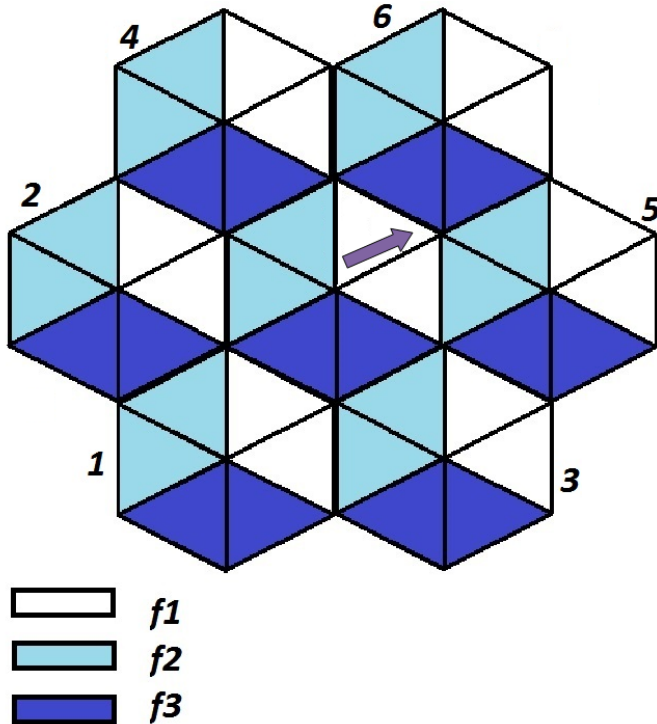


Fig. 1. Central cell with neighbour cells.

The distance from every neighbor interfering antenna to every point of the movement is different, but we can say that

from 1 and 2 cells it will be the same. The same can be noted for the pairs 3,4 and 5,6.

$$R_{1,2} = \sqrt{\left(\frac{2\sqrt{3} \cdot d}{2}\right)^2 + R_i^2} + \frac{2 \cdot 2\sqrt{3} \cdot d}{2} \cdot R_i \cdot \frac{\sqrt{3}}{2} \quad (3)$$

$$R_{3,4} = \sqrt{\left(\frac{2\sqrt{3} \cdot d}{2}\right)^2 + R_i^2} \quad (4)$$

$$R_{5,6} = \sqrt{\left(\frac{2\sqrt{3} \cdot d}{2}\right)^2 + R_i^2} - \frac{2 \cdot 2\sqrt{3} \cdot d}{2} \cdot R_i \cdot \frac{\sqrt{3}}{2} \quad (5)$$

Where:

R_i : distance from the antenna to 10 points on line of movement (km)

Considering all the losses and noise from the signal source and the sources of interferences, the signal-to-interference-plus-noise ratio was calculated. (Table III)

TABLE III. SIGNAL/NOISE RATIO

Distance, m	Signal/noise without interference, dB	Signal/noise with interference, dB
100	49.9	39.4
200	39.3	29.1
300	33	23.3
400	28.7	19.1
500	25.3	15.8
600	22.5	13.1
700	20.1	10.7
800	18	8.5
900	16.3	6.5
1000	16.7	4.6

IV. THROUGHOUT

For the calculation of throughput, the Shannon theorem was used.

$$R_{C1} = \frac{\Delta F}{3 \cdot 10^6} \cdot \log_2(1 + 10^{0.1(P-L-N)}) \quad (6)$$

$$R_{C2} = \frac{\Delta F}{3 \cdot 10^6} \cdot \log_2(1 + 10^{0.1 \cdot h}) \quad (7)$$

Where:

R_{C1} : throughput without noise

R_{C2} : throughput with noise

P: power of transmitter

N: level of noise

h: signal-to-noise ratio

But to compare the parameters was also used the calculation of throughput according to the adaptive MCS.

$$R_C = \gamma \cdot \Delta F \quad (8)$$

Where:

R_C : throughput according to the adaptive MCS

γ : Spectral efficiency, bit/s/Hz

ΔF : bandwidth.

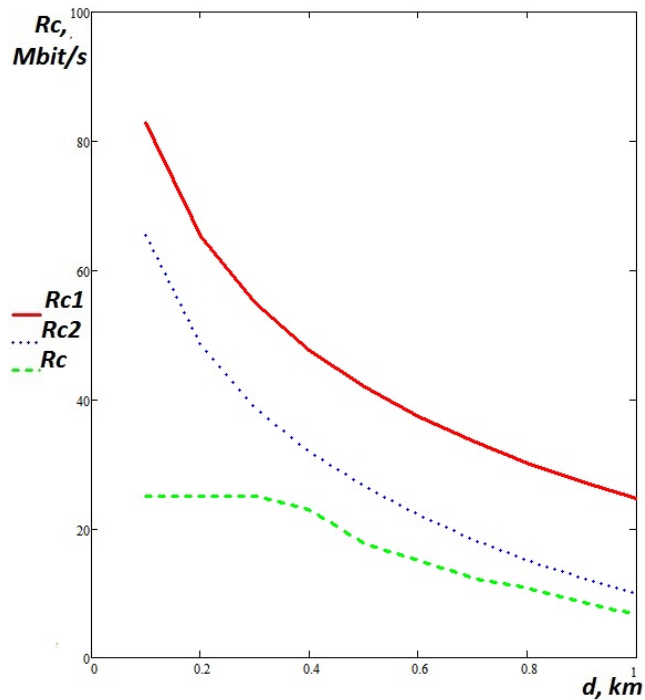


Fig. 2. Throughput

From the graphs (Fig.2), it can be seen that the channel throughput, calculated according to the Shannon theorem, in the vicinity of the station is much larger than that calculated by the real parameters of the MCS system.

The results of calculations, that is illustrated on Fig.2 is in the Table IV

TABLE IV. THROUGHOUTS COMPARATION

R_{C1}	R_{C2}	R_C	Modul. type	Code rate	Spectral efficiency, bit/s/Hz	SNR, dB
82.86	65.42	25	64QAM	0.93	5.5547	21,5
65.24	48.43	25	64QAM	0.93	5.5547	21,5
54.94	38.66	25	64QAM	0.93	5.5547	21,5
47.64	31.79	23	64QAM	0.85	5.1152	18
41.98	26.47	17.6	64QAM	0.65	3.9023	14,1
37.37	22.1	15	64QAM	0.55	3.3223	11,6
33.48	18.37	12.3	64QAM	0.45	2.7305	10,5
30.13	15.13	10.8	16QAM	0.6	2.4063	7
27.19	12.3	8.6	16QAM	0.48	1.9141	5,3
24.56	9.85	6.6	16QAM	0.37	1.4766	3,3

The throughput of the channel calculated by real indicators of AMC is much lower than that calculated by the Shannon theorem. At the first calculated point, it differs near 3 times more, at the second - about 2 times more, at the third - by 1.5 times. As the receiving station becomes more distant from the transmitting station, the difference in throughput decreases. But this is true, only up to a certain distance. In the remaining gap, the difference between them is approximately the same (1.5 times).

REFERENCES ЛІТЕРАТУРА

- [1] P.I. Chernobay. Analysis of increase capacity E-UTRA using SFR / P.I. Chernobay, D.A. Makoveyenko // Refrigeration engineering and technology. – 2014. – № 1. – P. 76 – 80.
- [2] Tihvinskij V.O. Seti mobilnoj svjazi LTE. Tehnologii i arhitektura / V.O. Tihvinskij, S. V. Ter-ent'ev, A. B. JURchuk – M.: JEko-Trendz, 2010. – 284 p.
- [3] Babkov V.U. Seti mobilnoj svjazi. Chastotno-territorialnoe planirovanie / V.U. Babkov, Voznjuk M.A., Mihajlov P.A. – M.: Gorjachaja linija-Telekom, 2007. – 224 p.
- [4] Sajid Hussain, "Dynamic Radio Resource Management in 3GPP LTE", Blekinge Institute of Technology, 2009. – 58 p.
- [5] Report ITU-R SM.2028-2. Monte Carlo simulation methodology for the use in sharing and compatibility studies between different radio services or systems, 2017. – 57 p.
- [6] Specification 3GPP TR 25.913. Requirements for evolved UTRA (E-UTRA) and evolved UTRAN (E-UTRAN), 2009. – 20 p.

Огляд Методів Моделювання Капілярних Поверхонь для Методу Висячої Краплі

Сусак Я.Я.

Кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет
імені Василя Стефаника
Івано-Франківськ, Україна
susak.slava@gmail.com

Малько О.Г.

Кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет
імені Василя Стефаника
Івано-Франківськ, Україна
malko.pochta@gmail.com

An Overview of the Methods of Modeling Capillary Surfaces for the Method of Pendant Drop

Susak Y.Y.

Department of computer sciences and information systems
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine
susak.slava@gmail.com

Malko O.G.

Department of computer sciences and information systems
Vasyl Stefanyk Precarpathian National University
Ivano-Frankivsk, Ukraine
malko.pochta@gmail.com

Анотація—Визначено математичні основи методу висячої краплі.

Ключові слова: висяча крапля, діаметр, екватор, екваторіальний діаметр, конічна насадка, крапля з шийкою, кругла трубка.

Abstract— The mathematical foundations of the method of pendant drop were determined.

Keywords: pendant drop, diameter, equator, equatorial diameter, conical nozzle, drop with neck, circular tube.

I. ВСТУП

Об'єктом дослідження в даній роботі є метод висячої краплі.[1]

Метод був створений Ендрасом, Хаусером і Такером в ХХ столітті і активно розвивається до цих пір.

Актуальність роботи пояснюється необхідністю виконувати складні розрахунки стану висячої краплі, що вимагає математичного моделювання.

Метою роботи є провести аналіз наукових праць, що описують метод висячої хвилі та вивести необхідні для моделювання математичні формули.

Робота пов'язана з науковими працями Ендраса, Хаусера, Такера, Уортінгтона, Місака, Ро, Бачетти, Рамакрішна, Бабу.

II. ОГЛЯД МЕТОДУ ВИСЯЧОЇ КРАПЛІ

Міжфазна тензіометрія вивчає явище поверхневого натягу.

Поверхневий натяг - термодинамічна характеристика поверхні поділу двох середовищ, що знаходяться в рівновазі фаз, обумовлена роботою оборотного ізотермокінетического утворення одиниці площі цієї поверхні поділу за умови, що температура, обсяг системи і хімічні потенціали всіх компонентів в обох фазах залишаються постійними [2].

Поверхневий натяг має подвійний фізичний зміст - енергетичний (термодинамічний) і силовий (механічний). Енергетичне (термодинамічне) визначення: поверхневий натяг - це питома робота збільшення поверхні при її розтягуванні за умови сталості температури. Силове (механічне) визначення: поверхневий натяг - це сила, що діє на одиницю довжини лінії, яка обмежує поверхню рідини.

Сила поверхневого натягу спрямована по дотичній до поверхні рідини, перпендикулярно до ділянки контуру, на який вона діє і пропорційна довжині цієї ділянки. Коефіцієнт пропорційності - сила, яка припадає на одиницю довжини контуру - називається коефіцієнтом поверхневого натягу. Він вимірюється в ньютонках на метр. Але більш правильно дати визначення поверхневого натягу, як енергії (Дж) на розрив одиниці поверхні (m^2). У цьому випадку з'являється ясний фізичний зміст поняття поверхневого натягу.

У 1983 році було доведено теоретично і підтверджено даними з довідників, що поняття поверхневого натягу рідини однозначно є частиною поняття внутрішньої енергії (хоча і специфічною: для симетричних молекул близьких за формою до кулястих). Наведені в цій журнальній статті формули дозволяють для деяких речовин теоретично розраховувати значення поверхневого натягу рідини за іншими фізико-хімічними властивостями, наприклад, по теплоті пароутворення або по внутрішньої енергії.

Поверхневий натяг може бути на кордоні газоподібних, рідких і твердих тіл. Зазвичай мається на увазі поверхневий натяг рідких тіл на кордоні «рідина - газ». У разі рідкої поверхні розділу поверхневий натяг правомірно також розглядати як силу, що діє на одиницю довжини контуру поверхні і яка прагне скоротити поверхню до мінімуму при заданих обсягах фаз.

У загальному випадку прилад для вимірювання поверхневого натягу називається тензіометр.

Одними з основних методів, що застосовуються в міжфазній тензіометрії, є методи лежачої та висячої краплі. Метод висячої краплі буде детально розглянутий в цій роботі.

Творцями методу висячої краплі прийнято вважати Ендреса, Хаусера і Такера [3], які з аналізу профілю висячої краплі вперше отримали значення поверхневого натягу з достатньою точністю. Раніше подібні, але невдалі спроби здійснювались низкою авторів, в тому числі Уортінгтоном [4] і Фергюсоном [5].

В роботі [3] використовувались висячі краплі з екватором, на профілю яких виміряли екваторіальний діаметр $2x_e$ і діаметр $2x_{(m)}$ на відстані mx_e (при $m=2$) від вершини краплі, як зображено на рис. 1:

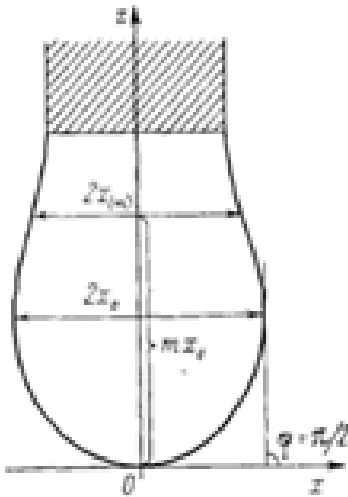


Рис. 1. Профіль висячої краплі

Крапля звисає з фіксованого круглого отвору радіусу r , розміщеного на висоті z_0 , де z – координата по вертикалі (додатній напрям вгору) і $z = 0$ – рівень, на якому тиск в рідині рівний нулю.

Якщо V – “виштовхуваний” об’єм рідини, Ω – поверхня краплі (межа середовищ рідина-повітря) та $|\Omega|$ – площа цієї поверхні, то потенціальна енергія конфігурації задається формулою [6]

$$E_0(\Omega) = \sigma|\Omega| + \rho \cdot g \int_{\Omega} z dV$$

де σ – поверхневий натяг,

ρ – густина рідини,

g – постійна сили тяжіння.

Умовою рівноваги є те, що перша варіація потенціальної енергії $\partial(E_0(\Omega, N))$ рівна нулю для всіх збурень N поверхні Ω , які перпендикулярні до неї і стають рівні нулю на границі отвору. Відповідно рівняння Ейлера матиме наступний вигляд

$$2H = -\chi u$$

$$\chi = \rho \cdot g / \sigma \text{ – постійна.}$$

Тут H – середня кривизна поверхні Ω , знак цієї кривизни вибирається так, щоб вона була додатня на кінчику краплі. Виконавши, якщо потрібно, необхідне

нормування, можна рахувати, що. Умова стійкості полягає в тому, що друга варіація енергії додатня для будь-якого нетривіального нормального збурення поверхні:

$$\partial^2 E_0(\Omega, N) > 0 \text{ для всіх } N \neq 0$$

Процедура обчислення γ в [1] здійснювалась на основі залежності $1/4cx_e^2$ від $x_{(m)} / x_e$ (остання величина знаходиться з експерименту). Ця залежність була знайдена емпіричним шляхом, а саме з незмінних діаметрів $2x_e$ і $2x_{(m)}$ для різних по розміру і формі крапель води. При цьому величина $x_{(m)} / x_e$ варіювалась від 0,7 до 1.

В більш пізніх роботах отримання залежності $1/4cx_e^2$ від $x_{(m)} / x_e$ також для $m = 2$ здійснювалось вже на основі інтегрування диференціального рівняння

$$\frac{d\varphi}{d(\frac{s}{b})} + \frac{\sin \varphi}{\frac{x}{b}} = 2 + \beta(\frac{z}{b})$$

в безрозмірному вигляді. Точність отриманих даних була вище, більше того, охоплювалась ширша область зміни $x_{(m)} / x_e$.

Мисак в [7] на основі табличних залежностей $1/4cx_e^2$ від $x_{(m)} / x_e$ (для $m=2$) отримав набір апроксимаційних формул, кожна з яких добре працює в інтервалі:

$$T = \left(\frac{0,32720}{S^{2,56651}} \right) - 0,97553S^2 + 0,84059S - 0,18069$$

$$\text{(для } 0,401 \leq S \leq 0,46 \text{)}$$

$$T = \left(\frac{0,31968}{S^{2,59725}} \right) - 0,46898S^2 + 0,50059S - 0,13261$$

$$\text{(для } 0,46 \leq S \leq 0,59 \text{)}$$

$$T = \left(\frac{0,31522}{S^{2,62435}} \right) - 0,11714S^2 + 0,14756S - 0,05282$$

$$\text{(для } 0,59 \leq S \leq 0,68 \text{)}$$

$$T = \left(\frac{0,31345}{S^{2,64267}} \right) - 0,09155S^2 + 0,14701S - 0,05877$$

$$\text{(для } 0,68 \leq S \leq 0,9 \text{)}$$

$$T = \left(\frac{0,30715}{S^{2,84636}} \right) - 0,69116S^3 + 1,08315S^2 - 0,18341S - 0,20970$$

$$\text{(для } 0,9 \leq S \leq 1 \text{)}$$

де $T = 1/4cx_e^2$.

Можливості описаного в [3] підходу були розширені в роботах [8], [9]: наряду з екваторіальним діаметром $2x_e$ в [5] пропонувалось вимірювати діаметр $2x_{(m)}$ на відстані mx_e від вершини краплі для $m = 1,6; 1,8; 2,2; 2,4$, а в [9] – для $m=1,25; 1,5; 1,75; 2,25$ і $2,5$ (в обох роботах випадок $m=2$ також розглядався). Що стосується залежності, то в [5] вона мала такий же функціональний вигляд, що і в [3], а в [9] для знаходження γ використовувалась таблична залежність $x_{(m)} / x_e$ від $x_e c^{0,5}$.

Інший варіант методу висячої краплі запропонований Уінкелом [10]. Для його реалізація необхідна крапля з шийкою (рис. 2), а вимірюваними величинами служать екваторіальний діаметр краплі $2x_e$ і діаметр шийки $2x_n$.

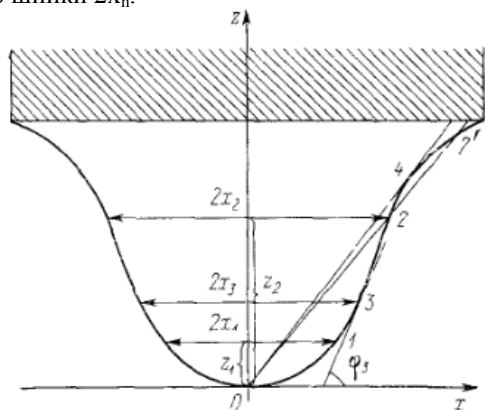


Рис. 2. Профіль висячої краплі з шийкою

Описана в цій роботі схема пошуку величини γ з $2x_e$ і $2x_n$ по суті така ж, як і в [3]: спочатку знаходиться відношення діаметрів, потім з залежності, що пов'язує це відношення і безрозмірний екваторіальний діаметр визначається c і звідси відповідно γ .

III. ПОРІВНЯННЯ ВАРІАНТІВ МЕТОДУ ВИСЯЧОЇ КРАПЛІ І ВИВЕДЕННЯ ФОРМУЛ

Порівнюючи два варіанти методу — представлений в [3], [8], [9] і в [10], — варто зазначити, що варіант Уінкела, незважаючи на труднощі в формуванні стійкої краплі з шийкою, з експериментальної точки зору більш точний, бо вимірювані в ньому відрізки $2x_e$ і $2x_n$ мають однакову похибку. В першому варіанті діаметр $2x_n$ визначається з більшою похибкою, оскільки до власної похибки $2x_n$ додається подвоєна похибка вимірювання $2x_e$.

Розглянемо ще один варіант методу висячої краплі, в якому, на відміну від традиційного способу формування краплі на кінці круглої трубки, використовується конічна насадка [11]. Автор цього варіанту запропонував вимірювати в ході досвіду величини x_t і z_t — декартові координати вершини конуса (рис. 3). При цьому повинні бути відомі крайовий кут θ і кут при вершині конуса 2β :

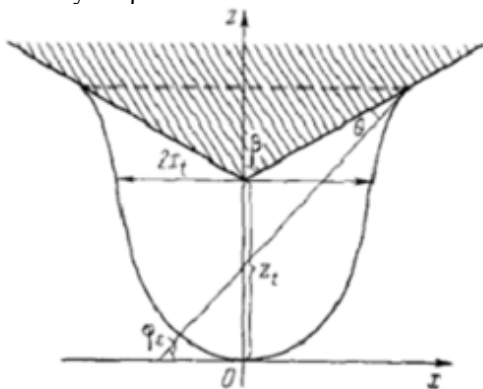


Рис. 3. Профіль висячої краплі на конічній насадці

З урахуванням вказаних величин пара додаткових умов буде мати вигляд:

$$z|_{x=x_t} = z_t$$

$$\left. \frac{x}{z - z_t} \right|_{\varphi=\varphi_c=\frac{\pi}{2}-\beta+\theta} = \operatorname{tg} \beta$$

Зробимо заміну: $x=x(\varphi, b, c)$, $z-z_t = z(x, b, c) - z_t$. Звідси $x/(z-z_t)=f(\varphi, b, c, z_t)$, отримуємо систему алгебраїчних рівнянь з двома невідомими b

[8] для спрощення розрахунків Бабу знайшов інтерполяційні поліноміальні формули, які описують більш частину діапазону z_t/x_t :

$$x_t \sqrt{c} = d_1 + d_2 \frac{z_t}{x_t} + d_3 \left(\frac{z_t}{x_t} \right)^2$$

$$z_t \sqrt{c} = e_1 + e_2 \frac{z_t}{x_t} + e_3 \left(\frac{z_t}{x_t} \right)^2 +$$

$$+ e_4 \left(\frac{z_t}{x_t} \right)^3 + e_5 \left(\frac{z_t}{x_t} \right)^4 + e_6 \left(\frac{z_t}{x_t} \right)^5$$

ВИСНОВКИ

В даній роботі було здійснено аналіз і узагальнення наукових праць провідних вчених, що створювали і розвивали метод висячої краплі, виведено основні математичні формули і рівняння, що описують поведінки висячої хвилі.

Дані результати можуть бути використані в дослідженнях методу висячої краплі в міжфазній тензіометрії.

Актуальність роботи пояснюється необхідністю виконувати складні розрахунки стану висячої краплі, що вимагає математичного моделювання.

Проведення досліджень за даною тематикою є доцільним, оскільки дозволяє суттєво спростити використання методу висячої краплі в хімії, що може мати використання в теоретичній хімії, прикладній хімії, медицині та інших галузях, що пов'язані з хімією.

ЛІТЕРАТУРА REFERENCES

- [1] Русанов А. И. Межфазная тензиометрия / А. И. Русанов, В. А. Прохоров. - СПб. : Химия : Санкт-Петербург. отд-ние, 1994. - 397
- [2] Адамсон А. Физическая химия поверхностей / Пер. с англ. - М.: Мир, 1979.-568с.
- [3] Andreas J M, Hauser E A, Tucker W B //J Phys Chem 1938 V 42, № 8 [с.1001 — 1019]
- [4] Worthington A M //Proc Roy Soc 1881 V 32, № 214 [с.362 — 377]
- [5] Ferguson A //Phil Mag Ser 6 1912 V 23, № 135 [с.417 — 430]
- [6] Финн, Р. Равновесные капиллярные поверхности. Математическая теория :пер. с англ. /Р. Финн ; Москва : Мир, 1989. 312 с.
- [7] Misak M D //J Colloid Interface Sci 1968 V 27, №1 [с.141-142]
- [8] Roe R -J, Bacchetta V I, Wong P M G //J Phys Chem 1967 V 71, № 13 [с.4190 — 4193]
- [9] Ramakrishnan S, Princz J F, Hartland S //Indian J Pure Appl Phys 1977 V 15, № 4 [с.228 — 233]
- [10] Winkel D //J Phys Chem 1965 V 69, № 1 [с.348 — 350]
- [11] Babu S R //J Phys Chem 1986 V 90, № 18 [с.4337 — 4340]

Вдосконалення Параметрів АЦП Шляхом Введення Надлишковості

Святослав Гриджан
кафедра комп'ютерних наук та
інформаційних систем
Прикарпатський національний
університет імені Василя Стефаника
Івано-Франківськ, Україна
sviatyk.hrydzhan@icloud.com

Тарас Піварчук
Кафедра нафтогазової геофізики
Івано-Франківський національний
технічний університет
нафти і газу
Івано-Франківськ, Україна
taras.pivarchuk@gmail.com

Михайло Петришин
кафедра комп'ютерних наук та
інформаційних систем
Прикарпатський національний
університет імені Василя Стефаника
Івано-Франківськ, Україна
m.l.petryshyn@gmail.com

Improving the ADC Parameters by Introducing Redundancy

Sviatoslav Hrydzhan
Department of Computer Science and
Information Systems
Vasyl Stefanyk Precarpathian National
University
Ivano-Frankivsk, Ukraine
sviatyk.hrydzhan@icloud.com

Taras Pivarchuk
Department of oil and gas geophysics
Ivano-Frankivsk National Technical
oil and gas University
Ivano-Frankivsk, Ukraine
taras.pivarchuk@gmail.com

Mykhailo Petryshyn
Department of Computer Science and
Information Systems
Vasyl Stefanyk Precarpathian National
University
Ivano-Frankivsk, Ukraine
m.l.petryshyn@gmail.com

Анотація— В даній роботі здійснено аналіз основних підходів покращення параметрів функціонування сучасних аналого-цифрових перетворювачів шляхом введення структурної надлишковості, а саме структурно-схематичної, функціонально-алгоритмічної та інформаційної надлишковості.

Ключові слова —аналого- цифрові перетворювачі, структурно- схематична надлишковість, функціонально-алгоритмічна надлишковість, інформаційна надлишковість.

Abstract— In this paper, an analysis of the main approaches to improving the parameters of the functioning of modern analog-digital converters by introducing structural redundancy, namely, structural-schematic, functional-algorithmic and informational redundancy.

Keywords— analogue-digital converters, structural and schematic redundancy, functional-algorithmic redundancy, information redundancy.

I. ВСТУП

Використання структурно-схематичної, функціонально-алгоритмічної та інформаційної надлишковості [1] дозволяє покращити характеристики і пришвидшити перебіг процесу АЦ перетворення за рахунок компенсування похибок або коригування результату перетворення. Застосування систем числення з ваговою надлишковістю (СЧВН) у техніці АЦП і ЦАП започатковано в СРСР науковою школою професора О.П. Стахова [2] та в наступному продовжено школою професора О.Д. Азарова [1,3] незалежно від аналогічних робіт, що розгорнулися водночас у США дослідниками фірми Intersil Inc.

II. СТРУКТУРНО-СХЕМОТЕХНІЧНА НАДЛИШКОВІСТЬ

Структурно-схематична надлишковість полягає в використанні додаткового обладнання, одним з методів реалізації є комбінування алгоритмів перетворення даних, що передбачає знаходження таких комбінацій класичних алгоритмів, які дозволяють за рахунок незначного ускладнення структури АЦП підвищити точність та швидкість процесу АЦ перетворення. Різноманіття комбінованих АЦП зумовлене поєднанням здебільшого двох алгоритмів з трьох типових, а саме методу послідовного наближення; методу порозрядного врівноваження та методу безпосереднього перетворення [4]. Можливим є побудова наступних варіантів комбінованих АЦП (КАЦП):

- КАЦП на основі порозрядного кодування та порозрядного врівноваження (послідовного наближення);
- КАЦП на основі порозрядного врівноваження (послідовного наближення) та безпосереднього перетворення;
- КАЦП на основі порозрядного кодування та безпосереднього перетворення.

Функціонування КАЦП полягає в тому, що діапазон перетворення, представлений n розрядами вихідного коду, розбивається на два піддіапазони з розрядностями перетворення n_1 та n_2 в співвідношенні, закладеному архітектурами складових компонентів АЦП, що забезпечують максимальний вииграш часу перетворення. Тобто, в піддіапазоні із розрядністю перетворення n_1 працює один метод перетворення, в n_2 – інший, відмінний від попереднього. Проте існують техніки, коли діапазон розрядності n розбивається на

кілька рівних піддіапазонів, в кожному з яких застосовано один і той самий метод перетворення.

Перспективним поруч з методами комбінованих структур є підхід, пов'язаний з введенням структурно-схемотехнічної надлишковості, який дозволяє уникнути фізичного впливу на елементи схем[8, 9]. В разі використання ЦАП на основі двійкової системи числення, зменшення статичних похибок досягається корекцією вихідної величини шляхом введення поправки в аналоговій формі, що формується додатковим корегувальним ЦАП. В цьому випадку перетворюваний код корекції K_{ex} поступає (рис. 1), на вхід основного ЦАП і в цифровий обчислювальний пристрій (ЦОП). Вихідний результат перетворення $A_{вих}$ формується за допомогою аналогового суматора. Такий принцип підвищення точності ЦАП є розповсюдженим і застосовується провідними виробниками засобів ПФІ.



Рис. 1 – Коригування вихідної величини ЦАП

АЦП з корекцією може працювати у двох режимах: калібрування, при якому відбувається визначення кодів відхилень та безпосередньо порозрядного перетворення з коректуванням значень компенсації неточності розрядів. Застосування такого підходу зазвичай дозволяє покращити один з параметрів. Наприклад підвищення точності АЦ перетворення призводить до погіршення інших параметрів, а саме, зменшується швидкодія процесу АЦ перетворення та зростає складність обладнання [2].

III. ІНФОРМАЦІЙНА НАДЛИШКОВІСТЬ

Введення інформаційної надлишковості у вигляді надлишкових систем числення дозволяє організувати процедуру самокалібрування і компенсації динамічних похибок значень розрядів АЦП і використовувати ці принципи для розробки багаторозрядних АЦП з високою точністю та швидкодією. Основною перевагою використання двійкової системи числення у ПФІ порівняно з НПСЧ та основою числення відмінною від двійкової є класична (не подовжена) розрядна мережа, відсутність необхідності збільшення кількості обладнання, а також відсутність необхідності перетворення отриманого коду в двійковий. Водночас, як одну з переваг використання НПСЧ з $\alpha < 2$ варто виділити можливість компенсації динамічних похибок I і II роду.

Недоліком двійкової системи числення порівняно з НПСЧ є відсутність можливості коригування в АЦП статичних похибок аналогових компонентів без затрат часу перетворення на обчислення і введення коригувальних виправлень у процесі основного перетворення. Такі процедури виконуються в режимі самокалібрування пристрою. Однак при цьому частину надлишковості НПСЧ необхідно затратити для забезпечення нерозривності

характеристики вхід-вихід перетворювача, побудованого на неточних аналогових компонентах.

Перспективним є напрямком використання систем числення, відмінних від двійкової. Так, запропоновано застосування трійкового числення [5], на базі якого розроблено адаптивний трійковий АЦП для використання в ширококутових системах зв'язку. В високошвидкісній системі зв'язку 1-трійковий тернарний АЦП має динамічну порогову адаптацію і забезпечує формування вихідного коду у трійковій симетричній формі $[+1, 0, -1]$. Трійковий АЦП включає пару 1-розрядних бінарних АЦП, один з яких налаштовується в неінвертованій формі, а інший - в інвертованій. Кожен бінарний АЦП охоплений зворотним зв'язком, що дозволяє одночасно і незалежно адаптувати пару порогів, компенсуючи ефекти зсуву постійного струму. В роботах[6,7] розроблено схема АЦП порозрядного зрівноваження починаючи від наймолодших ваг розрядів, що пришвидшує процес перетворення в вузьких під діапазонах. Введено додаткову ланку, що виконує порівняння різниці вхідної аналогової величини та сформованої суми вагових значень з одиничним еталонним значенням, що дозволяє завершити процес перетворення за меншу кількість кроків, що в свою чергу призведе до підвищення швидкодії.

ВИСНОВКИ

Проаналізовано підходи щодо покращення параметрів швидкодії та точності АЦ перетворення. Визначено, що покращення параметрів елементної бази, введення інформаційної та структурної надлишковості, а також комбінування структур дозволяють покращити параметри АЦП, проте кожен з методів володіє визначеними обмеженнями.

REFERENCES ЛІТЕРАТУРА

- [1] О. Д. Азаров, Аналого-цифрове порозрядне перетворення на основі надлишкових систем числення з ваговою надлишковістю. Вінниця, Україна: УНІВЕРСУМ, 2010.
- [2] А. П. Стахов, *Алгоритмическая теория измерения*.– Москва: Знание, 1979.
- [3] О. Д. Азаров, Методи та засоби високоточного слідкувального аналого-цифрового перетворення з ваговою надлишковістю. Вінниця, Україна: УНІВЕРСУМ, 2014
- [4] П. А. Арутюнов, *Теория и применение алгоритмических измерений*. Москва: Энергоатомиздат, 1990.
- [5] M. McLaughlin, M. McCullagh and C. McElroy, "Adaptive ternary A/D converter for use in an ultra-wideband communication system", *US 9325338 B2*, 26 Apr 2016
- [6] M. Petryshyn, "Modeling of the IFT processes in binary numeral systems based on the vector-branching diagrams", *IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON)*, 2017, p. 1077-1083.
- [7] М.Л. Петришин, "Моделювання процесів ПФІ в трійковій симетричній системі числення", *Системи обробки інформації*. Вип. 2, с. 60-65, 2017.
- [8] Побережний Л.Я. Ровінський В.А. Вимірювальний комплекс з АЦП на основі мікросхеми AD7714 Стаття. Збірник наукових праць II-ї Міжнародної науково-технічної конференції аспірантів та студентів. Донецьк, 2002 .
- [9] Васильків О.В. Заміховський Л.М., Шумада В.М., Ровінський В.А. Тензометричний давач для діагностування штангових глибинно-насосних установок Розвідка і розробка нафтових і газових родовищ. Серія: Технічна кібернетика та електрифікаційно-енерг. компл. - Івано-Франківськ, 2001. вип.37 (т.6).-С.201-208

Детектування Високочастотних Характеристик Цифрових Сигналів на Основі Трійкового Симетричного Вейвлет-Перетворення

Ігор Голуб

Прикарпатський національний університет
Івано-Франківськ, Україна
igorjokg@gmail.com

Артем Ізмайлов

кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет
Івано-Франківськ, Україна
aiartefact@gmail.com

Detecting of High Frequency Characteristics of Digital Signals Based on Symmetric Ternary Wavelet Transform

Ihor Holub

Prekarpathian National University
Ivano-Frankivsk, Ukraine
igorjokg@gmail.com

Artem Izmailov

dept. of Computer Science and Information Systems
Prekarpathian National University
Ivano-Frankivsk, Ukraine
aiartefact@gmail.com

Анотація—Неперервне трійкове симетричне вейвлет-перетворення застосовано для детектування високочастотних характеристик цифрових сигналів. Здійснено аналіз ефективності запропонованого застосування у порівнянні з існуючими вейвлет-перетвореннями.

Ключові слова—цифрова обробка інформації; вейвлет-перетворення; трійкове симетричне вейвлет-перетворення; високочастотні характеристики цифрових сигналів

Abstract—Continuous symmetric ternary wavelet transform was applied for detection of high-frequency characteristics of digital signals. Effectiveness of the introduced application was estimated in comparison with existing wavelet transforms.

Keywords—digital information processing; wavelet transform; symmetric ternary wavelet transform; high-frequency characteristics of digital signals

I. ВСТУП

Цифрова обробка інформації є ключовим елементом багатьох технічних систем, які використовуються у різних галузях економіки, управління, виробництва, зв'язку та медицини [1 –]. Ефективні рішення у галузі цифрової

обробки інформації підвищують ефективність перебігу процесів, які включають цифрову обробку інформації, у прикладних галузях.

Одним із актуальних завдань цифрової обробки інформації є детектування високочастотних характеристик цифрових сигналів [1 – 6]. Для розв'язання цього завдання застосовуються різні вейвлет-перетворення, причому, кожне з них, пристосоване для обробки лише певного класу сигналів, тобто, має обмежений спектр застосування [3 – 7]. Звідси випливає, що актуальною є проблема детектування високочастотних характеристик цифрових сигналів за допомогою нових синтезованих вейвлет-функцій та відповідних їм вейвлет-перетворень.

Аналіз останніх досліджень у галузі детектування високочастотних характеристик цифрових сигналів вказує на те, що дослідження щодо розв'язання даного завдання за допомогою трійкового симетричного вейвлет-перетворення (ТСВП) не проводились [2 – 7]. Водночас, застосування методів цифрової обробки інформації на основі трійкових симетричних функцій володіє рядом переваг [2, 3, 8], які вказують на перспективність розвідок у даному напрямі.

Метою дослідження є аналіз ефективності застосування синтезованого неперервного ТСВП для детектування високочастотних характеристик цифрових сигналів.

Наукова новизна отриманих результатів полягає в успішному доведенні ефективності застосування синтезованого неперервного ТСВП, у порівнянні з найбільш уживаними вейвлет-перетвореннями, для детектування високочастотних характеристик цифрових сигналів.

II. НЕПЕРЕРВНЕ ТРИКОВЕ СИМЕТРИЧНЕ ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ

В основі довільного вейвлет-перетворення лежить функція, яку називають материнським вейвлетом, а відповідне вейвлет-перетворення синтезується на основі системи функцій, які є стисненими та зсунутими по осі часу копіями материнського вейвлета [4 – 6].

Неперервним вейвлет-перетворенням функції $f(x) \in L_2(\mathbb{R})$ називають функцію двох змінних

$$W_f(a,b) = \langle f(\bullet) | \psi_{a,b}(\bullet) \rangle, a,b \in \mathbb{R}, a \neq 0, \quad (1)$$

де вейвлети $\psi_{a,b}(x)$ є масштабованими та зсунутими копіями материнського вейвлета $\psi(x) \in L_2(\mathbb{R})$

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right), a,b \in \mathbb{R}, a \neq 0. \quad (2)$$

Трійковий симетричний материнський вейвлет визначено у вигляді [3]

$$\psi_1(t) = \begin{cases} -\sqrt{\frac{3}{2}}, t \in [0, \frac{1}{3}), \\ \sqrt{\frac{3}{2}}, t \in [\frac{2}{3}, 1), \\ 0, t \notin [0, \frac{1}{3}) \cup [\frac{2}{3}, 1). \end{cases} \quad (3)$$

Неперервне ТСВП синтезовано підстановкою функції (3) у вираз (1), що у інтегральній формі записано наступним чином:

$$W_f(a,b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}(t) dt, a,b \in \mathbb{R}, a \neq 0, \quad (4)$$

де вейвлети $\psi_{a,b}(x)$ є масштабованими та зсунутими копіями материнського вейвлета $\psi_1(x)$

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}} \psi_1\left(\frac{x-b}{a}\right), a,b \in \mathbb{R}, a \neq 0. \quad (5)$$

Обернене до (4) перетворення визначено наступним чином:

$$f(x) \approx 0,10996 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a,b) \psi_{a,b}(x) \frac{da db}{a^2}. \quad (6)$$

Неперервне ТСВП у вигляді (4) використано для подальших досліджень ефективності застосування даного перетворення для детектування високочастотних характеристик цифрових сигналів.

III. ЕФЕКТИВНІСТЬ ЗАСТОСУВАННЯ ТРИКОВОГО СИМЕТРИЧНОГО ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ ДЛЯ ДЕТЕКТУВАННЯ ВИСОКОЧАСТОТНИХ ХАРАКТЕРИСТИК ЦИФРОВИХ СИГНАЛІВ

Для оцінки ефективності застосування ТСВП (4) застосовано порівняльний аналіз скейлограм (англ. scalogram) [4, 5], одержаних для множини заданих тестових сигналів на основі перетворення (4) (ST) та інших перетворень на основі вейвлетів Хаара (haar), Добеші 2-го порядку (db2), Добеші 3-го порядку (db3), Добеші 4-го порядку (db4), біортогонального з параметрами 1.3 (bior1.3), біортогонального з параметрами 2.2 (bior2.2), біортогонального з параметрами 3.7 (bior3.7).

Множина тестових сигналів поділена на дві групи: translated та superimposed. Сигнали з першої групи складаються з двох копій заданої хвилі, які є змаштабованими (мають різні частоти) та зміщеними по осі часу. Сигнали з другої групи складаються з двох копій заданої хвилі, які є змаштабованими, але незміщеними по часовій осі, тобто, відбувається накладання хвиль різної частоти.

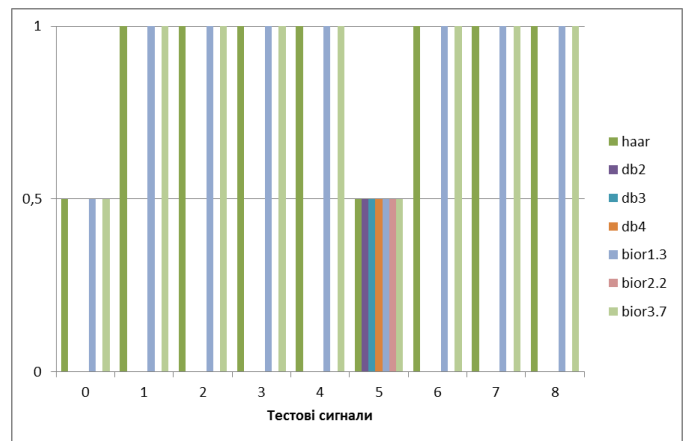


Рис. 1. Діаграма переваги неперервних вейвлет-перетворень над ТСВП для сигналів типу translated («1» – перевага заданого перетворення над ТСВП, «0» – протилежна ситуація, «0,5» – однаковий рівень ефективності)

У випадку тестових сигналів з групи translated (рис. 1), проаналізовані вейвлет-перетворення детектують обидві копії хвилі (з різними частотами) з однаковою ефективністю. Тобто, якщо неперервне вейвлет-перетворення детектує низькочастотну складову тестового

сигналу, то воно детектує і високочастотну складову з тою самою ефективністю. З діаграми переваги неперервного ТСВП над іншими перетвореннями у випадку тестових сигналів типу translated (рис. 1) випливає, що ТСВП (4) має перевагу у детектуванні ізольованих високочастотних характеристик над перетвореннями db2, db3, db4 та bior2.2 для всіх тестових сигналів, за винятком сигналу testsig5 translated. Водночас, перетворення haar, bior1.3 та bior3.7 володіють перевагою над ТСВП у випадку 78% тестових сигналів.

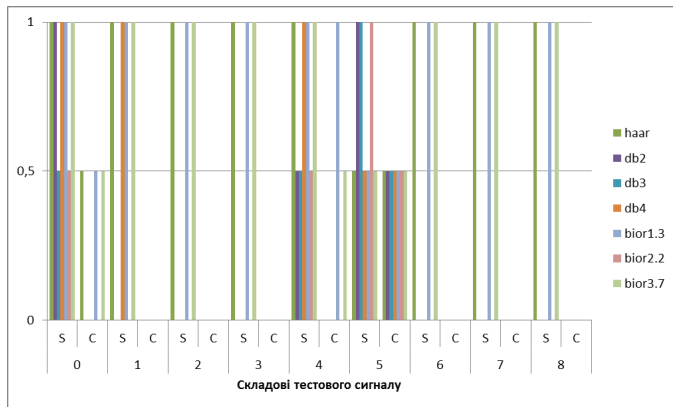


Рис. 2. Діаграма переваги неперервних вейвлет-перетворень над ТСВП для сигналів типу superimposed («S» – низькочастотна складова сигналу, «c» – високочастотна складова сигналу)

У випадку тестових сигналів групи superimposed (рис. 2), проаналізовані неперервні вейвлет-перетворення детектують обидві компоненти сигналу з різною ефективністю. У загальному випадку, показники ефективності детектування низькочастотних складових за допомогою ТСВП аналогічні показникам, отриманим для тестових сигналів типу translated. Водночас, ТСВП детектує високочастотні складові (які накладені на низькочастотні) з вищою ефективністю (з більшими на 10%-30% коефіцієнтами перетворення), ніж інші перетворення, у випадку всіх тестових сигналів, крім testsig5 superimposed.

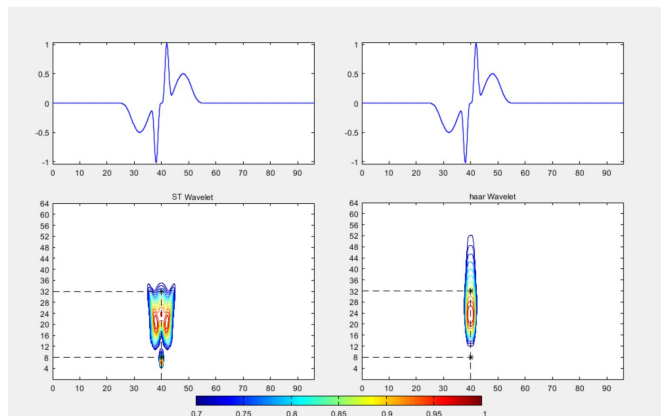


Рис. 3. Скейлограми ТСВП та неперервного вейвлет-перетворення Хаара для тестового сигналу testsig8 superimposed

Наприклад, у випадку тестового сигналу testsig8 superimposed ТСВП та неперервне вейвлет-перетворення Хаара детектують низькочастотну складову з однаковою ефективністю (рис. 3). Однак, у випадку перетворення Хаара, відсутні нормалізовані коефіцієнти перетворення більші, ніж 0,7, у області, близькій до точки детектування (ідеального співвідношення між масштабом та зміщенням вейвлета для детектування характеристики сигналу) високочастотної складової сигналу, яка позначена зірочкою на відповідній скейлограмі (рис. 3). Водночас, у випадку ТСВП, спостерігається локальний максимум значень коефіцієнтів перетворення у області, близькій до зазначеної точки (рис. 3). Відповідно, ТСВП чітко детектує та розділяє обидві накладені складові сигналу.

ВИСНОВКИ

Проведений аналіз ефективності показав, що ТСВП є добре пристосоване для детектування характеристик цифрових сигналів і, у загальному, забезпечує вищу ефективність застосування у порівнянні з іншими вейвлет-перетвореннями у 50% випадків. Максимальна перевага (до 30%) ТСВП досягається у випадку детектування високочастотних характеристик, які накладені на низькочастотні. Відповідно, ТСВП може бути ефективно застосоване у системах цифрової обробки інформації для моніторингу довготермінових процесів, оскільки, дозволить детектувати короткотермінові відхилення у їх перебігу.

Подальші дослідження полягають у вивченні властивостей та характеристик ТСВП, які дозволять детально визначити перспективу застосування даного перетворення.

ЛІТЕРАТУРА REFERENCES

- [1] E. Ifeachor, B. Jervis: Digital Signal Processing: A Practical Approach (2nd Edition), Prentice Hall, Harlow, 2001.
- [2] А.В. Измайлов. “Ефективність застосування ортогонального перетворення на основі трійкових симетричних функцій для цифрової обробки інформації.” Методи та прилади контролю якості. 2018. № 1 (40). С. 97-104.
- [3] А. В. Измайлов та Л. Б. Петришин, “Дискретне трійкове симетричне вейвлет-перетворення та його застосування для цифрової обробки інформації у розподілених системах управління,” Інформаційні технології та комп’ютерне моделювання: матеріали статей Міжнародної науково-практичної конференції, Івано-Франківськ, 2018.
- [4] P.S. Addison. The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance (Second Edition), CRC Press, 2016.
- [5] S. Mallat. A Wavelet Tour of Signal Processing (2nd edition), Academic Press, 1999.
- [6] I. Daubechies. Ten Lectures on Wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, Rutgers University and AT&T Bell Laboratories, 1992.
- [7] S. Prasad, W. Li, J.E. Fowler, L.M. Bruce. Information Fusion in the Redundant-Wavelet-Transform Domain for Noise-Robust Hyperspectral Classification, IEEE Transactions on Geoscience and Remote Sensing, Vol. 50, No. 9., September 2012, pp. 3474-3486. doi: 10.1109/TGRS.2012.2185053
- [8] B. Hayes. Computing science. Third base. A reprint from American Scientist, the magazine of Sigma Xi, the Scientific Research Society, Vol. 89, No. 6., November–December 2001, pp. 490-494.

Застосування Трійкового Симетричного Вейвлет-Перетворення для Підвищення Ефективності Керування Процесом Буріння Нафтових і Газових Свердловин

Артем Ізмайлов

*кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет
Івано-Франківськ, Україна
aiartefact@gmail.com*

Application of Symmetric Ternary Wavelet Transform to Increase Effectiveness of Controlling the Drilling Process for Oil and Gas Wells

Artem Izmailov

*dept. of Computer Science and Information Systems
Precarpathian National University
Ivano-Frankivsk, Ukraine
aiartefact@gmail.com*

Анотація—Розроблено пристрій реалізації трійкового симетричного вейвлет-перетворення. Оцінено ефективність застосування розробленого пристрою у якості компонента комплексу засобів наземного контролю і керування процесом буріння нафтових і газових свердловин.

Abstract—Device for performing symmetric ternary wavelet transform was synthesized. The introduced device was applied as the component of the complex for ground-based control and management of drilling oil and gas wells. Effectiveness of the described application was estimated.

Keywords—*digital information processing; discrete wavelet transform; symmetric ternary wavelet transform; controlling the drilling process for oil and gas wells*

I. ВСТУП

Одним із актуальних завдань цифрової обробки інформації (ЦОІ) є стиснення даних [1–3]. Одним з

шляхів розв'язання даного завдання є застосування вейвлет-перетворень (ВП), при якому зберігаються не самі дані, а коефіцієнти відповідного перетворення, причому, близькі до нуля коефіцієнти відкидаються [1–3]. Водночас, успішний синтез дискретного трійкового симетричного вейвлет-перетворення (ТСВП) та доведена ефективність його застосування у системах ЦОІ [2] вказують на необхідність застосування даного перетворення для розв'язання заданої задачі.

Метою дослідження є розробка апаратно-програмного засобу реалізації ТСВП та його застосування для стиснення даних у комплексі засобів наземного контролю і керування процесом буріння нафтових і газових свердловин.

Наукова новизна отриманих результатів полягає в успішній розробці апаратно-програмного засобу реалізації ТСВП та його застосуванні для підвищення ефективності функціонування комплексу засобів наземного контролю і керування процесом буріння нафтових і газових свердловин.

II. ОСНОВНА ЧАСТИНА

Детальну інформацію про ТСВП та показники його ефективності можна знайти у [2]. З метою спрощення апаратної реалізації пристрою ЦОІ на основі ТСВП, використано плату розробки Arduino Mega 2560 Rev3, яка містить мікроконтролер ATmega2560, у якості обчислювального елемента, а мікроконтролер ATmega16U2 використано для перетворення інтерфейсів зв'язку. Для забезпечення взаємодії розробленого пристрою з комп'ютером та управління перебігом процесу ЦОІ розроблено програмний засіб «Symmetric Ternary Wavelet Transform MC Performer», написаний мовою програмування C#.

Розроблений апаратно-програмний засіб застосовано у якості компонента комплексу засобів наземного контролю і керування процесом буріння нафтових і газових свердловин СКУБ-М2 [4] (рис. 1).

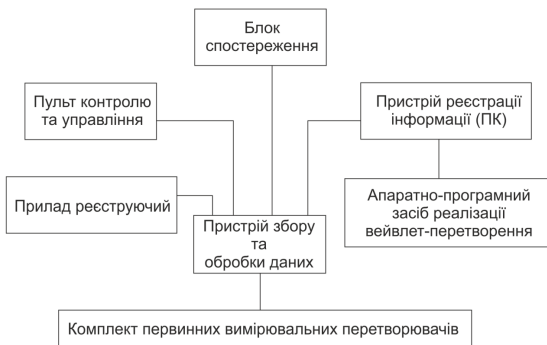


Рис. 1. Структурна схема комплексу СКУБ-М2 з включенням до його складу апаратно-програмним засобом реалізації ТСВП

Функції пристрою реєстрації інформації комплексу СКУБ-М2, які стосуються ведення архіву, пов'язані з даними, які, в більшості випадків, є надлишковими, оскільки, пов'язані з циклічними процесами у буровому обладнанні, і допускають похибку до 5% при стисненні [5], зокрема, шляхом відкидання близьких до нуля коефіцієнтів вейвлет-перетворення.

Для оцінювання ефективності розробленого засобу здійснено порівняння значень критерію середньої абсолютної похибки у відсотках $MAPE$ (mean absolute percentage error) відновлення даних (1) [3] за частиною коефіцієнтів перетворення, одержаних для ТСВП (ST), вейвлета Хаара (haar), вейвлета Добеші 4-го порядку (db4) та біортогонального вейвлета з параметрами 2.6 (bior2.6) при відновленні послідовностей значень технологічного параметра моменту механічного лівого

$$MAPE = \frac{1}{N} \sum_{i=0}^{N-1} \frac{|X(i) - X_r(i)|}{X(i)} \cdot 100\%, \quad (1)$$

де N – кількість відліків у векторі вхідних даних X , $X(i)$ – вхідні дані, $X_r(i)$ – відновлені дані після зменшення надлишковості.

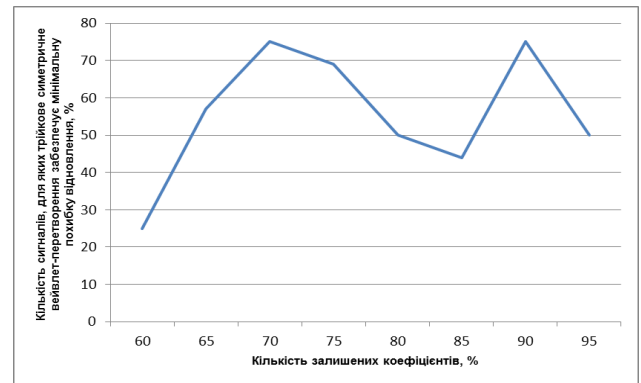


Рис. 2. Графік залежності кількості сигналів, для яких ТСВП забезпечує мінімальну похибку відновлення (1), від кількості залишених коефіцієнтів ВП

З графіка на рисунку 2 випливає, що мінімальна похибка відновлення (1) забезпечується ТСВП для 75% тестових послідовностей значень технологічного параметра моменту механічного лівого, у випадку відновлення за 70% та 90% коефіцієнтів перетворення. Максимальне зменшення надлишковості при допустимих значеннях похибки відновлення досягається усіма проаналізованими ВП при збереженні 70% коефіцієнтів перетворення. При цьому, застосування ТСВП забезпечило, в загальному випадку, наступний приріст ефективності за критерієм мінімальної похибки відновлення (1): у порівнянні з haar – 20%, db4 – 8%, bior2.6 – 18,5%.

ВИСНОВКИ

Таким чином, застосування розробленого апаратно-програмного комплексу реалізації ТСВП для зменшення надлишковості значень технологічного параметра моменту механічного лівого у комплексі засобів наземного контролю і керування процесом буріння нафтових і газових свердловин СКУБ-М2 дозволило зменшити похибку відновлення до 20%, у порівнянні з іншими ВП. Подальші дослідження полягають у вдосконаленні розробленого засобу реалізації, а також, у його застосуванні у якості компонента відмінних від проаналізованих, комплексів обробки цифрових даних.

ЛІТЕРАТУРА REFERENCES

- [1] E. Ifeachor, B. Jervis: Digital Signal Processing: A Practical Approach (2nd Edition), Prentice Hall, Harlow, 2001.
- [2] А.В. Измайлов та Л.Б. Петришин, Дискретне трійкове симетричне вейвлет-перетворення та його застосування для цифрової обробки інформації у розподілених системах управління, Інформаційні технології та комп'ютерне моделювання: матеріали статей Міжнародної науково-практичної конференції, м. Івано-Франківськ, Івано-Франківськ: Супрун В.П., 2018, С. 152-155.
- [3] Д. Сэлмон, Сжатие данных, изображений и звука, пер. с англ. В.В. Чепыжова. М.: Техносфера, 2004.
- [4] "Івано-Франківське СКБ ЗА. Каталог продукції. СКУБ-М2.," Івано-Франківське СКБ ЗА, 2018. [Online]. Available: http://skbza.if.ua/cat_5ua.htm. [Accessed 11 October 2018].
- [5] Л. М. Заміховський, В. А. Ровінський та О. В. Євчук, Діагностика технічного стану штангових глибинно-насосних установок, Івано-Франківськ: Симфонія форте, 2006.

Застосування Трійкового Симетричного Ортогонального Перетворення для Підвищення Ефективності Функціонування Пристроїв Обліку та Контролю Параметрів Передачі Газу

Артем Ізмайлов
кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет
Івано-Франківськ, Україна
aiartefact@gmail.com

Application of Symmetric Ternary Orthogonal Transform to Increase Operation Effectiveness of the Devices for Gas Accounting and Transmission Control

Artem Izmailov
dept. of Computer Science and Information Systems
Precarpathian National University
Ivano-Frankivsk, Ukraine
aiartefact@gmail.com

Анотація—Розроблено пристрій реалізації трійкового симетричного ортогонального перетворення. Оцінено ефективність застосування розробленого пристрою у якості компоненту пристроїв обліку та контролю параметрів передачі газу.

Abstract—Device for performing symmetric ternary orthogonal transform was synthesized. The introduced device was applied as the component of the devices for gas accounting and transmission control. Effectiveness of the described application was estimated.

Ключові слова—цифрова обробка інформації; дискретне ортогональне перетворення; трійкове симетричне ортогональне перетворення; пристрій обліку газу

Keywords—digital information processing; discrete orthogonal transform; symmetric ternary orthogonal transform; gas accounting devices

I. ВСТУП

Одним із актуальних завдань цифрової обробки інформації (ЦОІ) є стиснення даних [1–4]. Одним з шляхів розв'язання даного завдання є зниження рівня кореляції вхідних даних за допомогою ортогональних перетворень (ОП) [1–4]. При цьому, зберігаються не самі дані, а коефіцієнти відповідного перетворення. Водночас, успішний синтез дискретного трійкового симетричного ортогонального перетворення (ТСОП) та доведена ефективність його застосування у системах ЦОІ [2] вказують на необхідність застосування даного перетворення для розв'язання заданої задачі.

Метою дослідження є розробка та моделювання пристрою реалізації ТСОП та його застосування для стиснення даних у пристроях обліку та контролю параметрів передачі газу.

Наукова новизна отриманих результатів полягає в успішній розробці та моделюванні пристрою реалізації

ТСОП та його застосуванні для підвищення ефективності функціонування пристроїв обліку та контролю параметрів передачі газу.

II. ОСНОВНА ЧАСТИНА

Детальну інформацію про ТСОП та показники його ефективності за різними критеріями можна знайти у [2]. Структурну схему розробленого пристрою реалізації ТСОП наведено на рисунку 1.

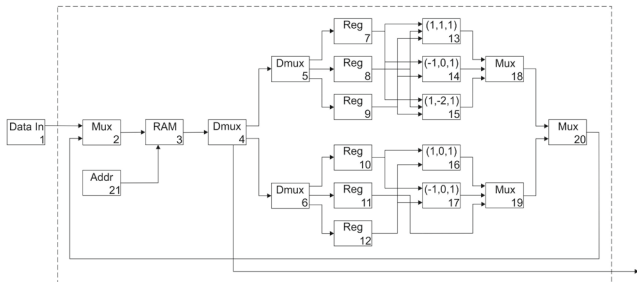


Рис. 1. Структурна схема пристрою реалізації ТСОП

У якості компонентів розробленого пристрою використано оперативний запам'ятовуючий пристрій (3), регістри пам'яті (7-12), мультиплексори (2, 18-20), демультимплексори (4-6), генератор адрес (21) та операційні блоки: $f(x_0, x_1, x_2) = (x_0 + x_1) + x_2$ (13), $f(x_0, x_1) = (-x_0 + x_1)$ (14, 17), $f(x_0, x_1, x_2) = (x_0 + x_1) - 2x_2$ (15), $f(x_0, x_1) = (x_0 + x_1)$ (16).

Ефективність застосування розробленого пристрою оцінена для пристроїв комплектації вузлів обліку та контролю параметрів передачі газу. Зокрема, оцінено ефективність застосування розробленого пристрою реалізації перетворення у якості компонента наступних пристроїв: ОЕ-22ЛА (обчислювач об'єму газу), ОЕ-RW (пристрій переносу інформації), ТЕМР-ОЕ (лічильник газовий роторний з електронним відліковим пристроєм). Повні назви та призначення цих пристроїв наведено у [5].

Перелічені пристрої дозволяють архівувати значення показників, які отримані з їх допомогою (ОЕ-22ЛА, ТЕМР-ОЕ), та передавати ці значення на вузли контролю, які реалізовані на основі комп'ютера (ОЕ-RW) [5]. Застосування пристрою реалізації ТСОП дозволило знизити значення ступеня кореляції архівованих даних і, відповідно, збільшити ступінь їх стиснення. У результаті, збільшено кількість даних, які можуть зберігатися у пам'яті пристрою обліку без необхідності введення додаткових запам'ятовуючих елементів. Результати оцінки приросту кількості даних для зберігання у пристроях обліку та контролю параметрів передачі газу наведено у таблиці 1.

Приріст ефективності зберігання даних у таблиці 1 обчислено за наступною формулою

$$\left(\frac{T}{S} - I\right) \cdot 100\% \quad (1)$$

ТАБЛИЦЯ 1. ПРИРІСТ ЕФЕКТИВНОСТІ ЗБЕРІГАННЯ ДАНИХ ПРИСТРОЯМИ ОБЛІКУ ТА КОНТРОЛЮ ПАРАМЕТРІВ ПЕРЕДАЧІ ГАЗУ (S – КІЛЬКІСТЬ ЗАПИСІВ У АРХІВІ БЕЗ ЗАСТОСУВАННЯ ТСОП, T – КІЛЬКІСТЬ ЗАПИСІВ У АРХІВІ ІЗ ЗАСТОСУВАННЯМ ТСОП)

№	Шифр пристрою	Тип запису	S	T	Приріст ефективності зберігання даних
1	ОЕ-22ЛА	Добовий	64	77	20,3125
2	ОЕ-RW	–	8	9	12,5
3	ТЕМР-ОЕ	Добовий	126	153	21,4286
4		Зміна стану лічильника	540	657	21,6667
5		Зміна у конфігурації лічильника	128	155	21,0938

Результати проведеного аналізу ефективності (табл. 1) застосування розробленого пристрою реалізації ТСОП у якості компонента пристроїв обліку та контролю параметрів передачі газу, вказали на підвищення ефективності зберігання даних в середньому на 20%. У випадку пристроїв, для яких здійснювався аналіз, такий приріст зумовив відповідне збільшення часу автономного функціонування та підвищення точності обліку, у зв'язку з можливістю збереження більшої кількості відліків змін конфігурації обладнання розширило можливості моніторингу та діагностики обладнання з метою виявлення відхилень у перебігу процесів обліку, що зумовило підвищення надійності функціонування вузлів обліку та контролю параметрів в цілому.

ВИСНОВКИ

Одержаними результатами аналізу ефективності (табл. 1) застосування розробленого пристрою реалізації ТСОП доведена ефективність застосування запропонованого методу ЦОІ у системах обробки та передачі даних, у яких пріоритетним напрямом оптимізації є стиснення даних при передачі та/або достовірність передачі даних.

Подальші дослідження полягають у вдосконаленні структури розробленого ТСОП та відповідного пристрою перетворення, а також, у застосуванні розробленого пристрою у якості компонента відмінних від проаналізованих, пристроїв обробки цифрових даних.

ЛІТЕРАТУРА REFERENCES

- [1] E. Ifeachor, B. Jervis: Digital Signal Processing: A Practical Approach (2nd Edition), Prentice Hall, Harlow, 2001.
- [2] А.В. Измайлов. "Ефективність застосування ортогонального перетворення на основі трійкових симетричних функцій для цифрової обробки інформації," Методи та прилади контролю якості. 2018. № 1 (40). С. 97-104.
- [3] Г. Н. Мальцев, А. В. Харченко и А. С. Гарагуля, "Границы повышения помехоустойчивости радиотехнических систем передачи информации при цифровой обработке сигналов с компенсацией помех," Информационно-управляющие системы, №4, С. 111-116, 2014.
- [4] Д. Сэлмон, Сжатие данных, изображений и звука, пер. с англ. В.В. Чепыжова. М.: Техносфера, 2004.
- [5] "ТОВ СЛОТ – Каталог продукції," ТОВ СЛОТ, 2018. [Online]. Available: <http://www.slot.if.ua/catalog/>. [Accessed 11 October 2018].

A Generalized Methodology for the Construction of Predictive Analysis Systems as Exemplified by the Mining Equipment in the BIG DATA Environment

Andrey Kupin

*Department of Computer Systems and Networks
Kryvyi Rih National University
Vitaliya Matushevycha str. 11, Kryvyi Rih, Ukraine
kupin.andrew@gmail.com*

Rodion Ivchenko

*Department of Automation Computer Science and Technologies
Kryvyi Rih National University
Vitaliya Matushevycha str. 11, Kryvyi Rih, Ukraine
ivchenko.ra@gmail.com*

Узагальнена Методологія Побудови Систем Предиктивного Аналізу на Прикладі Гірського Устаткування в Умовах BIG DATA

Андрій Купін

*Кафедра комп'ютерних систем та мереж
Криворізький національний університет
вул. Віталія Матусевича, 11, м. Кривий Ріг, Україна
kupin.andrew@gmail.com*

Родіон Івченко

*Кафедра автоматизації, комп'ютерних наук і технологій
Криворізький національний університет
вул. Віталія Матусевича, 11, м. Кривий Ріг, Україна
ivchenko.ra@gmail.com*

Abstract – It is necessary to determine the optimal methodology for the system of predictive analysis of equipment to prevent emergency situations. The system may include, in particular: data input/reading from sensors, processing/storage of information in a database using algorithms for processing Big Data and decision trees [1]. Identifying possible types of problems and making decisions on how to respond to them; training the system for more accurate response and decision-making.

Keywords—Industry 4.0; Big Data; Predictive analytics.

I. INTRODUCTION

The issue of introducing smart management of technical maintenance services at large industrial enterprises on the basis of new approaches within the framework of the modern concept of Industry 4.0 is being considered.

The current trends in the scientific and technological progress of the global industry are quite often described in such terms as “Smart Plant”, “Smart Production”, “Smart Factory” and “Factory of the Future”. At present, the development of those research areas is quite well illustrated by the concept of the 4th industrial revolution (Industry 4.0). The implementation of that concept involves the use of some key technological

trends, such as Big Data processing, cyber-physical systems, autonomous robots with various smart sensors, simulators for 2D and/or 3D modeling, 3D printers, Internet things, augmented reality, etc. Consequently, according to the estimates of the world's leading experts, those trends will determine the main vector of the modern competitive industries [1].

The predictive analytics is a class of data analysis methods that focuses on predicting the future behavior of objects and subjects in order to make optimal decisions.

The predictive analytics uses statistical methods, smart data analysis methods, analyzes current and historical facts to compile predictions about future events. In business, predictive models use patterns found in historical data and the data being acted upon in order to identify risks and opportunities. Models capture relationships among many factors to make possible assessing risks or the potential associated with a particular set of conditions thereby guiding decision making.

II. PROBLEM STATEMENT

Another task for the predictive algorithms is equipment maintenance and repair. Companies mainly use basic control mechanisms provided by equipment manufacturers. But the

potential of those means is limited, because they do not make it possible to analyze additional factors influencing the state of the equipment, and to predict a critical situation in advance. So, the maintenance department employees receive a lot of data, but they do not know how the items of that data are related to each other. As a result, the response from the repair services follows only after the actual equipment failure, which leads to downtime and, consequently, additional costs. The predictive analytics conducts, by means of machine learning and artificial intelligence, continuous analysis of Big Data, performs data visualization on the current state of the equipment, and predicts scenarios for the occurrence of equipment failures. As a result, unplanned downtime is reduced, equipment maintenance and repair work are optimized, maintenance time is decreased, and the management personnel receive an in-depth analysis of the causes of equipment failures.

Modern IoT and Big Data capabilities, along with the advanced predictive analytics methods, are becoming an efficient tool to reduce costs, improve product quality and increase company productivity. The predictive analytics has become a new trend of modern times that opens up broad prospects for the further development of companies. In addition to manufacturing industry, forecast platforms are successfully used in banking, insurance, retail, logistics, marketing and many other areas. Thus, banks and other financial institutions use the predictive analytics tools to forecast their customers' behavior, e. g. to find out whether they will acquire a new banking product or whether they will continue to cooperate with the bank. Exploring their historical behavior in conjunction with external factors, mathematical algorithms can help make profitable personal offers, anticipate sales increase, customer churn, etc. In marketing, mathematical modeling makes it possible to segment customers by certain characteristics, predict their future behavior (indicate customer value, calculate the likelihood of another purchase, etc.) and determine services and products that may be of interest to the customers. Predictive models accurately determine the efficient channels of interaction with the target groups. Analyzing social media data helps marketing researchers better understand customers and potential buyers. In retail business and logistics, predictive models help forecast demand, optimize inventories and analyze the risks of delays in deliveries [2].

When embarking on the introduction of a predictive analytics system, it is necessary to keep in mind that such a system cannot work without a large amount of the historical data and continuous collection of the current data. The less data is involved, the less accurate the predictions will be. When promoting predictive analysis tools, companies often face the problem of data shortage, which becomes an insurmountable barrier to forecasting.

It is therefore necessary to start collecting data immediately. Otherwise the company risks not sustaining rivalry with more technologically equipped competitors. Organizing stable high-quality data collection and storage is the primary task that must be accomplished in the near future if the company intends to remain on the market.

III. DESCRIPTION OF CONTROL OBJECTS AND THEIR MATHEMATICAL MODELS

Usually, concentrating plants are classified according to the nature of the concentrating processes used (flotation, gravitational, magnetic ones) or according to the raw material being processed (coal, ore, etc.). For any type of concentrating plants, the concentrating process technology is determined by a large and diverse number of factors influencing the said process.

The complexity of analyzing the mineral dressing technology as an object of control makes it natural and necessary to observe for its study a certain hierarchical structure. The distinctive feature of such a structure is the sequential division of the technical system into subsystems, between which the relationships of subordination are established. The hierarchical structure of ore concentrating plants from the standpoint of control tasks can be represented in the form of three stages [3]:

The lower stage is the so-called typical processes of the mineral dressing technology (crushing, grinding, separation, etc.);

The middle stage is a group of processes or pieces of equipment that perform an independent technological process task for the production of a given product, e. g. the flotation area of an ore concentrating plant; grinding, classification and magnetic separation stages of magnetic concentration plants; ore concentrate drying area, etc.;

The upper stage is a technological line for dressing a particular mineral raw material considered as a whole.

Each hierarchy of mineral dressing technology is characterized by its own control tasks that involve their respective functions, which actually determines the type of a mathematical model describing the process of functioning of a given object.

In general, the mathematical model of each stage of the hierarchy can be viewed as a complex object of the mineral dressing technology and presented as a function of variables. Here, three types of actions serve as the input of the object:

1) uncontrolled (but monitored) input variables $Y = \{y_1, \dots, y_r\}$ constitute a disturbance vector and, as a rule, characterize, as far as concentrating production is concerned, quality indicators of the source material to be processed and those of its intermediate products obtained during the concentrating process;

2) controlled input variables $U = \{u_1, \dots, u_n\}$ constitute a control vector and characterize, as a rule, quantitative indicators (expenditure) of material and energy flows;

3) the uncontrolled factors $Z = \{z_1, \dots, z_k\}$ constitute an interference vector. Basically, this is a disturbance vector, about which the developer of the control system knows very little or nothing at all. Most often that vector is not taken into account at all.

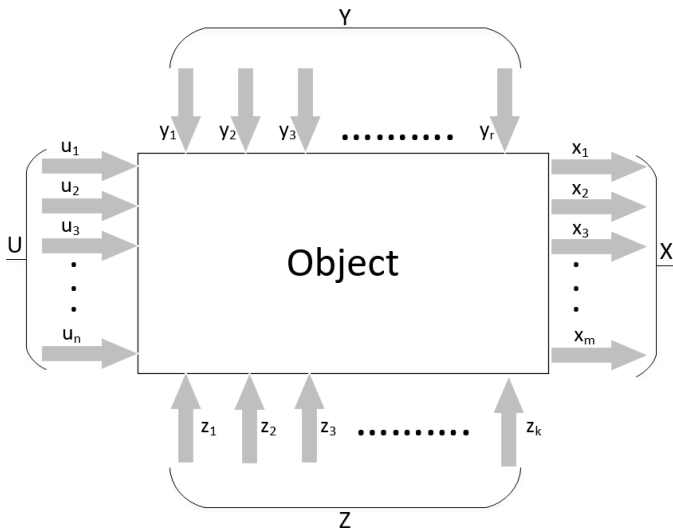


Fig. 1 – The structure of a complex control object

The essence of any technological process, including the process of concentrating mineral raw materials, is in converting the Y , U , Z input actions into the $X = \{x_1, \dots, x_m\}$ output ones. The X vector is called the state vector, and characterizes, as far as concentrating production is concerned, qualitative and quantitative features of the output products (concentrate, tailings).

The objects of the concentrating process can be conveniently represented in the form of models describing typical processing operations, as far as the distribution of flows of mineral raw materials or intermediate products is concerned, resulting in a quantitative and qualitative change in the parameters of those flows. In total, one can identify four typical operations that characterize the mineral dressing technology: simple, mixing, separating, and combination ones.

The mixing operation is characterized by the presence of the "n" process flows $Y = \{y_1, \dots, y_n\}$ at the input of the object and by one "x" output flow. The operation equation looks like this:

$$x = f(Y, U) \quad (1)$$

A mixing operation can be exemplified by a ball mill used before the flotation operation. At its entrance, the mill is fed by the flows of ore Q_r (t/h), water W_v (m^3/h), and reactants W_r , i. e. $Y = \{Q_r, W_v, W_r\}$, while at its exit, a flow "x" of the slurry from finely ground particles of ore is formed. The control actions can be represented by the consumption Q_b (t/h) of balls for the mill, ore consumption (Q_r (t/h), and water consumption W_v (m^3/h), i. e. $U = \{Q_r, W_v, W_r\}$.

IV. SENSOR TYPES

A sensor is an item of a measuring, signaling, regulating or controlling device that converts a monitored parameter (temperature, pressure, frequency, luminous intensity, voltage, current, etc.) into a signal convenient for measuring, transmitting, storing, processing, recording, and sometimes for influencing processes being controlled. Or more simply put, a

sensor is a device that converts the input action of any physical parameter into a signal convenient for further use.

The sensors used are very diverse, and can be classified according to various criteria [4]:

Depending on the type of the input (measured) parameter, the following sensors are identified: those of mechanical displacements (linear and angular); pneumatic and electric ones; flow meters; speed sensors; those of acceleration, force, temperature, pressure, etc.

Currently, there is approximately the following distribution of the shares of measurements of various physical parameters in the industry: temperature – 50%, consumption (by mass and by volume) – 15%, pressure – 10%, level – 5%, quantity (mass, volume) – 5%, time – 4%, electrical and magnetic parameters – less than 4%.

According to the type of the output, into which the input is converted, non-electrical and electrical sensors are distinguished: direct current sensors (respective EMF or voltage sensors), alternating current amplitude sensors (respective EMF or voltage sensors), alternating current frequency sensors (respective EMF or voltage sensors), resistance sensors (effective resistance sensors, inductance sensors or capacitance sensors), etc.

Most sensors are electric. This is due to the following advantages of electrical measurements:

- electrical parameters are convenient to transmit over a distance, with the transmission being carried out at a high speed;
- electrical parameters are universal in the sense that any other parameters can be converted into electrical ones and vice versa;
- electrical parameters can be accurately converted into a digital code, and make it possible to attain high accuracy, sensitivity and operation speed of measuring instruments

By the operating principle, the sensors can be divided into two classes: generator and parametric ones (modulator sensors). The generator sensors (transducers) directly convert the input parameter into an electrical signal.

Parametric sensors convert an input parameter into the change in some electrical parameter (R , L , or C) of the sensor.

According to the operating principle, the sensors can also be divided into ohmic, rheostat, photoelectric (optical-electronic), inductive, capacitive ones, etc.

Three classes of sensors are identified:

- analog sensors, i. e. sensors that produce an analog signal that is proportional to the change in the input parameter;
- digital sensors generating a sequence of pulses or a binary word;
- binary sensors that produce a signal of only two levels: "on/off" (in other words, "0" or "1"); they are widespread due to their simplicity

Requirements for sensors:

- non-ambiguous one-to-one relationship between the output parameter and the input one;
- performance stability over time;
- high sensitivity;
- small size and weight;
- no counter-effects on the process and the parameter being controlled;
- operation under various conditions;
- various mounting options

Parametric sensors (modulator sensors) convert the X input parameter into a change in some electrical parameter (R, L or C) of the sensor. It is impossible to transmit over a distance a change in the above sensor parameters without an electric signal (voltage or current). It is possible to identify the change in the corresponding sensor parameter only by the sensor's response to current or voltage, since it is the above parameters that characterize that response. Therefore, parametric sensors require the use of special measuring circuits supplied by direct or alternating current.

Ohmic (resistance) sensors – their operating principle is based on the change in their effective resistance when the length L , the sectional area S or the resistivity ρ change:

$$R = \rho l / S \quad (2)$$

In addition, the relationship between the effective resistance value and the contact pressure and the illumination of the photocells is used. In accordance with that, ohmic sensors are divided into contact, potentiometric (rheostat), strain gauge, thermistor, and photo-resistor ones.

Contact sensors are the simplest type of resistor sensors that convert the movement of a primary element into an abrupt change in the resistance of an electrical circuit. With the help of contact sensors, forces, displacements, temperature, dimensions and shapes of objects, etc. are measured and controlled. The contact sensors include position and limit switches, contact thermometers and so-called electrode sensors that are mainly used to measure the critical levels of electrically conducting fluids.

Contact sensors can operate using both direct and alternating current. Depending on the measurement range, contact sensors can be single-range and multi-range ones. The latter are used to measure values that vary significantly, with the parts of resistor R incorporated into the electrical circuit being successively short-circuited.

The disadvantage of contact sensors is the complexity of continuous control and a limited service life of the contact system. But owing to the extreme simplicity of those sensors, they are widely used in automation systems.

Rheostat sensors are actually resistors with varying effective resistance. The input parameter of the sensor is the displacement of the contact, while the output one is the change in its resistance value. The moving contact is mechanically

connected with the object, whose displacement (angular or linear) needs to be converted.

The most widespread is the potentiometric circuit for operating a rheostat sensor, into which the rheostat is incorporated according to the voltage divider circuit. It shall be reminded that the voltage divider is an electrical device for dividing DC or AC voltage into parts; the voltage divider makes it possible to take up (to use) only a portion of the available voltage through the items of an electrical circuit consisting of resistors, capacitors or inductors. A variable resistor included in the voltage divider circuit is called a potentiometer.

Usually, rheostat sensors are used in mechanical measuring devices to convert their readings into electrical parameters (current or voltage), e. g. in float level meters for liquids, in various pressure gauges, etc.

A sensor in the form of a simple rheostat is almost not used due to the considerable nonlinearity of its static characteristic line $I_n = f(x)$, where I_n is the current in the load.

The output parameter of such a sensor is the voltage drop U_{out} between the movable contact and one of the fixed ones. The relationship between the output voltage $U_{out} = f(x)$ and the x displacement of the contact corresponds to the regularity of the electrical resistance change along the potentiometer. The regularity of the resistance distribution over the length of the potentiometer that is determined by the design thereof can be either linear or nonlinear.

Potentiometric sensors, that are constructively variable resistors, are made of various materials, viz. winding wire, metal films, semiconductors, etc.

Strain gauge resistors (strain gauge transducers) are used to measure mechanical stresses, small deformations, and vibration. The operating principle of strain gauges is based on the strain-resistive effect, which is actually the change in the effective resistance of conductor and semiconductor materials under the influence of forces applied to them.

Thermometric sensors (thermistors) – their resistance depends on temperature. Thermistors, as sensors, are used in two ways:

1) The temperature of the thermal resistor is determined by the environment; the current passing through the thermal resistor is so small that it does not cause heating-up thereof. In view of that, the thermal resistor is used as a temperature sensor and is often called a "resistance thermometer".

2) The temperature of the thermal resistor is determined by the degree of heating by constant-value current and cooling conditions. In this case, the steady-state temperature is determined by the heat transfer conditions of the thermal resistor surface (the speed of the medium – gas or liquid – relative to the thermal resistor, the density of the medium, its viscosity and temperature); the thermal resistor can therefore be used as a flow rate sensor, thermal conductivity sensor of the medium, gas density sensor, etc. In sensors of such a type, something like a two-step conversion occurs: the parameter being measured is first converted into a change in the

temperature of the thermal resistor, which is then converted into a change in the electrical resistance.

Thermal resistors are made both from pure metals and semiconductors. The material, from which such sensors are made, should have a high temperature resistance coefficient, a linear, if possible, relationship between the electrical resistance and the temperature, good reproducibility of properties and inertness to environmental influences. Platinum satisfies all the above properties to the highest degree, copper and nickel do that to a slightly lesser degree.

Semiconductor thermal resistors (thermistors) have a higher sensitivity as compared to metal ones.

Inductive sensors are used to receive information about the movements of the working parts of machines, mechanisms, robots, etc. in a contactless manner and to convert that information into an electrical signal.

The operating principle of an inductive sensor is based on a change in the magnetic core winding inductance relative to the position of the individual elements of the magnetic circuit (armature, core, etc.). In such sensors, linear or angular displacement X (input parameter) is converted into a change in inductance (L) of the sensor. Inductive sensors are used to measure angular and linear displacements, deformations, for dimensional control, etc.

In the simplest case, the inductive sensor is an inductance coil with a magnetic core, the movable element of which (armature) is displaced under the action of the parameter being measured.

An inductive sensor recognizes all electrically conductive objects and responds accordingly. An inductive sensor is contactless, does not require mechanical action, and operates in a contactless manner by utilizing changes in the electromagnetic field.

Advantages:

- no mechanical wear and tear, no failures related to the state of the mechanical contacts
- no contact rattling sound and false response
- high switching frequency of up to 3000 Hz
- resistant to mechanical effects

The disadvantages of inductive sensors are the relatively low sensitivity, the relationship between the induction and the frequency of the supply voltage, a significant reverse effect of the sensor on the parameter being measured (due to the attraction of the armature to the core).

Capacitive sensors – their operating is based on the relationship between the capacitor capacitance and the size and the relative position of its plates, as well as the permittivity of the medium between the plates.

For a double-plate flat capacitor, the capacitance is determined by the following formula:

$$C = \epsilon_0 \epsilon S/h \quad (3)$$

where ϵ_0 – the dielectric constant; ϵ – the relative permittivity of the medium between the plates; S – the effective area of the plates; h – the distance between the plates of the capacitor.

The $C(S)$ and $C(h)$ relationships are used to convert mechanical displacements into changes in capacitance.

Capacitive sensors, as well as inductive ones, are supplied with alternating voltage (usually of heightened frequency – up to dozens of megahertz). As measuring circuits, bridge circuits and resonant ones are usually used. In the latter case, the relationship between the generator oscillation frequency and the capacitance of the resonant circuit is as a rule used, i. e. the sensor has a frequency output.

The advantages of capacitive sensors are their simplicity, high sensitivity and low response time. Their disadvantages are their propensity to the influence of external electric fields and the relative complexity of measuring devices, into which they are incorporated.

Capacitive sensors are used to measure angular displacements, very small linear ones, vibrations, speeds of movement, etc., as well as to reproduce given functions (harmonic, saw-tooth, rectangular ones, etc.).

Capacitive transducers, whose permittivity ϵ changes relative to a displacement, deformation or changes in the composition of the dielectric medium, are used as level sensors of non-conducting fluids, bulk and powdery materials, thickness sensors of the layer of non-conducting materials (thickness gauges), as well as control devices for the humidity and composition of a substance.

V. SUBSEQUENT INFORMATION PROCESSING

Interfaces control the data flow. A clear definition of the required data and its connection to the network data world are the basis for reliable data exchange over the network. That being the case, an important role is played by the choice of the correct data transfer protocols for a given section of the path. The Ethernet-based solutions are at the forefront. But IO-Link also makes connection to the network possible, especially for devices that need only a reduced data transfer capability.

In smart manufacturing, many sensors collect a large amount of data in many places. As a result, the importance of decentralized data processing is rising. Additional interfaces in the data or software systems make possible new analyses and functions, help improve the flexibility, quality, efficiency and transparency of the production.

Through the application field-oriented connection technologies and bypassing the control system, the sensor data can in the future be directed immediately into the cloud.

Thanks to the successful and comprehensive integration of all sensors into a network with centralized or decentralized data processing systems, a previously unknown number of solutions appears, while the whole process becomes transparent due to data transfer and communication protocols.

BIG DATA methods make it possible to handle structured and unstructured data of very large volumes to obtain results that are efficient under the conditions of continuous growth and

distributed over the nodes of the computing network. The need for such methods is caused by the ever-increasing development of technological processes and the equipment itself at the companies. Those methods can be used to collect information from sensors for the predictive analyses and data processing. They are also used to improve security and modularity (e. g. for the purpose of preventing equipment breakdowns).

The known technologies include [5]:

- 1) Recognition of graphic elements as a new part of speech recognition implementation.
- 2) Adaptation operations and related automated vehicles.
- 3) Semi-automatic flexible machines for additional services.
- 4) Fully automated quality assurance for adapting to rapid changes in demand.
- 5) Smart automatic control of objects for greater efficiency.
- 6) Enhancing security and modularity.

Enhancing the safety of production is of prime importance. The safety improvement is therefore possible mainly due to the predictive analysis to prevent emergency situations.

In recent years, SAP Predictive Analytics [6] has focused on the development of machine learning, Big Data processing and the IoT development. Those are the three most important technological areas, where the company is developing its solutions. SAP works not only on the development of a tool, but also on the application of the respective technologies in practice. If you have a large number of customers, automating your business processes with the use of SAP products allows you to analyze customer needs in a comprehensive way and offers you new approaches in utilizing client data to increase the efficiency of those processes.

You can download a temporary series into a data analysis tool, but without data pre-processing, the resulting model will be of poor quality. When preparing data, it is necessary to fulfill two stages of processing. The first stage is Data Engineering, i. e. collecting, understanding, clearing and primary data processing. The second stage is Feature Engineering: the formation of descriptive data features that contain information on various aspects of the behavior of the object, whose model is under construction. In terms of the CRISP-DM [2] methodology, those stages are similar to Data Understanding and Data Preparation (cf. Fig. 2).

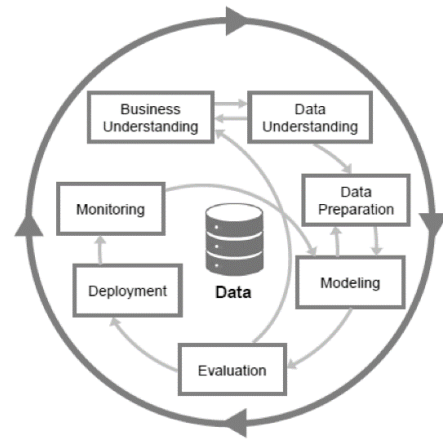


Fig. 2 – Stages in the CRISP-DM methodology with the possible transition paths between the stages [7]

An integral part of the initial stages of the machine learning process is the feature selection, i. e. the selection of variables on the basis of which the model undergoes the process of learning. The selection can be done using various tools, and also be dependent on many factors, e. g. the correlation of features with the target variable or the data quality. The next (and more important) step may be the creation of new features based on the already existing ones, the so-called feature engineering. That operation can allow engineers to improve the quality of the model, while at the same time making it possible to obtain a more complete explanation of the data, if the model is interpreted. In our case, the first step in constructing a SAP Predictive Analytics model was to create new features using the built-in Data Manager solution.

In the data set prepared, there are indicators that affect the target variable at the given moment of time. However, you can get additional information if you determine the impact of those indicators during a certain period up to the given moment. In our case, the following time intervals were chosen: during 1 hour and 1 day until the given moment of time. Even more informative may be the degree of the indicator change from the moment in the past until the given moment. As part of the method, the natural logarithm of the quotient of the current indicators and those with an interval of 1 hour and 2 days was selected. Consequently, we managed to obtain the degree of the indicator change from the moment in the past (whether it had increased or decreased and if so, then to what extent).

All modern techniques of working with Big Data invariably follow the three principles given below. In order to comply with them, it is necessary to use methods, techniques and paradigms of data processing. One of the most established methods is called MapReduce.

MapReduce is a distributed data processing model offered by Google for processing large volumes of data in computer clusters. MapReduce:

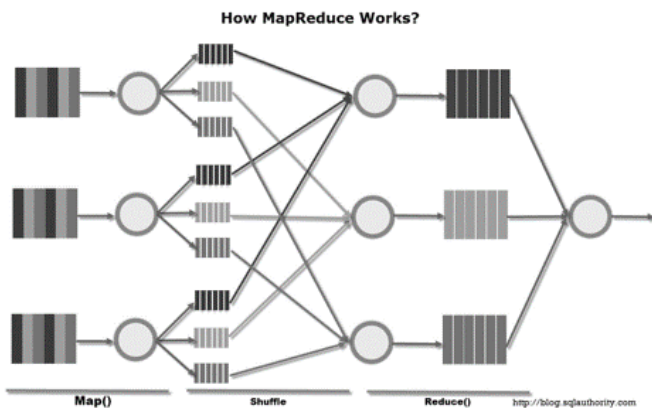


Fig. 3 The MapReduce model [8]

MapReduce assumes that the data is organized in the form of some entries. Data processing takes place in 3 stages:

1. The "Map" stage. At this stage, the data is converted using the "map ()" function, which is defined by the user. The role of this stage lies in the preliminary processing and filtration of the data. This operation is very similar to the "map" operation in functional programming languages. This user-intended function is applied to each input section.

The "map ()" function is applied to one input entry pair and returns a set of key-value pairs. The word "set" means that it may return only one entry, may not return anything, and may return several key-value pairs. What is in the key and in the value is defined by the user, but the key is a very important thing, since the data with one key will later on find its way into one instance of the "reduce" function.

2. The "Shuffle" stage. It goes unnoticed by the user. At this stage, the output of the "map" function is "distributed into baskets" – each basket corresponds to one output key of the "map" stage. Later on, those baskets will serve as inputs for the "reduce" function.
3. The "Reduce" stage. Each "basket" with the values formed at the "shuffle" stage finds its way into the "reduce ()" function.

The "reduce" function is determined by the user, and calculates the final result for a separate "basket". The set of all values returned by the "reduce ()" function is the final result of the MapReduce task.

VI. DEVELOPMENT OF KEY EFFICIENCY INDICATORS

The methodology and tools for working with structured data have long been created. This is a relational data model and database management systems. But under the present-day conditions, companies need to process large amounts of unstructured data of various types (Table 1), so the previous methods are not quite suitable for that task. New methods of handling data are needed. At present, the model of work with Big Data implemented in the Apache Hadoop project is becoming increasingly popular [9].

TABLE I. PREVAILING TYPES OF INFORMATION FOR VARIOUS FIELDS OF ACTIVITY WITH THE DEGREE OF USE

Field of activity\Type of information	Video	Images	Audio	Text/Numbers
Banking sector	Medium	Medium	Medium	Medium
Insurance	Low	Low	Low	High
Securities and investments	Low	Low	Low	High
Manufacture	Medium	Medium	Low	High
Retail trade	Medium	Low	Low	High
Wholesale trade	Low	Low	Low	High
Professional services	Medium	Medium	Medium	High
Entertainment	Medium	Low	Medium	Medium
Healthcare	Low	High	Low	High
Transportation	Medium	Medium	Low	High
Mass media	High	Medium	High	High
Utilities	Medium	Medium	Low	High
Civil engineering	Low	High	Low	Medium
Resources	Medium	Medium	Low	High
Government	High	Medium	High	High
Education	High	Medium	High	Low

Most products for working with Big Data have a highly efficient system for processing huge amounts of information and analyzing it in real time. The expected effect of the Big Data implementation may vary depending on the type of activity and the actual policy of a particular company (Fig. 4). When working with Big Data, methods of knowledge manipulation are used: various methods of the theory of recognition and classification, methods of intelligence analysis and data generalization, smart approaches in the form of genetic algorithms, neural networks and other branches of artificial intelligence. The sources indicate the relationship between the expected effect of the Big Data implementation and the field of activity and policy direction of a particular company [9].

The main tasks of the Hadoop platform are data storage, processing and management. The main components of the Hadoop platform are:

- Failure-resistant Hadoop Distributed File System (HDFS), which is used for storage purposes;
- The Map Reduce software interface, which is the basis for writing applications that process in parallel large volumes of structured and unstructured data on a cluster of thousands of machines;
- Apache Hadoop YARN performing data management function.

This reflects the Google citation index (Fig. 4). The Hadoop platform makes it possible to reduce the time for data processing and preparing, expands the possibilities for analysis, enables one to handle new information and unstructured data.

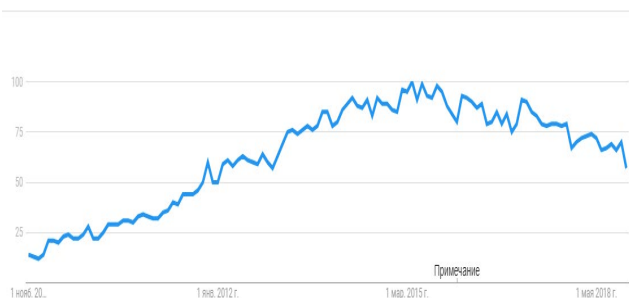


Fig. 4 Google citation index

The results of the project on the implementation of the Hadoop technology confirm the feasibility of its use (Table 2).

TABLE II. PROJECT RESULTS [9]

Platform	Description of the equipment	Approximate cost of the equipment, \$	Average time for one report, min.
Oracle database	Hi End Class Server	300 thousand	59
Hadoop cluster	10 workstations	7 thousand	66
Optimized Hadoop cluster	10 workstations	8 thousand	40

Solutions based on the Hadoop technology have a number of significant advantages. The main ones are given in Table 3.

TABLE III. ADVANTAGES OF A SOLUTION BASED ON HADOOP [9]

Advantage	Short description
Reduced data processing time	When processing data using a cluster, it is possible to significantly reduce the time for data processing.
Reduced equipment cost	The use of the Hadoop technology makes it possible to reduce the cost of equipment required for data storage and processing, by dozens of times.
Increased failure resistance; the technology makes it possible to come up with a failure-resistant solution.	Failure of one or several cluster nodes affects only the system productivity, while the system as such continues to operate properly providing service to the end users.
Linear scalability	The solution makes it possible to increase productivity simply by adding new cluster nodes. With that said, cluster productivity increases linearly.
Unstructured data operation capability	The technology makes it possible to conduct complex processing of any files including unstructured ones, so that such data can be efficiently processed and

used.

The Big Data storage solution is based on business requirements, workload management and smart storage ideas. As shown here, end users have many reporting possibilities and indicators that help them understand the use of tables, table spaces and workloads, since they relate to the frequency of access to data and its changes.

VII. CONCLUSION

Scientific methods aimed at introducing IT for processing large amounts of data with a distributed infrastructure based on smart agents and parallel algorithms have been examined. The emphasis is laid upon innovative methods based on smart agents and the principles of Industry 4.0. The implementation and simulation of parallel processing algorithms for Big Data and decision trees are being done.

The methods of operating BIG DATA have been formulated. The SAP Predictive Analytics modeling will be important in the process of collecting information from equipment sensors and in terms of the possibility of the follow-up analysis of the equipment, for example, for its proper functionality.

A review of technologies and business process control models has been carried out, during analysis of which recommendations were given that would be used for the predictive analysis in the BIG DATA environment.

REFERENCES ЛІТЕРАТУРА

- [1] A. Kupin, I. Muzyka, R. Ivchenko: "Information Technologies of Processing Big Industrial Data and Decision-Making Methods", Problems of Infocommunications. Science and Technology, 2018.
- [2] "Predictive analytics features: case from Beltel Datanomics" [Electronic resource] – Available: <https://iot.ru/promyshlennost/vozmozhnosti-prognoznoy-analitiki-keys-ot-beltel-datanomics>
- [3] A.M. Mariuta, Yu. G. Kachan, V. A. Bunko: "Automatic Control of Technological Processes of Concentrating Plants". Moscow «Nedra». p 1983.
- [4] "Electrotechnical Encyclopedia. Sensors" [Electronic resource] – Available: http://www.electrolibrary.info/subscribe/sub_16_datchiki.htm
- [5] R. Ivchenko: "Technology predictive analysis based on IoT TA BIGDATA". III International Scientific and Practical Conference "Information Security and Computer Technologies". Central Ukrainian National Technical University, Kropivnitsky, April 19-20, 2018.
- [6] Magazine "Habrahabr". "How to predict the exchange rate of the ruble to the dollar using SAP Predictive Analytics" [Electronic resource] – Available: <https://habrahabr.ru/company/sap/blog/345108/>
- [7] "Comparison of metadata editors". [Electronic resource] – Available: https://en.wikipedia.org/wiki/Comparison_of_metadata_editors
- [8] Magazine "Habrahabr". "Big Data from A to Z. Part 1: Principles of working with big data, the MapReduce paradigm". [Electronic resource] – Available: <https://habrahabr.ru/company/dca/blog/267361/>
- [9] Ivanov P.D., Vampilov V.Zh. "Big Data technologies and their application in a modern industrial enterprise." Engineering Journal: Science and Innovation, 2014. [Electronic resource] – Available: <http://engjournal.ru/catalog/it/asu/1228.html>
- [10] "Using big data in marketing research." [Electronic resource] – Available: <http://www.ovtr.ru/stati/bolshie-dannye-big-data-v-marketingovyh-issledovaniyah>.

Метод Вибору Маршруту для Перенаправлення Трафіку в SDN

А.В. Коган
ст. викл. каф. АСОІУ, ф-тет інформатики та
обчислювальної техніки
НТУУ «КПІ ім. І. Сікорського»
м.Київ, Україна
a.v.kohan433@gmail.com

Є.О. Павленкова
студентка каф. АСОІУ, ф-тет інформатики та
обчислювальної техніки,
НТУУ «КПІ ім. І. Сікорського»
м.Київ, Україна
pavlenkovayevheniia@gmail.com

Route selection method for traffic redirection in SDN

Alla Kohan,
Department of Technical Systems of
Automation and Control of Production,
Kyiv, Ukraine
a.v.kohan433@gmail.com

Yvheniia Pavlenkova
Department of Technical Systems of
Automation and Control of Production,
Kyiv, Ukraine
pavlenkovayevheniia@gmail.com

Анотація—У роботі розглянуто задачу прийняття рішення щодо вибору маршруту для перенаправлення трафіку з ушкодженої ділянки мережі з оцінюванням кожного варіанту за допомогою багатокритеріальної оцінки. Запропоновані основні параметри маршрутів, які впливають на його оцінку та набір варіантів маршрутів для перенаправлення.

Ключові слова— конструювання трафіку, перенаправлення трафіку, прийняття рішень, контролер SDN.

I. ВСТУП

Головним завданням комп'ютерної мережі є передача трафіку найбільш ефективно, надійно та швидко. Конструювання трафіку (Traffic Engineering, TE) є одним з головних методів для оптимізації роботи мережі [1]. Задачею конструювання трафіку є аналіз, передбачення та управління поведінкою даних у мережі, тобто вирішення того яким чином передавати дані по встановленим маршрутам, як розділяти ці дані між маршрутами.

Останнім часом спостерігається невинний процес зростання об'ємів трафіку [2] призводить до збільшення розмірності мереж і через це дедалі складнішим стає процес керування мережею та конструювання трафіку. Тому традиційні підходи до архітектури мережі та керування нею в найближчому майбутньому стануть неефективними, через це виникла необхідність пошуку та прийняття нової мережевої моделі, яка дозволить полегшити процес керування великими мережами.

Під час функціонування мережі можуть виникати такі ситуації (наприклад коли вузол мережі чи канал зв'язку виходить з ладу або через високу завантаженість ділянки маршруту виникає загроза утворення затори [3]), коли необхідно перенаправити трафік з того маршруту по якому він передавався на інший. В таких випадках пристрій мережі вирішує як найкраще «обійти» проблемну ділянку, маючи лише локальну інформацію про мережу. Такий підхід є традиційним, але щоб прийняти краще рішення щодо перенаправлення потрібно використовувати підхід при якому можна мати глобальне бачення стану мережі, тобто мати інформацію про всю мережу. Таким підходом є SDN (Software-Defined Networking), в ньому площина передачі даних та площина керування мережею розділені, управління мережею здійснює окремий пристрій – контролер SDN, завдяки чому досягається централізований контроль, це

дозволяє приймати краще рішення в описаних ситуаціях та спростити і підвищити ефективність управління мережею.

II. ПОСТАНОВКА ЗАДАЧІ

Так як контролер має всю необхідну інформацію про топологію мережі, то у випадку виникнення проблемної ситуації він може швидко її вирішити здійснивши реконфігурацію мережі та перенаправлення трафіку. Зазвичай є декілька варіантів для перенаправлення, і контролер, маючи глобальне представлення про мережу, може обрати найкращий з них, щоб забезпечити якомога кращу подальшу передачу даних. Отже, важливою задачею є прийняття рішення щодо вибору варіанту перенаправлення трафіку.

III. ВИБІР ВАРІАНТУ ПЕРЕНАПРАВЛЕННЯ

Теорія прийняття рішень є сильним математичним апаратом, це

поняття і методи математики, статистики, яка вивчає закономірності вибору шляхів вирішення різного роду завдань, а також досліджує способи пошуку найбільш вигідних з можливих рішень

Вона дозволяє формалізувати та впорядкувати процес прийняття рішення, завдяки чому простіше обирати найкраще рішення.

Теорію прийняття рішень можна застосовувати для задачі вибору найкращої серед кількох альтернатив, або їх впорядкування. Зазвичай будь-який об'єкт характеризується декількома властивостями або параметрами, і для того щоб прийняте рішення було якомога кращим потрібно оцінювати об'єкти по всім параметрам, а не лише по одному. Також потрібно брати до уваги відносну важливість обраних параметрів один до одного.

Пропонується застосувати даний підхід для вирішення задачі вибору контролером варіанту перенаправлення трафіку.

У якості варіантів маршрутів для перенаправлення пропонується обирати:

- маршрут, який обходить проблемну ділянку та повертається на свій початковий шлях;
- маршрут, який переходить на інший маршрут – найближчий до даного;

- маршрут, який переходить на інший маршрут – другий найближчий до даного.

На рис. 1 проілюстровані вказані варіанти.

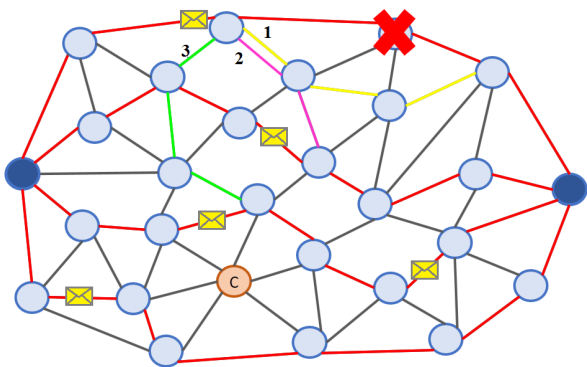


Рис. 1 – Ілюстрація маршрутів для перенаправлення

Маємо множину альтернатив $U = \{u_1, u_2, \dots, u_n\}$, в нашому випадку це маршрути по яким можна перенаправити трафік. Щоб обрати кращу альтернативу потрібно визначити їх оцінки. Для визначення оцінок альтернатив пропонується використати метод визначення багатокритеріальної нечіткої оцінки [4], його доцільно використовувати тому, що кожен маршрут характеризується кількома параметрами. Тобто потрібно зробити вибір на основі ступеню відповідності альтернатив деякій системі з m різних критеріїв C_1, C_2, \dots, C_m , для кожного з яких будується нечітка множина:

$$A_{C_i} = \{\mu_{C_i}(u_1), \mu_{C_i}(u_2), \dots, \mu_{C_i}(u_n)\},$$

де $\mu_{C_i}(u_j) \in [0,1]$ - оцінка j -ї альтернативи по i -му критерію, іншими словами це характеристика ступеню відповідності альтернативи даному критерію.

Критеріями виступають:

- завантаженість – вказує яка частина пропускної спроможності маршруту зайнята;
- довжина – кількість хопів у маршруті;
- час передачі – час передачі по даному маршруту у незавантаженій мережі;
- надійність – процент успішно переданих даних.

Найкращою альтернативою буде та, яка найбільшою мірою задовольняє усій множині критеріїв. Тобто правило знаходження оцінки по всім критеріям буде мати вигляд:

$$D = A_{C_1} \cap A_{C_2} \cap \dots \cap A_{C_m}$$

У відповідності до операції перетину нечітких множин функція належності для кожної альтернативи знаходиться за формулою:

$$\mu_D(u_j) = \min_{i=1,m} (\mu_{A_{C_i}}(u_j)), j=1,n$$

Найкраща альтернатива визначається за формулою:

$$\mu_D(u_j^*) = \max_{j=1,n} (\mu_D(u_j))$$

Найкращу альтернативу можна також визначити іншим методом [5], згідно якого на множині альтернатив будується нечітке відношення R з функцією належності $\mu_R(u_i, u_j) \in [0, 1]$ значення якої трактується як ступінь переваги i -ї альтернативи над j -ю.

Для побудови відношення R можна використати оцінки альтернатив отримані за допомогою наведеного вище методу.

По відношенню R будується нечітка множина невідомованих альтернатив $U_R^{nd} \subset U$, функція належності до множини U_R^{nd} :

$$\mu_R^{nd}(u_i) = \min_{u_j \in U} \{1 - \mu_R^S(u_j, u_i)\} = 1 - \max_{u_j \in U} \{\mu_R^S(u_j, u_i)\}, u_i \in U$$

Для кожної вершини u_i значення $\mu_R^{nd}(u_i)$ розуміється як ступінь невідомованості.

Обирається альтернатива з множини невідомованих з найбільшим значенням функції належності:

$$u^* = \arg \max_{u_i \in U} \mu_R^{nd}(u_i)$$

Такий спосіб вибору маршруту для перенаправлення трафіку дозволяє врахувати необхідну інформацію про мережу та обрати найкращий варіант.

ВИСНОВОК

В роботі описана задача прийняття рішення щодо вибору найкращого маршруту для перенаправлення трафіку з проблемної ділянки. Запропоновані підходи до її вирішення. Архітектура SDN надає можливість контролеру мати глобальне представлення про стан топології мережі, що дозволяє використовувати всю необхідну інформацію для прийняття рішення щодо перенаправлення трафіку таким чином щоб не викликати заторів або перевантаження на інших ділянках.

REFERENCES ЛІТЕРАТУРА

- [1] Zhaogang Shu Traffic Engineering in Software-Defined Networking: Measurement and Management/ Special section on green communications and networking for 5G wireless. - 2016. - Volume 4. – Pp. 3246- 3256.
- [2] Cisco, Inc. – https://www.cisco.com/c/ru_ua/index.html
- [3] Клеппман М. Высоконагруженные приложения. Программирование, масштабирование, поддержка/ Клеппман М. / 2017г. – С. 422-427.
- [4] Пономарев А.С. Нечеткие множества в задачах автоматизированного управления и принятия решений / Пономарев А.С. / Харьков, 2006г. – С. 104-107.
- [5] Блюмин С.Л. Модели и методы принятия решений в условиях неопределенности / Блюмин С.Л., Шуйкова И.А. / Липецк, 2000г. – С. 49- 52.

Система Захисту Програмного Забезпечення з Прив'язкою до Сервера

Куперштейн Леонід
Факультет інформаційних
технологій та комп'ютерної
інженерії

Вінницький національний технічний
університет, Вінниця, Україна
kupershtein.lm@gmail.com

Войтович Олеся
Факультет інформаційних
технологій та комп'ютерної
інженерії

Вінницький національний технічний
університет, Вінниця, Україна
voytovych.op@gmail.com

Беляєв Ігор
Факультет інформаційних
технологій та комп'ютерної
інженерії

Вінницький національний технічний
університет, Вінниця, Україна
ihor777770@gmail.com

Software Protection System With Server Binding

Kupershtein Leonid
Faculty of Information Technology and
Computer Engineering
Vinnytsia National Technical
University,
Vinnytsia, Ukraine

Voitovych Olesya
Faculty of Information Technology and
Computer Engineering
Vinnytsia National Technical
University,
Vinnytsia, Ukraine

Beliaev Igor
Faculty of Information Technology and
Computer Engineering
Vinnytsia National Technical
University,
Vinnytsia, Ukraine

Анотація — Запропоновано систему захисту програмного забезпечення на основі клієнт – серверної архітектури. Реалізація системи базується на основі токен авторизації та dll-ін'єкції.

Abstract — The Software protection system based on client-server architecture is proposed. The implementation of the system is based on the token-authorization and dll-injection.

Ключові слова — захист програмного забезпечення, клієнт-серверна архітектура, токен-авторизація, dll-ін'єкція.

Keywords — software protection, client-server architecture, token-authorization, dll-injection.

I. ВСТУП

Інформаційні технології дуже швидко розвиваються і їх використання у різних областях людської діяльності призводить до того, що крім задач передачі, збереження й обробки інформації виникла важлива задача її захисту.. Під рівнем безпеки програмного забезпечення (ПЗ) розуміється ймовірність того, що при заданих умовах у процесі його експлуатації буде отриманий функціонально придатний результат. Актуальність теми даної роботи зумовлюється тим, що на сьогоднішній день системи захисту програмного забезпечення надзвичайно поширені і знаходяться в постійному розвитку, а також завжди існує взаємодія між розвитком захисту та способами його зламу. Метою роботи є підвищення захищеності програмного забезпечення від несанкціонованого доступу з використанням методів прив'язки до серверу.

II. ОГЛЯД ІСНУЮЧИХ МЕТОДІВ ЗАХИСТУ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ВІД НЕСАНКЦІОНОВАНОГО ДОСТУПУ

Захист програмного продукту від несанкціонованого доступу (НСД) – актуальна задача у зв'язку зі збереженням комерційних і авторських прав фірм і розробників. За висновками закордонних фахівців, економічний збиток від піратського копіювання програмного забезпечення складає мільярди доларів [1].

Точні втрати установити неможливо через відсутність повних відомостей про число “піратських” копій, але вважається, що з кожної програми їх робиться від 2 до 15. Розрізняють декілька загальних методів захисту ПЗ від НСД.

Парольний захист. Такий захист є надійним інструментом від зловмисника-непрофесіонала, мінімальні незручності для користувача, можливість передачі пароля по мережі, відсутність конфліктів з системним та прикладним ПЗ. До недоліків можна віднести низьку стійкість більшості систем даного типу, користувачеві доводиться запам'ятовувати пароль /код.

Системи «прив'язки» ПЗ до апаратного забезпечення. При такому підході не потрібні додаткові апаратні засоби для організації захисту, ускладнення несанкціонованого доступу до скопійованого ПЗ, простота застосування. Недоліком є помилкові спрацьовування системи захисту ПЗ при будь-яких змінах в параметрах ПК, можливість конфліктів з системним ПЗ [2].

Засоби захисту ПЗ на основі електронних ключів. Такий підхід забезпечує значне ускладнення нелегального використання ПЗ, позбавлення виробника ПЗ від розробки власної системи захисту, висока автоматизація процесу захисту ПЗ, наявність API систем для більш глибокого захисту. Недоліками є ускладнення розробки і налагодження ПЗ через обмеження з боку засобів захисту, додаткові витрати на придбання системи захисту, зниження відмовостійкості ПЗ, несумісність систем захисту і системного або прикладного ПЗ користувача [3].

Мережевий захист. До технічних переваг цього підходу можна віднести те, що ні перевстановлення продукту, ні чистка реєстру не дозволяють повторно використовувати пробний період для однієї і тієї ж робочої станції. До недоліків захисту можна віднести вимогу доступу до мережі Інтернет на момент запуску захищеного продукту (у разі відсутності з'єднання він не запуститься), складність реалізації (необхідність підтримки різних мережевих протоколів) [4].

Для захисту ПЗ від НСД пропонується скористатися мережевим підходом на основі архітектури клієнт-сервер, при якому для запуску захищеної програми потрібно буде зв'язатися з сервером для передачі і перевірки ключа доступу до захищеної програми.

III. СТРУКТУРА СИСТЕМИ ЗАХИСТУ

Мережевий метод полягає в використанні захисту з прив'язкою до сервера на основі JWT-токенів, якими обмінюється захищене ПЗ з сервером, для автентифікації, використання ПЗ за призначенням і запобігання несанкціонованого використання.

На рис. 1 наведено загальну схему роботи системи захисту.

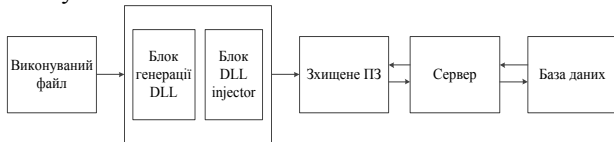


Рис. 1. Схему роботи системи захисту

Спочатку вибираємо виконуваний файл для накладання на нього захисту. Створюємо запис на сервері з логіном та паролем. На основі цих даних відбувається формування відкритого ключа та генерації dll-файлу. Згенерований dll-файл прив'язується до виконуваного файлу для накладання захисту шляхом виконання dll-ін'єкції. Отримане захищене ПЗ звертається до сервера та передає відкритий ключ. Сервер, після отримання відкритого ключа, звертається до бази даних, для отримання закритого ключа який формується на основі відкритого і відбувається перевірка. Останнім кроком є відповідь від сервера. Якщо вона позитивна – надає можливість доступу до захищеного ПЗ, інакше – відмовляє в доступі.

На рисунку 2 зображена структурна схема системи захисту.

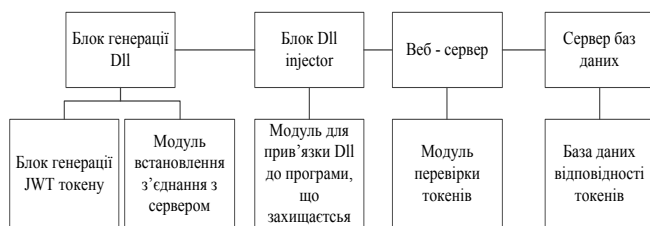


Рис. 2. Структурна схема системи захисту програмного забезпечення

Система захисту (рис. 2) складається з 4 головних компонентів:

- 1) Блок генерації dll, який містить модулі для генерації JWT-токенів та встановлення з'єднання із сервером.
- 2) Блок DLL-injector, який містить модуль прив'язки dll-файлу до програми.
- 3) Веб-сервер, який використовується для отримання запиту до програми, перевірки JWT-токену,

і відправлення дозволу (чи його відхилення) на запуск програми.

4) База даних, яка є сховищем відповідностей відкритих токенів до закритих.

Для роботи веб-сервера пропонується використання Apache HTTP Server і СУБД MySQL.

Для впровадження захисту використовується DLL – ін'єкція.

DLL – динамічна бібліотека, яка дозволяє багаторазове використання різними програмами.

Коли функції в DLL змінюються, додатки, які використовують їх, не потребують повторної компіляції або повторного компонування до тих пір, поки параметри функції, угоди про виклики і значення, які повертаються не зміняться;

Для передачі даних між програмою і сервером використовується сесійний ключ, який генерується в системі захисту. Для безпечної передачі JWT-токену від програми до сервера використовується асиметричний алгоритм шифрування HS256. При цьому секретний токен для розшифрування зберігається тільки на сервері, а відкритий – прив'язується до dll-бібліотеки додатку.

При першому запуску захищеної програми, модуль перевіряє наявність закритого ключа. Якщо його немає, виконуються наступні дії:

- 1) програма створює з'єднання до API сервера, і відправляє MAC-адресу, логін, згенерований JWT-токен на сервер.
- 2) веб-сервер отримує дані і проводить пошук в БД на наявність відповідного JWT-токену;
- 3) якщо такий токен існує, то веб-сервер відправляє команду для подальшого завантаження додатку;
- 4) ПЗ отримує відповідь від серверу і час життя токену, і працює у повному функціоналі.

ВИСНОВКИ

Представлений підхід до захисту програмного забезпечення є ефективним альтернативним варіантом захисту від ряду згроз порушення НСД, який може бути використаний компаніями, що займаються розробкою програмного забезпечення та хочуть захистити свій продукт.

ЛІТЕРАТУРА REFERENCES

- [1] Защита программ Назр [Електронний ресурс]. – Режим доступу : URL: <http://www.aladdin-rd.ru>. - Назва з екрану
- [2] Марков А. С., Фадин А. А., Цирлов В. Л. Систематика дефектов и уязвимостей программного обеспечения // Сборник трудов Второй всероссийской НТК «Безопасные информационные технологии» : [Специальный справочник] / под ред. В. А. Матвеева. – М: НИИ МГТУ им. Н. Э. Баумана, 2011, С. 83–87.
- [3] Галатенко В. А. Основы информационной безопасности : / А. В. Галатенко - М.: Мир, 2004. - 264 с.
- [4] Вихорев С. В. Классификация угроз информационной безопасности. [Електронний ресурс]. – Режим доступу : URL: http://www.cnews.ru/reviews/free/oldcom/security/elvis_class.shtml - Назва з екрану.

Елементна база та моделювання роботи прототипу пристрою для неінвазивного контролю рівня глюкози в крові

Ігор Когут

кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет
Івано-Франківськ, Україна
igorkohut2202@gmail.com

Михайло Котик

кафедра комп'ютерної інженерії та електроніки
Прикарпатський національний університет
Івано-Франківськ, Україна
mikl.kotykh@gmail.com

Elemental base and simulation of the prototype of the device for non-invasive blood glucose control

Ihor Kogut

dept. of Computer Science and Electronics
Precarpathian National University
Ivano-Frankivsk, Ukraine
igorkohut2202@gmail.com

Mykhailo Kotyk

dept. of Computer Science and Information Systems
Precarpathian National University
Ivano-Frankivsk, Ukraine
mikl.kotykh@gmail.com

Анотація— В даній роботі визначено ключові елементи неінвазивного пристрою для визначення рівня глюкози в крові, за допомогою САД-систем здійснено відлагодження параметрів їх роботи, промодельовано режими роботи даних пристроїв і визначено оптимальні складові для побудови прототипу даного виду пристроїв, здійснено аналогово-цифрове перетворення сигналу, отриманого із виходу фоточутливого елемента.

Abstract— In this article are determined the key elements of the non-invasive device for determining the level of glucose in the blood, with the help of CAD-systems, was performed debugging of their parameters, the operating modes of these devices were modeled and the optimal components for constructing a prototype of this type of devices were determined, analog-digital transformation of the signal received from the output of a photosensitive element is carried out.

Ключові слова— неінвазивний, САД-системи, відлагодження, моделювання, прототип

Keywords— non-invasive, CAD-systems, debugging, modeling, prototype

I. ВСТУП

Проблема моніторингу цукрового діабету зумовлена як значною поширеністю захворювання, так і розвитком на цій основі складних супутніх ускладнень, ранньої інвалідності та смертності. Тому актуальним є

оперативний та безперервний контроль рівня глюкози в крові (РГК), що підвищить показники лікування цукрового діабету.

На даний час розроблені малогабаритні системи діагностики РГК як для умов клінік, так і особистого користування. Такі системи в більшості є інвазивними, що вимагають безпосереднього забору крові. Водночас, набувають розвитку діагностичні пристрої, в основі яких лежать неінвазивні методи, які за точністю поступаються інвазивним, але є перспективнішими, оскільки володіють більшими функціональними можливостями з точки зору простоти вимірювань, накопичення і обробки результатів. Такі системи містять вимірювальний пристрій та сенсорні елементи. На жаль, немає достатньо даних щодо розробки вітчизняних малогабаритних пристроїв неінвазивного контролю РГК. [1-2]

Розробка та вдосконалення пристроїв для безперервного моніторингу глюкози в крові (БМГК) в даний час є важливим напрямком досліджень в діагностиці діабету. Актуальною залишається проблема вдосконалення приладів для цілодобового контролю РГК. Важливими первинними сенсорними елементами (СЕ) для неінвазивних пристроїв контролю РГК людини є фотодіоди і фотоприймачі, які функціонують на довжині хвилі $\lambda=940$ нм, на якій спостерігається максимальне значення поглинання глюкози та мінімальний вплив

інших факторів, наприклад, таких як вплив поглинання водою з максимальною інтенсивністю на довжині хвилі 960 нм. Для портативних систем неінвазивного моніторингу РГК з автономним джерелом живлення критичним параметром є споживана потужність. Тому однією із задач було знижене енергоспоживання пристрою. [3-4]

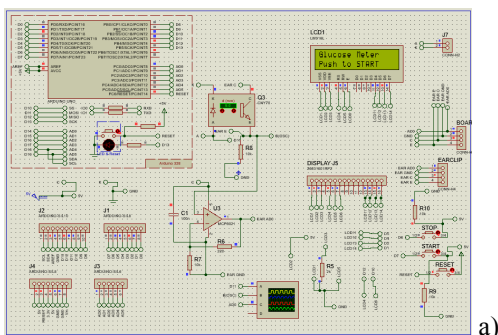
II. ЕЛЕМЕНТИ НЕІНВАЗИВНОГО ПРИСТРОЮ КОНТРОЛЮ РІВНЯ ГЛЮКОЗИ В КРОВІ

Розробка апаратно-програмованої платформи пристроїв неінвазивного контролю РГК та її елементів і трактів проходження сигналів є актуальною науково-прикладною задачею. Її вирішення дозволить отримати спеціалізовану елементну базу для проектування автоматизованих пристроїв неінвазивного контролю РГК.

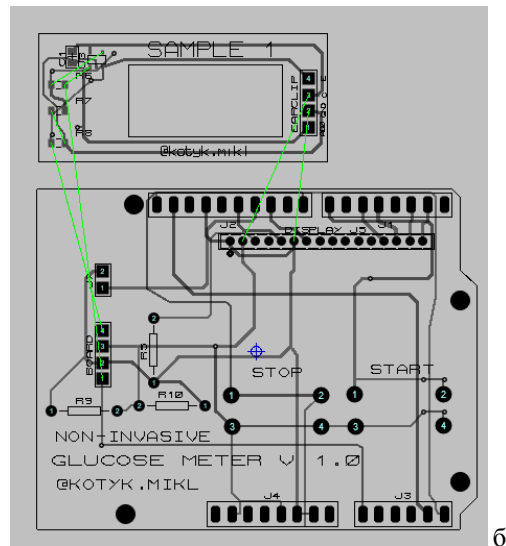
Було досліджено, що піки поглинання глюкозою ближнього інфрачервоного випромінювання становлять 650, 940, 1050 нм відповідно. В зв'язку із такою залежністю довжини випромінювання від кількості поглинання, було обрано оптимальну довжину хвилі - 940 нм. [5]. Пристрій складається із головного блоку та персональної вушної кліпси. [6] Головний блок побудований на базі мікроконтролера сімейства ATmega, і призначений для керування процесом неінвазивного вимірювання рівня глюкози в крові, задавання живлення світлодіода за допомогою ШІМ-сигналу, зчитування сигналу із фоточутливого елемента та аналогово-цифрової обробки сигналу. Протягом процесу вимірювання за допомогою спеціального програмного забезпечення проводиться збір 300 результатів, які в подальшому передаються в масив, і мікроконтролер проводить обробку даних та виведення їх у придатній для читання формі на LCD дисплей[7]

III. РЕЗУЛЬТАТИ САД-МОДЕЛЮВАННЯ СКЛАДОВИХ ПРИСТРОЮ ТА ЙОГО РОБОТИ В ЦІЛОМУ

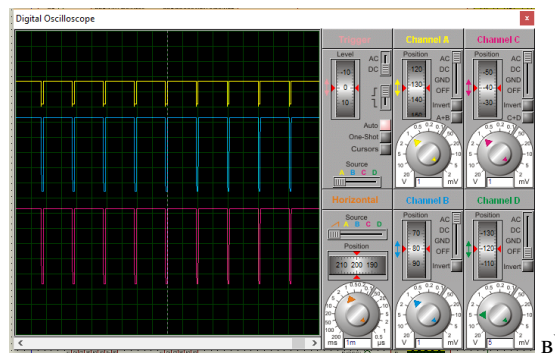
Проаналізувався складові системи неінвазивного контролю рівня глюкози в крові, було здійснено РСВ-дизайн та моделювання цих складових в САД-системі, та відлагоджено параметри цих елементів, зокрема було підібрано резистори обв'язки для операційного підсилювача. Результати РСВ-дизайну та моделювання роботи прототипу пристрою неінвазивного контролю рівня глюкози в крові зображено на рис. 1.



a)



b)



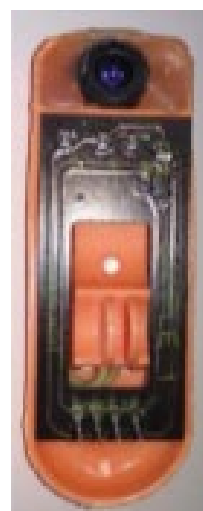
b)

Рис. 1. Схема, РСВ-дизайн та результати моделювання прототипу пристрою неінвазивного контролю рівня глюкози в крові, побудованого в САД системі.

На основі даних моделювання і дизайну було зібрано експериментальний зразок прототипу пристрою для неінвазивного дослідження рівня глюкози в крові. Його фото, а також складових елементів зображено на рис. 2.



a)



b)



Рис. 2. Загальний вигляд а), персональна вушна кліпса б) та головний модуль в) пристрою неінвазивного контролю рівня глюкози в крові.

ВИСНОВКИ

При проведенні аналізу експериментального зразка прототипу пристрою неінвазивного контролю рівня глюкози в крові були виявлені основні конструктивні та технологічні недоліки, при усуненні яких можна буде очікувати суттєве збільшення точності показання приладу, що в свою чергу дасть змогу використовувати даного виду пристрої для особистого неінвазивного визначення рівня глюкози в крові в домашніх умовах з достатньою тоністю.

Зокрема, як одну із основних причин втрати точності показу прототипу приладу було виявлено здатність навколишнього освітлення впливати на його показання. Цей недолік може бути усунено за рахунок обладнання приймача оптичного випромінювання смуговим фільтром, робочий діапазон котрого співпадає із діапазоном піку фоточутливості чутливого елемента..

ЛІТЕРАТУРА REFERENCES

- [1] International Diabetes Federation (IDF) diabetes atlas (2015) Brussels, Belgium: International Diabetes Federation.
- [2] Ong WM, SS Chua, CJ Ng (2014) Barriers and facilitators to self-monitoring of blood glucose in people with type 2 diabetes using insulin: a qualitative study. *Patient Prefer Adherence* 8: 237-246.
- [3] Vashist SK (2012) Non-invasive glucose monitoring technology in diabetes management: a review. *Anal Chim Acta* 750: 16-27.
- [4] do Amaral CEF, B Wolf (2008) Current development in non-invasive glucose monitoring. *Med Eng Phys* 30(5): 541-549.
- [5] Котик М. В. Сенсорні елементи і пристрої оперативної діагностики рівня глюкози в крові / М. В. Котик, І. Т. Когут // Журнал "Фізика хімія твердого тіла". Івано-Франківськ, 2016. – № 1 – С. 108-113.
- [6] Котик М.В, Особливості неінвазивного вимірювання рівня глюкози в крові із застосування ближнього інфрачервоного випромінювання./ Когут І.Т., Котик М.В, Саффури Х.Г //, *Фізика і хімія твердого тіла* - 2017. - Т.4, №4. - С. 618-624
- [7] М.Л. Петришин, В.А. Ровінський, "АЦП порозрядного врівноваження для пришвидшеного перетворення технологічних параметрів.", *Інформаційні технології та комп'ютерна інженерія*, Т.1., № 2 (39), с. 42-52, 2017.

Метод Вдосконалення Одношляхових Протоколів Динамічної Маршрутизації

Катерина Шевчук
Вінницький національний технічний університет
м. Вінниця, Україна
kateryna.shevchuk092@gmail.com

Сергій Захарченко
Вінницький національний технічний університет
м. Вінниця, Україна
zahap@i.ua

Method for Improving Single Path Protocols Dynamic Routing

Kateryna Shevchuk
Vinnytsia National Technical University
Vinnitsia, Ukraine
kateryna.shevchuk092@gmail.com

Sergiy Zakharchenko
Vinnytsia National Technical University
Vinnitsia, Ukraine
zahap@i.ua

Анотація: Принцип функціонування сучасних протоколів динамічної маршрутизації та способи призначення метрик зв'язків, виявили, що більшість з них є одношляховими, обирають один маршрут з мінімальною метрикою або здійснюють балансування між маршрутами з однаковою метрикою. Як результат функціонування мережі у такий спосіб спричиняє максимальне використання знайденого найкращого або альтернативного шляху. Вирішення цієї проблеми можливо здійснити через модифікацію процесу маршрутизації, скомбінувавши найкращі особливості розподіленого та централізованого способів маршрутизації. Після впровадження в мережу централізованого контролера маршрутизації, від всіх маршрутизаторів періодично надходить інформація про їх функціональний стан. Контролер при виникненні потреби через протокол SNMP вносить зміни в параметри маршрутизатора з подальшою зміною таблиць маршрутизації, а модуль керування проводить аналіз та розрахунки для ініціює зміни.

Ключові слова: трафік, протокол, динамічна маршрутизація, метрика, контролер, модуль керування, балансування, альтернативні маршрути, оптимізація потоку

Abstract: The principle of functioning of modern protocols of dynamic routing and methods of assigning communication metrics, found that most of them are one-way, choose one route with a minimal metric or carry out balancing between routes with the same metric. As a result of network operation in this way causes the maximum use of the found best or alternative path. This problem can be solved by modifying the routing process by combining the best features of distributed and centralized routing techniques. After the implementation of the centralized router controller, all routers will periodically receive information about their functional state. When required by the controller, the SNMP changes the router's parameters with the subsequent change of the routing tables, and the control module conducts the analysis and calculations to initiate the change.

Keywords: traffic, protocol, dynamic routing, metric, controller, control module, balancing, alternative routes, stream optimization

I. ВСТУП

В мережі може виникати ситуація, коли загальний трафік, що передається по вибраному шляху є дещо більшим за пропускну спроможність каналів. Для уникнення перевантаження необхідно ділити трафік на частини і перенаправляти частину пакетів по іншим шляхам. Однак при високій швидкості руху інформації, пакети надходять до вузла не частинами, а відразу повністю. В таких випадках, трафік має виражений пульсуючий характер, що підвищує ймовірність перевантаження на вузлах мережі.

Для уникнення перевантажень необхідними діями є:

1. Зміна маршруту передачі даних шляхом перенаправлення трафіку в обхід проблемного вузла.
2. Перенаправлення всього потоку на інший, менш безпечний шлях.
3. Розподілення трафіку з основного шляху на всі не завантажені маршрути, незалежно від рівня їх безпеки.

Разом з тим, характер подій в мережі в певний часовий проміжок залежить від попередніх віддалених подій. Це означає, що при великих масштабах мережі, трафік наділений властивостями подібності, тобто виглядає майже однаково при достатньо великих масштабах часового проміжку.[1]

Для усунення недоліку зацикловування пакетів у цих моделях, що проявлявся у мережах з напівдуплексними і (або) дуплексними каналами зв'язку [1] пропонується введення додаткових нелінійних обмежень, що гарантуватиме відсутність ефекту зацикловування пакетів без зниження якості балансування.

Інший недолік полягає в тому, що максимальне покращення балансування по каналам зв'язку можливе не для всіх топологій мережі. Особливо критичним зниженням якості балансування відбувалось у мережах з неоднорідною топологією, представлених у вигляді розподіленого графа.

Для мінімізації цього недоліку пропонується замінити централізоване керування методом маршрутизації «по підмережам», оскільки повна

ефективність балансування трафіку напряму залежить від якості балансування у «вузьких місцях» мережі, особливо це стосується неоднорідних мереж у вигляді розподіленого графа[1].

Слід зауважити, що дослідження, проведені з визначення проблем при передачі інформації в мережах та перспективних рішень не в повній мірі дозволяють усунути всі ймовірні недоліки, тому тематика статті, присвячена методу вдосконалення одношляхових протоколів динамічної маршрутизації доцільною та актуальною.

II. ПРИНЦИП РОЗПОДІЛЕНО-ЦЕНТРАЛІЗОВАНОЇ МАРШРУТИЗАЦІЇ З ВПРОВАДЖЕННЯМ КОНТРОЛЕРА МАРШРУТИЗАЦІЇ ТА МОДУЛЯ УПРАВЛІННЯ.

Для успішного адміністрування мережі необхідно знати стан кожного її елемента з можливістю змінювати параметри його функціонування. Зазвичай мережа складається з пристроїв різних виробників, і здійснювати управління нею було б нелегким завданням, якби кожен з мережевих пристроїв розумів тільки свою систему команд. Тому виникла необхідність у створенні єдиної мови управління мережевими ресурсами, яку б розуміли всі пристрої, і яка в силу цього, використовувалася усіма пакетами управління мережею для взаємодії з конкретними пристроями. Подібною мовою стала SNMP – Simple Network Management Protocol.

SNMP виконує такі функції:

- відправлення та прийом пакетів SNMP через протокол IP;
- збір інформації про статус і поточну конфігурацію мережевих пристроїв;
- зміна конфігурації мережевих пристроїв. •

Крім управління пристроями, SNMP використовують для моніторингу, оскільки він може отримувати різну інформацію від будь-яких мережевих пристроїв (роутер, свіч або комп'ютер, в яких є підтримка даного протоколу). Вміст одержуваної інформації може бути різним, наприклад : час аптайм, лічильники продуктивності CPU, мережі та ін, мережеві параметри пристроїв.

Для вирішення проблеми одношляхових протоколів маршрутизації пропонується здійснити структурні зміни в мережі, а саме модифікацію процесу маршрутизації шляхом впровадження в мережу централізованого контролера маршрутизації (рис. 1). Від всіх маршрутизаторів періодично надходить інформація про їх функціональний стан. Контролер надсилає запити до маршрутизаторів та при виникненні потреби через протокол SNMP надсилає команди та вносить зміни в параметри маршрутизатора з подальшою зміною таблиць маршрутизації. В свою чергу маршрутизатори при виникненні небажаних подій, аналогічним способом надсилають інформаційне повідомлення про це. Модуль керування вступає в мережу в якості допоміжного пристрою, який проводить аналіз отриманих параметрів, розраховує порогові значення, необхідні зміни до метрик маршрутів та розмір завантаження буферу черги для перерозподілу, в дуплексному режимі обмінюється інформацією з контролером.

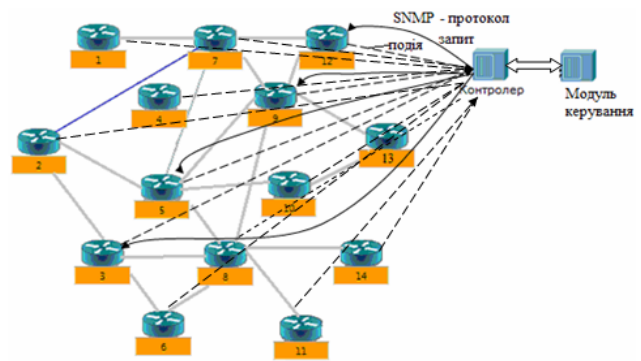


Рис. 1 – Функціональна схема мережі після модифікації.

У випадку виходу з ладу контролера, маршрутизатори продовжують працювати автономно за допомогою протоколів динамічної маршрутизації, тобто недолік централізованого способу маршрутизації (втрата працездатності мережі через вихід з ладу центру управління) в модифікованому методі буде усунуто, як і недолік змішаного способу (складність прогнозування стану мережі), оскільки при працюючому контролері зміни виявляються миттєво через постійний моніторинг за допомогою протоколу SNMP.

III. ЗАВДАННЯ МОДИФІКОВАНОГО ПРОЦЕСУ МАРШРУТИЗАЦІЇ

Загалом, для подальшої реалізації методу вдосконалення одношляхових протоколів динамічної маршрутизації, досягнення мети та модифікації процесу маршрутизації необхідно:

1. Визначити можливі варіанти та ознаки за допомогою яких може здійснюватися перевірка стану пристроїв або трафіку:

- Черги – накопичення пакетів на входних інтерфейсах, в тому числі буферні черги.
- Статистика переданого трафіку на інтерфейсі маршрутизатора.
- Завантаження центрального процесора кожного маршрутизатора.

При врахуванні всіх ознак та параметрів, модуль керування отримає можливість більш повно оцінити стан мережі та розподілити трафік без негативних наслідків. Мається на увазі в таких ситуаціях, коли черги та навантаження на інтерфейс незначне, однак у пристрої перевантажений центральний процесор. В такому випадку швидкодія маршрутизатора зменшується і, ймовірно, через це трафік може досягнути порогових значень та може виникнути перевантаження. Вирішення цієї проблеми можливо здійснити за допомогою введення комбінованого показника, який враховуватиме всі ознаки за рівнем важливості, зокрема, з введенням вагових коефіцієнтів.

2. Визначити маршрути для перенаправлення трафіку.

Перенаправлення трафіку має відбуватися розподілено. На кожний обраний для цього шлях, потрібно визначити розмір трафіку, який може передатися без перевантаження пристроїв та дотримання рівня порогових значень. Адже кожен шлях має різну пропускну спроможність, і маршрут з більшою пропускну спроможністю не може

передавати таку ж кількість перенаправленого трафіку, як і маршрут з меншою пропускною спроможністю (рис. 2).

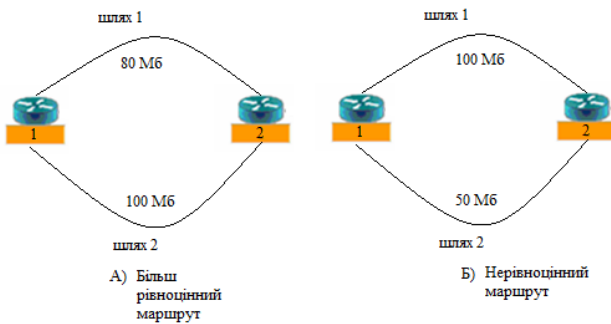


Рис. 2– Рівноцінні та нерівноцінні альтернативні шляхи

В даному випадку контролеру необхідно знати кількість трафіку, який він може передавати на альтернативні маршрути. Для моніторингу трафіку в мережі можна використати пропрієтарний відкритий протокол NetFlow, розроблений Cisco. Netflow надає можливість аналізу мережевого трафіку на рівні сеансів, роблячи запис про кожну транзакцію TCP / IP.

Архітектура системи будується на сенсорі, колекторі і аналізаторі:

- Сенсор (встановлюється на граничних маршрутизаторах сегментів мережі) збирає статистику по трафіку, який проходить через нього.

- Колектор здійснює збір інформації від сенсорів. Отримані дані він скидає в файл для подальшої обробки.

- Аналізатор, або система обробки, зчитує ці файли і генерує звіти у більш зручній формі.

Всі маршрутизатори будуть передавати інформацію про потоки на контролер, тоді контролер зможе через SNMP подавати статичні записи для певних підмереж, щоб пустити цей трафік іншим маршрутом, розділяючи його відповідно до можливої кількості передаваного трафіку без перевантаження альтернативного маршруту.

3. Способи перенаправлення трафіку та варіанти вплив на інші маршрутизатори для балансування трафіку та підтримки багатозв'язності.

- зміна стандартної метрики – розпаралелювання потоків.

- через параметр Unequal Cost Path Load Balancing динамічного протоколу маршрутизації EIGRP.

- зміна адміністративної відстані протоколу (ступінь довіри протоколу) – перенаправлення всього трафіку на інший маршрут. При одночасній роботі різних протоколів маршрутизації (RIP, OSPF, EIGRP) необхідно визначити тих, кому варто віддати перевагу. Плаваючі маршрути (Floating Routes) - статичний плаваючий маршрут такий самий, як і звичайний статичний маршрут, за винятком того, що цей тип маршруту має адміністративну відстань, якій можна налаштувати значення вище 1. Шляхом зміни цього параметру можна проводити перенаправлення трафіку на інші шляхи, це рішення є способом перевизначення протоколів маршрутизації для деяких випадків.

Використання статичного маршруту як резервного маршрут з змінюваною адміністративною

відстанню дозволить потім віддати перевагу шляху, який підтримує IGP (Internal Gateway Protocol), наприклад, OSPF. Відомо, що більшість маршрутів, отриманих протоколом OSPF, мають адміністративну відстань 110. У такому випадку адміністративна відстань статичного маршруту повинна бути менше 110, якщо маршрут OSPF на певний час потрібно виключити з таблиці маршрутизації. У випадку повного перевантаження маршрутизаторів та черг, доцільно використовувати цей спосіб перенаправлення.

ВИСНОВКИ

Характеристика сучасних протоколів динамічної маршрутизації, показала що більшість з них є одношляховими, тобто вибирають один маршрут з мінімальною метрикою, або здійснюють балансування між маршрутами з однаковою метрикою, що спричиняє максимальне використання знайденого маршруту, його перевантаження, в той час як інші ресурси мережі будуть в стані очікування - не задіяні в процесі передачі трафіку.

Для усунення недоліків передачі інформації в комп'ютерних мережах, зменшення затримки передавання трафіку шляхом модифікації процесу маршрутизації без внесення змін в протокол маршрутизації, розроблено комбінований розподілено-централізований варіант маршрутизації з демонстрацією можливості функціонування модифікованої мережі.

Використання запропонованого методу доцільно в процесі досягнення стану повноцінної рівноваги, балансованого розподілу навантаження між всіма можливими альтернативними шляхами.

REFERENCES ЛІТЕРАТУРА

- <http://ela.kpi.ua/handle/123456789/22143>.
- [2] Том М. Томас П. Структура и реализация сетей на основе протокола OSPF. Руководство Cisco = OSPF Network Design Solutions. — 2-е изд. — М.: «Вильямс», 2004. — С. 816. — ISBN 1-58705-032-3.
 - [3] Остерлох Х. Маршрутизация в IP-сетях. Принципы, протоколы, настройка / Х. Остерлох. — СПб.: BHV. — СПб., 2002. — 512 с.
 - [4] Medhi D., Ramasamy K. Network routing: algorithms, protocols, and architectures. Morgan Kaufmann, 2007. — 788 p.
 - [5] Sheng Huang, Biswanath Mukherjee. Adaptive Reliable Multi-Path Provisioning in WDM Networks. / in Proc. of IEEE ICC — Beijing, China, 2008. Pp. 5300-5304.
 - [6] Lei Song, Biswanath Mukherjee. On the Study of Multiple Backups and Primary-Backup Link Sharing for Dynamic Service Provisioning in Survivable WDM Mesh Networks / IEEE Journal on selected areas in Telecommunication, 2008. Vol. 26, No6. Pp. 84-91.

Дослідження Математичної Моделі Процесу Теплопередачі та Розробка на її Основі Пристрою для Контролю Температури Рідини

Михайло Шевчук
Факультет математики та інформатики
ДВНЗ «Прикарпатський національний університет імені Василя Стефаника»
Івано-Франківськ, Україна
shevchukmisha@i.ua

Анотація - У статті розглянуто математичну модель процесу теплопередачі в рідинах. Розроблено пристрій для контролю процесу теплопередачі з використанням датчика температури DS18B20 та Arduino-контролера.

Ключові слова—математична модель, теплопередача, рідина, arduino, DS18B20

I. ВСТУП

Сучасні технології математичного і розрахункового експерименту є ефективним інструментом досліджень, без розвитку якого неможливе подальше вдосконалення приладів та технологій, вирішення фундаментальних та прикладних проблем. Однією з таких проблем є забезпечення необхідних теплових режимів. Таким чином, актуальність вдосконалення математичних моделей, методів та алгоритмів, розробки на їх основі дослідницьких комплексів програмного забезпечення є очевидною.

На основі моделей, алгоритмів та методів аналізу даних автором проведено дослідження конвективного теплообміну. Завданням дослідження є мінімізація використання енергії для нагрівання рідини та збільшити ККД систем нагрівання.

Об'єктом дослідження є процес теплопередачі у рідинах, які знаходяться у металевих баках.

Предметом дослідження є математична модель теплопередачі та пристрої для контролю температури рідини.

II. ПРОЦЕС ПЕРЕНЕСЕННЯ ТЕПЛА РІДИНАМИ

Під кипінням рідини розуміють її випаровування, що супроводжується інтенсивним утворенням пухирців пари на поверхні дзеркала рідини та в об'ємі під впливом тепла, що підводиться через стінку. Кипіння рідини відбувається при такій температурі та в тому разі, коли парціальний тиск пари рідини, що випаровується, дорівнює або перевищує загальний тиск у паровому просторі [1]. Отже, кипіння рідини виникає на межі розділу холодного теплоносія з нагрітою стінкою, а температура стінки повинна бути вищою від рівноважного значення температури при даному тиску для конкретної речовини. Рушійною силою процесу теплообміну при кипінні є різниця температур між гарячою стінкою й температурою кипіння середовища. Під впливом рушійної сили на гарячій поверхні стінки відбувається утворення парових пухирців, що відриваються від стінки, піднімаються угору й

розширюються в об'ємі за рахунок зменшення гідростатичного тиску.

Інтенсивність кипіння залежить від величини рушійної сили та питомого навантаження теплового потоку, при цьому розрізняють пухирчасте та плівкове кипіння. Дослідженнями встановлено, що найбільша інтенсивність теплообміну на межі між середовищем і стінкою досягається при певній критичній різниці температур і певному критичному навантаженні теплового потоку q_{kp} (рис. 1).

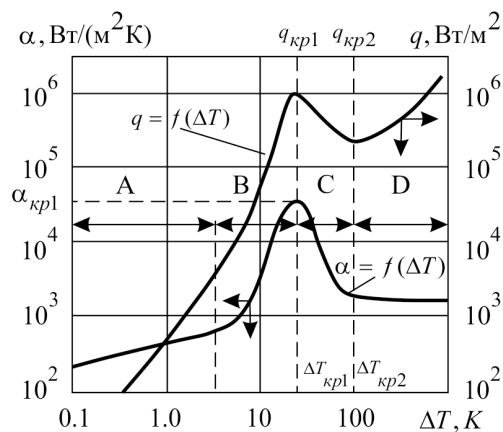


Рис. 1. Залежність щільності теплового потоку (q) і коефіцієнта теплообміну (α) при кипінні води від різниці температур та режимів кипіння: А – природна конвекція; В – пухирчасте кипіння; С – перехідна область; D – плівкове кипіння

Стінка покривається ізолюючим шаром тонкої парової плівки з низькою теплопровідністю, що призводить до перегріву стінки й можливого її пригорання. При значеннях різниці температур і щільності теплового потоку більше критичних величин (ΔT_{kp1} та q_{kp1}) коефіцієнт теплообміну різко зменшується, а потім зберігає майже постійне значення. Явище зменшення коефіцієнта теплообміну при переході від пухирчастого режиму кипіння до плівкового режиму одержало назву кризи теплообміну. Критична різниця температур і критична щільність теплового потоку залежать від природи киплячої рідини, тиску, матеріалу і стану поверхні стінки. На шорсткуватих і пористих стінках пухирчасте кипіння настає при меншому перепаді температур і коефіцієнт теплообміну вищий, ніж на гладких стінках [2]. Оребрення поверхні теплообміну дозволяє інтенсифікувати процес.

Інтенсивність теплообміну при кипінні рідин може досягати досить високих величин, що визначаються

щільністю питомого теплового потоку, близькою до $q=10^6$ Вт/м², та величиною коефіцієнта тепловіддачі на рівні $\alpha = 5 \cdot 10^4$ Вт/(м²·К).

У промислових теплообмінних апаратах, щоб уникнути переходу до нестабільного плівкового режиму кипіння та перегрівання стінок, прагнуть працювати при теплових навантаженнях, що не перевищують $q = 10^5$ Вт/м² та менші від критичних.

III. ВИМІРЮВАННЯ ТЕМПЕРАТУРИ

Вимірювати температуру у побутовій системі теплообміну на базі Arduino-контролера можна різними способами: за допомогою термопари, термометра опору чи цифрового давача температури [3].

При використанні термометра опору потрібно забезпечити стабільше джерело опорної напруги, застосування термопари вимагає можливості вимірювання напруги.

Для отримання фактичних значень температури у випадковій використанні термісторів можна використати рівняння Стейнхарта-Харта, що дозволяє реалізувати достатньо точну апроксимацію конвертованих значень.

Рівняння потребує значень змінних параметрів, які не завжди можуть бути відомими для конкретного термістора [4]. У такому випадкові використовують спрощене рівняння.

Простим та надійним варіантом є використання цифрового давача температури, наприклад, DS18B20.

Діапазон вимірювання температури складає від -55 до +125 °С. Для діапазону від -10 до +85 °С похибка не перевищує 0,5 °С.

У кожній мікросхеми DS18B20 є унікальний серійний код довжиною 64 розряди, який дозволяє декільком давачам використовувати спільну шину даних. Тобто, через один порт мікроконтролера можна обмінюватись даними з декількома датчиками, розташованими на різних відстанях. Такий режим зручний для використання у системах моніторингу температури, вузлах обладнання.

Особливості DS18B20:

- для однопровідного інтерфейсу 1-Wire достатньо одного порту зв'язку з контролером;
- кожен пристрій має унікальний серійний код довжиною 64 розряди;
- можливість підключення декількох давачів через одну лінію зв'язку;
- немає потреби у зовнішніх компонентах;
- можливість отримувати живлення від лінії зв'язку;
- напруга живлення може змінюватись у діапазоні від 3,0 В до 5,5 В;
- діапазон вимірювання температури -55 ... +125 °С;
- час замірів не перевищує 750 мс;
- можливість програмування параметрів тривожного сигналу;
- тривожний сигнал передає дані про адресу давачу, у якого температура вийшла за задані межі.

DS18B20 для обміну даними використано спеціалізований протокол 1-Wire корпорації Dallas. Усі пристрої фізично підключені до однієї загальної шини і використовують вихід з трьома станами або вихід типу відкритий колектор. У цій системі з однією шиною, мікроконтролер (майстер)

визначає наявність пристроїв на шині і обмінюється з ними, використовуючи унікальну адресу кожного пристрою – 64-розрядний код. Оскільки кожен термодавач має унікальний код, то кількість пристроїв, підключених до шини, практично нічим не обмежена.

Давачі температури DS18B20 під'єднують паралельно, для них досить одного резистора. За допомогою бібліотеки OneWire можна одночасно отримувати дані зі всіх давачів. Якщо кількість під'єднаних давачів – понад 10, потрібно підібрати резистор з опором не більше 1,6 кОм. Також для забезпечення точності вимірювання температури слід під'єднати додатковий резистор на 100 ... 120 Ом між виходом data на платі Ардуіно і data на кожному давачеві. Дізнатися, з якого давача отримано те чи інше значення, можна за допомогою унікального серійного 64-бітного коду, який буде виданий в результаті виконання програми. Приклад підключення декількох давачів в одну схему наведено на рис. 2.

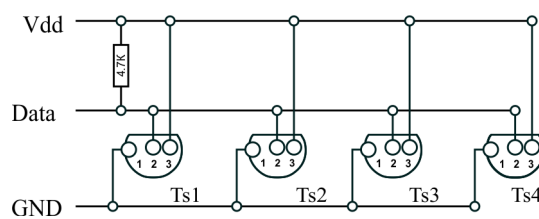


Рис. 2. Під'єднання 4 давачів DS18B20

ВИСНОВКИ

Використання моделі поширення тепла, особливо у відкритій системі, дозволяє керуючому контролеру оптимізувати витрати енергії та зменшити час нагріву рідини за рахунок врахування локальних особливостей теплопереносу.

Контрольоване нагрівання дозволяє урахувати інерційність процесу, таким чином нагрівач може бути вимкнено раніше, ніж досягнення верхніми шарами рідини заданої температури.

Множина давачів температури забезпечує можливість дослідження системи на відповідність ідеалізованій моделі для закритої системи та дозволяє врахувати невідповідність при використанні енергії при нагріванні рідини.

REFERENCES ЛІТЕРАТУРА

- [1] Федасюк Д. Моделирование теплообмена с учетом испарения на поверхности пластины / Д. Федасюк, Т. Муха // Вісник Національного університету "Львівська політехніка". – 2009. – № 650 : Комп'ютерні науки та інформаційні технології. – С. 151-157. – Бібліографія: 7 назв.
- [2] Основи теплоперенесення в елементах оптичного приладобудування : навчальний посібник / В. А. Ващенко, В. С. Антонюк, Г. С. Тимчак, І. В. Яценко, М. О. Бондаренко, О. В. Кириченко, М. П. Рудь ; Міністерство освіти і науки, молоді та спорту України, Національний технічний університет України "Київський політехнічний інститут". – Київ : НТУУ "КПІ", 2012. – 412 с. – Бібліографія: с. 359–370 (132 назви)
- [3] Основні положення конвективного теплообміну [Електронний ресурс] – Режим доступу URL: <http://posibnyky.vntu.edu.ua/tep/4.htm>. Дата звернення: 10.11.2017 р.
- [4] Петин В.А. Проекты с использованием контроллера Arduino. – СПб.: БХВ-Петербург, 2015. – 400 с.

СЕКЦІЯ 2

Інформаційні технології

Information Technologies

Інформаційне Забезпечення Діяльності Морського Агента в Сервісній Ергатичній Системі

Петров І.М.

Кафедра морських перевезень
Національний університет «Одеська морська академія»
Одеса, Україна
firmness@list.ru

Шибасва Н.О.

Кафедра інформаційних технологій
Одеський національний політехнічний університет
Одеса, Україна
nati.shibaeva@gmail.com

Рудніченко М.Д.

Кафедра інформаційних технологій
Одеський національний політехнічний університет
Одеса, Україна
murder8910@mail.ru

Шибасва Д.С.

Кафедра інформаційних технологій
Одеський національний морський університет
Одеса, Україна
denscreamer@gmail.com

Анотація—У даній статті наведені результати проектування інформаційного програмного забезпечення діяльності морського агента як ключового елемента сервісної ергатичної системи. Описана та формалізована специфіка діяльності морського агента, здійснено проектування схеми основних етапів розробки та розроблено діаграми засобами мови UML.

Ключові слова— сервісні ергатичні системи, морські перевезення, інформаційні системи, морські агенти

I. АКТУАЛЬНІСТЬ ТЕМАТИКИ

Підвищення ролі управління людськими ресурсами в морських компаніях вимагає врахування рівня професіоналізму і компетентності службовців при визначенні відповідності їх займаним посадам, виявлення слабо виражених або недостатньо розвинених якостей, вирішення питань атестації, кар'єрного просування, перепідготовки [1]. Застосовуваний в даний час анкетний підхід не забезпечує необхідного рівня ідентифікації функцій посади з властивостями і якостями самого працівника [2]. Згадані обставини зажадали створення адекватних моделей та засобів інформаційної підтримки праці службовців, які дозволяють відображати всі професійно необхідні характеристики фахівця, з можливістю їх формалізації для використання в сучасних сервісних ергатичних системах (СЕС). З огляду на, що такі моделі, як правило, є вербальними і відрізняються слабкою структурованістю, значною часткою нечіткості, невизначеності та суб'єктивності, розробка їх пов'язана з певними труднощами [3]. Ці обставини зумовили зміни в підходах до оцінки діяльності фахівців, зайнятих в сервісному сегменті на морському транспорті.

II. СПЕЦИФІКА ДІЯЛЬНОСТІ МОРСЬКОГО АГЕНТА В СЕС

Останнім часом значна увага приділяється питанням, пов'язаним з інформаційним забезпеченням СЕС на морському транспорті. Головним стає якість і доступність інформації, зручність її подання та використання в повсякденній діяльності. Інформація є потужним ресурсом, що дозволяє компаніям домогтися конкурентних переваг. Багато сервісних компаній переходять на нові стандарти і технології, пов'язані з використанням електронних форм забезпечення ділових

операцій. Прикладом тому є впровадження в морських портах України з серпня 2013 р Інформаційної системи портового співтовариства (ІСПС), учасниками якої стали власне порти, перевізники, карантинні служби, підприємницькі структури, які здійснюють сервісне обслуговування флоту. За допомогою ІСПС між контрагентами порту можуть бути здійснені обмін, перевірка, оформлення, розрахунки та передача інформації будь-якого виду, необхідних документів, для контролю і оформлення вантажів, що перевозяться і транспортних засобів, в електронному вигляді [4]. Впровадження ІСПС дозволило звести до мінімуму документообіг в паперовій формі і час оформлення вантажів, логістично вдосконалювати процес їх доставки до кінцевого одержувача, спростити і збалансувати схеми і процедури пропуску вантажів через порти України, і таким чином привести їх у відповідність з практикою передових зарубіжних портів. Це зажадало від усіх учасників ІСПС, в тому числі від сервісного сегменту, серйозних зусиль щодо вдосконалення інформаційного забезпечення виробничих процесів, в тому числі по розробці, модернізації та впровадження баз даних. Провідною постаттю в ньому є морський агент (МА), як «представник і помічник судновласника, який виступає від його імені і за його рахунок відповідно до повноважень, заснованими на законі або договорі, у всіх справах судновласника комерційного та адміністративного характеру, які пов'язані з морським перевезенням вантажів або пасажирів» [5]. У свою чергу, МА є елементом СЕС, її оператором. Для забезпечення якісного та ефективного функціонування таких систем повинні бути розроблені і враховані вимоги до навчання, добору, розстановки і діяльності МА як операторів СЕС та застосовані відповідні інформаційні технології з метою підтримки ключових виробничих процесів.

III. ПРОЕКТУВАННЯ ІНФОРМАЦІЙНОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

У зв'язку з наявністю великого числа функціональних обов'язків, що виконуються МА, який є ключовим елементом сучасних СЕС, необхідним є автоматизація процесу обліку, моніторингу та управління великими обсягами інформації. Оскільки

дані, що піддаються обробці МА, формуються на різних стадіях життєвого циклу роботи СЕС і можуть бути різнорідними, важливим аспектом є створення функціональної інформаційної системи (ІС) з управління такими даними, що включає до свого складу сховище інформації у вигляді бази даних (БД) і призначеного для користувача інтерфейсу прикладного програмного застосування.

З причини відсутності існуючих рішень по забезпеченню управління даними СЕС доцільним є розробка повноцінного інформаційного програмного забезпечення (ІПО), здатного функціонувати на різних операційних системах. Розроблюване ІПО управління СЕС має: бути стійким до паралельного використання системи різними користувачами, що володіють різними правами доступу до інформації; ефективно обробляти можливі позаштатні сценарії експлуатації і виняткові ситуації при модифікації даних шляхом інтеграції механізмів перевірки та затвердження даних; оперативно розгортатися на робочих станціях і підтримувати функції пошуку даних за необхідними критеріями. З метою структурного представлення основних етапів розробки ІПЗ СЕС складена схема основних завдань, наведена на рис.1. Першочерговим етапом є проектування структури БД у вигляді відповідних моделей даних, на базі стає можливим здійснення розробки проекту ІПО на основі застосування мови моделювання UML.



Рис. 1. Схема головних етапів розробки ІПЗ СЕС

Проект ІПЗ СЕС може бути розширений за допомогою розробки ряду додаткових діаграм, зокрема, схеми ключових компонентів системи (рис.2), які здійснюють процес обміну даними між собою в позначеній послідовності.

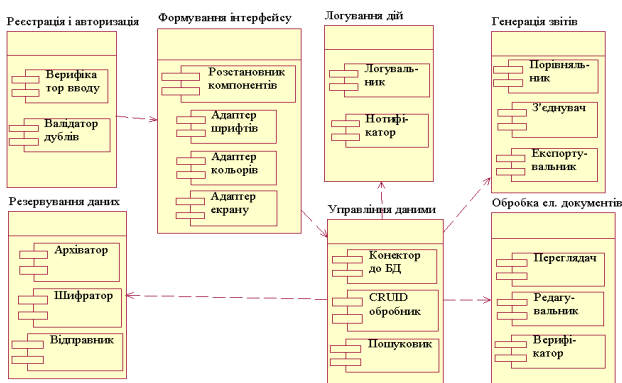


Рис. 2. Фрагмент діаграми головних компонентів ІПЗ СЕС

Це пов'язано з тим, що в процесі розробки системи в фізичному вигляді необхідно формалізувати ключові елементи логічного представлення в програмні сутності. Фізичне представлення проекту моделі ІПЗ СЕС реалізується за допомогою діаграми компонентів, найбільш значущими з них є наступні: реєстрація та авторизація, містять в своєму складі модулі верифікації даних, що вводяться в поля інтерфейсу користувача, а також валідатор наявності дублів в наявних облікових записах БД, щоб уникнути порушення цілісності і однозначності збережених даних; формування інтерфейсу, містить модулі розстановки компонентів графічного інтерфейсу користувача форм з управління даними, модулі адаптації базових властивостей і типів шрифтів, кольорних схем і дозволу екрану; управління даними, включає до свого складу модулі по організації сеансів зв'язку з БД за допомогою підключення до неї з програмного коду, обробки CRUD операцій (транзакційних запитів різного ступеня вкладеності) і здійснення пошуку за обраними критеріями; резервування даних, що містить модулі архівування даних, шифрування за обраним криптографічним алгоритмом і відправки на вказаний віддалений сервер або мережеву адресу; логування виконаних користувачем в системі дій з управління даними з модулем інформування в разі виникнення програмних колізій, виняткових ситуацій і помилок обробки даних; генерація звітів, що включає модулі порівняння звітів, злиття і сполучення зведених полів і експорту обраних масивів даних в файли обраного формату; обробка електронних документів, що містить модулі перегляду, редагування і верифікації вибраних файлів, які реалізують можливості нанесення спеціалізованих водяних знаків для захисту від несанкціонованого копіювання даних.

IV. ВИСНОВКИ

Розроблена концепція інформаційного програмного забезпечення діяльності морського агента в сервісній ергатичній системі є основою для реалізації повноцінної інформаційної системи, здатної інтегруватися в єдиний інформаційний простір (єдине морське вікно) в якості унітарного вузла. Використання запропонованої концепції дозволить підвищити ефективність праці МА як ключової ланки СЕС в процесі управління техніко-організаційними даними і документами в електронній формі.

REFERENCES ЛІТЕРАТУРА

- [1] V.A. Bobyr, "Management systems and ergatic systems in the navy." Scientific Herald of Kherson state maritime academy, no. 2 (9), 2013, pp. 9-19. (in Russian)
- [2] I.A.Garkina, A.M. Danilov, S.A. Pylaikin, "Practical methods of identifying transport ergatic systems " Almanac of modern science and education, no.8 (75), 2013, pp. 50-52. (in Russian)
- [3] A.M. Danilov, I.A. Garkina, E.A. Budylyna, "Practical methods of identification of the ergatic system." Industry Aspects of Technical Sciences, no.6 (30), 2013, pp. 3-5. (in Russian)
- [4] A.S. Bury, M.A. Shevkunov, "Intellectualization of decision making processes in ergatic systems." Transport business of Russia, no. 4, 2015, pp. 48-50. (in Russian)
- [5] A.S. Nugaev, A.M. Danilov, "Quality assessment of the control object of the ergatic system: quality functional, determination of weight constants." Magistracy Bulletin, no. 12-1 (39), 2014, pp. 16-19. (in Russian)

The measure of the information systems security risk assessment technique based on a fuzzy-multiple approach

N. V. Romashchenko
dept. SEMIT
NTU "KhPI"
Kharkiv, Ukraine
ronatavit@gmail.com

O. V. Shmatko
dept. SEMIT
NTU "KhPI"
Kharkiv, Ukraine
asu.spios@gmail.com

I. S. Litvinova
dept. SEMIT
NTU "KhPI"
Kharkiv, Ukraine
litjuli56@gmail.com

Abstract—The article provides an analysis of the main methods of information security risk assessment, identifying their advantages and disadvantages. The authors presented a technique that overcomes the uncertainty encountered in assessing the degree of information security risk. The presented technique provides the translation of the result obtained in a linguistic form.

Keywords—*information system, information security, risk assessment, fuzzy-multiple approach, linguistic form.*

I. INTRODUCTION

Information systems (IS) are now becoming increasingly popular in solving the issues of providing information base for various services that are engaged in solving technical, economic and a number of other different tasks.

When designing and developing reliable IS systems, it is necessary to anticipate a set of measures aimed at ensuring their protection against intentional or accidental impacts, which can lead to disruption of the system performance [1]. Such malfunctions in the work of IS incur serious losses in the form of information or real assets deprivation, the developer and owner of the IS reliability deterioration, and decrease in user credibility. That is why the task of ensuring Information Systems security is relevant.

Information security (InfoSec) assurance is part of the management of the information system as a whole. At the same time, one of the most important components of the InfoSec management system is risk assessment, which is designed to determine the effectiveness of the applied protection mechanisms. In the modern scientific community many scientists are engaged in the study of this issue. The main concern is to analyze existing techniques and improve their algorithms. For example, InfoSec risk analysis and the risk analysis process were classified, a comparison of software tools for managing InfoSec risks was presented in article [2]. Another example of research in this subject area is the work [3], which describes methods of risks assessment and management. In article [4] the author described the technique which uses.

In modern Information Systems, management functions are distributed over their various components connected to a common network. Unauthorized access by an attacker in order to override IS functions control can be obtained by directly manipulating applications or indirectly changing control parameters, as well as interfering with the data exchange between applications and devices. Therefore, it is necessary to protect the very management applications, as well as to protect the data exchange between applications and

devices. Thus, the purpose of this article is to develop a technique that allows to assess the degree of information system security risk in a linguistic form, comprehensible to the researcher.

II. INFORMATION SECURITY RISK ASSESSMENT TECHNIQUES ANALYSIS

There are a great number of risk assessment techniques in information systems. Consider a few of them.

The FAIR technique is based on the analysis of the factors that influence the risk components. The technique is divided into 4 stages: object assessment identification, threat frequency assessment, the rate of the potential losses probability assessment, risk results formalization [5].

The advantages of the technique include: complex analysis, simulation model use, high efficiency.

The FAIR technique is often applied to large enterprises. This technique is considered to be ineffective for small enterprises.

During the risk assessment procedure according to the MAGERIT technique, it is required to build an assessment model, draw up a risk map, evaluate existing protection mechanisms, and make an InfoSec risk rating. After that generating a system exploit report and developing a security plan is possible [6].

The advantages of the technique include: quantitative evaluation, flexibility, systematic method of analysis.

MAGERIT pays particular attention to the selection of participants in the InfoSec risk assessment process. The disadvantage of the method lies in the human factor influence on the resulting data.

Another popular technique for assessing the risk of InfoSec is NIST. The method consists of several stages. The initial steps determine the characteristics of the system, its vulnerability and threats. Next, an analysis of existing security measures and the threats realization likelihood is carried out and the incurred losses are assessed. After that, the risks magnitudes are determined and recommendations for the implementation of security mechanisms are made.

NIST has the advantage of a possible risks detailed description of the information assets suitable for IS of any scale.

The disadvantages of the technique include a long process of analysis, the lack of automation of some functions.

Security risks of IS are closely related to uncertainty. There are two cases of uncertainty: the identification of current and future system states. The above techniques do not take into account the uncertainty factor that must be overcome by a special mathematical apparatus.

III. GIVEN TECHNIQUE DESCRIPTION

The given risk assessment method for information systems can be determined by the sequence of the following steps.

Stage 1. At the first stage, term sets are introduced to describe the basic sets of the IS state and the subsets of states described in natural language.

- The full set of the InfoSec state E of the Information System assessment is divided into a subset of estimates.
- The corresponding set E of the complete set of the IS security risk degrees G is also divided into a subset of estimates. Hereinafter, we assume that the index G takes values from zero to one as defined.
- For an arbitrary given indicator of InfoSec estimate X_i , the full set of its values B_i is divided into subsets.

An additional condition for the correspondence of sets B , E and G of the following form is fulfilled: if all the indicators in the course of analysis have, according to the classification, the level of the subset B_{ij} , then the InfoSec condition qualifies as E_j , and the degree of InfoSec threat risk – as G_j . The fulfillment of this condition affects the correct quantitative classification of the indicators levels and the correct determination of the indicator significance level in the assessment system.

Stage 2. A set of individual indicators $X = \{X_i\}$ is built with the total number N , which, according to the expert analyst, on the one hand, affect the InfoSec threat risk assessment and, on the other hand, assess different in nature IS InfoSec.

Stage 3. Each indicator significance level for the analysis of r_i is compared.

Stage 4. At this stage, a classification of the current value g of the risk degree indicator G is made as a criterion for dividing this set into subsets.

Stage 5. Current values classification of x indicators of X is made as a criterion for dividing the full set of their values into subsets of type B .

Stage 6. Current level of indicators assessment is done and the obtained results are listed in a table, where each name of the indicator is correlated to its current value.

Stage 7. Current values x are classified according to the classification criterion corresponding to the fifth stage.

Stage 8. Bankruptcy risk assessment g is carried out at this stage. The main point of the stage is to estimate the importance coefficients of one or another subset B in assessing the state of IS InfoSec E and estimating the risk degree of IS InfoSec threats G .

These importance coefficients are further involved in external direct summation to determine the average value of g , where g_j is nothing more than average estimate g .

Stage 9. The obtained value of the IS InfoSec risk degree is classified on the basis of the classification obtained in the fourth stage. Thus, our conclusion about the threat risk degree of IS InfoSec takes on a linguistic form.

CONCLUSION

Information is one of the most important resources in modern IS, that is why it is necessary to assess the magnitude of vulnerability risk to anomalies and attacks. In this regard, already existing techniques for the information security risk assessment such as FAIR, MAGERIT, NIST are often used. At the same time, the above-mentioned methods do not take into account the fact that the IS security risks are closely related to the uncertainty that needs to be overcome. The proposed technique for the information security risk assessment can solve this problem – it overcomes the uncertainty and allows the researcher to assess the risk degree in a linguistic form. In the case of using the technique, the researcher gets the opportunity to formulate conclusions about the system security degree, and give advice on how to implement the necessary security mechanisms.

REFERENCES

- [1] C. P. Pfleeger and S. L. Pfleeger, Security in Computing, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2006, 921 p.
- [2] E. K. Baranova, "Methods of analysis and risk assessment Information security", Educational resources and technology, 2015'1(9), pp. 73-79.
- [3] I.V. Anikina, L. Yemaletdinova and A. Kirpichnikova, "Methods for assessing and managing information security risks in corporate information networks", Bulletin newspaper of technological university, 2015. Vol. 18, №6, pp. 195-197.
- [4] I.V. Sibkina, "Analysis of information security risks by using a fuzzy inference system", Science Bulletin of the NSTU, 2016, Vol. 65, No. 4, 2016, pp. 121-134.
- [5] I. Dobrynin, N. Maltseva, "Improvement of the method of factor analysis information risk", Information processing systems, 2017'3(49), pp.121-134.
- [6] O.V. Potij, A.V. Lenshin, "Research of information security assessment methods and guidelines design about its improvement on the ground of system approach", Coollection of scientific works of KhAI, 2010'2(24), pp. 85-91.

Детектування Наявності Додаткової Інформації в Цифрових Зображеннях за Умови Різного Ступеню Наповненості Контентів

Ахмамєтьєва Г.В.
кафедра інформатики та управління захистом
інформаційних систем
Одеський національний політехнічний університет
Одеса, Україна
a.v.akhmametieva@opu.ua

Мпуту Крістофер Бвабва
кафедра інформатики та управління захистом
інформаційних систем
Одеський національний політехнічний університет
Одеса, Україна
christopherbwabwa@gmail.com

Detection of the Presense of Additional Information in Digital Images in Conditions of Varying Degrees of Contents Fullness

Akhmametieva A.V.
Department of Informatics and Information Security
Management
Odessa National Polytechnic University
Odessa, Ukraine
a.v.akhmametieva@opu.ua

Mputu Chrisopher Bwabwa
Department of Informatics and Information Security
Management
Odessa National Polytechnic University
Odessa, Ukraine
christopherbwabwa@gmail.com

Анотація — В роботі пропонується удосконалення стеганоаналітичного методу, що здійснює аналіз просторової області кольорових цифрових зображень, шляхом розширення його можливостей з виявлення стеганоповідомлень, сформованих вбудовою додаткової інформації в дві та три колірні складові. Удосконалений метод забезпечує високу ефективність виявлення стеганоповідомлень за умови різного ступеню наповненості контейнерів.

Ключові слова — стеганоаналіз, цифрове зображення, послідовні тріади триплетів, просторова область, формат з втратами, формат без втрат.

Abstract — The improving of the steganalytic method which analyzes of the spatial domain of color digital images is presented. The developed method provides high efficiency in detecting stego formed with different degree of container fullness without reducing the accuracy of identifying the filled color components if the additional information was embedded into only one color component of the digital images.

Keywords — steganalysis, digital image, sequential triads of triplets, spatial domain, losses format, lossless format.

I. ВСТУП

Бурхливий розвиток інформаційних технологій та широке поширення електронного документообігу в усіх сферах життя людини сприяють зростанню загроз інформації, що зберігається, обробляється та передається між різними відділами, установами, організаціями. Зокрема, значного поширення набули випадки крадіжки конфіденційних даних за допомогою стеганографії, яка забезпечує приховування самого факту наявності додаткової інформації (ДІ) в будь-якому контейнері, який передається по відкритому каналу зв'язку. В якості

контейнера в стеганографії найчастіше використовують цифрові зображення (ЦЗ), результат вбудови ДІ в контейнер будемо називати стеганоповідомленням (СП).

Для попередження несанкціонованого отримання конфіденційних даних зловмисниками важливим є розвиток стеганоаналізу, основною задачею якого є виявлення факту наявності або відсутності ДІ в цифровому контенті, що аналізується [1].

Одним з сучасних стеганоаналітичних методів є метод, заснований на аналізі послідовних тріад триплетів в матриці унікальних кольорів [2], ефективність якого значно перевищує існуючі аналоги в умовах малих значень пропускної спроможності прихованого каналу зв'язку (ППС) за умови вбудови ДІ в одну довільну колірну складову ЦЗ. Однак означений метод не розглядає випадки, коли в процесі стеганографічного перетворення (СПП) задіяні дві або три колірні складові ЦЗ, що забезпечує приховування більшого об'єму інформації.

Тому метою роботи є удосконалення стеганоаналітичного методу, заснованого на аналізі послідовних тріад кольорних триплетів в матриці унікальних кольорів, шляхом розширення його можливостей у виявленні СП, сформованих в умовах різного ступеня наповненості кольорових ЦЗ.

II. ТЕОРЕТИЧНІ ОСНОВИ СТЕГАНОАНАЛІТИЧНОГО МЕТОДУ

В якості контейнерів будемо розглядати ЦЗ в форматі з втратами та в форматі без втрат, представлені колірною схемою RGB, де кожний піксель ЦЗ формується зі значень яскравості червоної, зеленої і синьої колірних складових – триплет (r_{mn}, g_{mn}, b_{mn}) . Всі унікальні триплети ЦЗ будемо називати унікальними кольорами.

В ході стеганоаналізу обчислюється кількість послідовних Red-, Green- і Blue-тріад в матриці унікальних кольорів UCT розміром $U \times 3$, яка містить U упорядкованих унікальних триплетів (r_i, g_i, b_i) , $i = \overline{1, U}$, де під послідовними Red-, Green- і Blue-тріадами розуміємо:

$$(r_k, g_k, b_k) \in UCT \ \& \ (r_k - 1, g_k, b_k) \in UCT \ \& \ (r_k + 1, g_k, b_k) \in UCT, \ k = \overline{1, U}; \quad (1)$$

$$(r_k, g_k, b_k) \in UCT \ \& \ (r_k, g_k - 1, b_k) \in UCT \ \& \ (r_k, g_k + 1, b_k) \in UCT, \ k = \overline{1, U}; \quad (2)$$

$$(r_k, g_k, b_k) \in UCT \ \& \ (r_k, g_k, b_k - 1) \in UCT \ \& \ (r_k, g_k, b_k + 1) \in UCT, \ k = \overline{1, U}; \quad (3)$$

відповідно. При підрахуванні кількості послідовних тріад триплетів послідовна тріада асоціюється з середнім триплетом $(r_k, g_k, b_k) \in UCT$, для якого виконуються умови (1), (2) або (3) в залежності від виду тріади.

Пустий контейнер в форматі з втратами містить не більше 2.5% середніх триплетів, які відповідають Red-, Green- і Blue-тріадам, але в результаті модифікацій просторової області ЦЗ кількість послідовних тріад значно зростає, що вказує на проведені перетворення в колірних складових ЦЗ та у випадку однієї заповненої колірної складової однозначно визначає модифіковану колірну матрицю. Однак при використанні контейнерів в форматі без втрат кількість послідовних тріад триплетів початково дуже висока (від 40 до 70%), що не дозволяє на даний час ідентифікувати такі зображення як СП.

III. РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

Оскільки в результаті СП просторової області ЦЗ стеганоповідомлення зберігаються в форматі без втрат, стеганоаналітику невідомо, чи є цифровий контент, що аналізується, пустим контейнером в форматі без втрат або контентом з ДІ, тому визначення формату оригінального контейнеру тісно пов'язано з процесом стеганоаналізу. Але характер змін у кількості середніх триплетів відрізняється в залежності від формату оригінального контейнеру та ступеню наповненості. Рис. 1 ілюструє вміст Red-, Green- і Blue-тріад у випадках незаповненого контейнеру в форматі з втратами (I), коли для вбудови ДІ була задіяна лише зелена колірна складова (II), задіяні червона і синя колірні складові (III), задіяні всі три колірні складові (IV), пустого контейнеру в форматі без втрат (V). У випадках II, III, IV ППС становить 0.2 біт/піксель.

Як видно з рис. 1, при визначенні СП слід враховувати різницю у кількості послідовних тріад триплетів ЦЗ в форматі без втрат та модифікованих зображень в форматі з втратами.

В ході проведених експериментів було встановлено, що повторна вбудова довільної ДІ в ЦЗ в форматі без втрат не призведе до значущих змін у кількості середніх триплетів, в той час як СП реагують на повторну модифікацію пікселів дуже сильно.

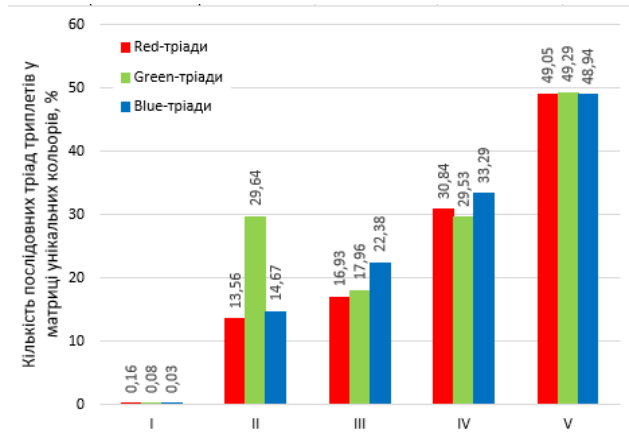


Рис. 1. Вміст послідовних тріад триплетів у матриці унікальних кольорів ЦЗ залежно від ступеню наповненості

На основі результатів дослідження характеру змін у кількості послідовних тріад триплетів у матриці унікальних кольорів, було удосконалено стеганоаналітичний метод для виявлення СП, сформованих в умовах різного ступеню наповненості ЦЗ, який містить наступні блоки:

- 1) Обчислення кількості середніх триплетів у матриці унікальних кольорів.
- 2)
- 3) Перевірка, чи сформоване ЦЗ вбудовою ДІ в одну колірну складову і визначення матриці, яка була задіяна в процесі СП.
- 4) Вбудова довільної ДІ в три колірні складові ЦЗ, що аналізується.
- 5) Відокремлення стеганоповідомлень, сформованих в умовах двох і трьох заповнених колірних матриць, від контейнерів в форматі без втрат.

Для виконання стеганоаналізу ЦЗ в середовищі MATLAB був реалізований програмний інтерфейс, представлений на рис. 2 і 3 для випадків однієї та двох заповнених колірних складових ЦЗ.

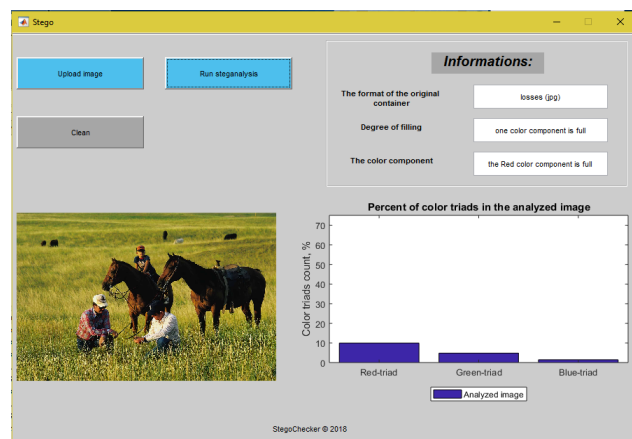


Рис. 2. Виявлення СП, сформованого вбудовою ДІ в червону колірну складову ЦЗ в форматі з втратами методом S-Uniward [3]

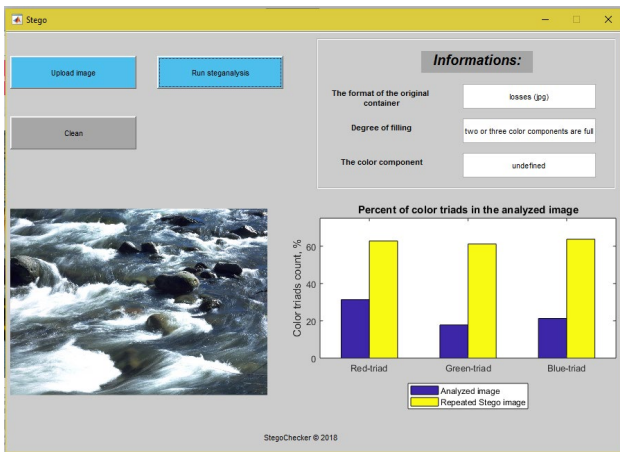


Рис. 3. Виявлення СП, сформованого вбудовою ДІ в зелену та синю колірні складові ЦЗ в форматі з втратами методом LSB Matching

IV. ОЦІНКА ЕФЕКТИВНОСТІ УДОСКОНАЛЕНОГО СТЕГАНОАНАЛІТИЧНОГО МЕТОДУ

Ефективність удосконаленого стеганоаналітичного методу будемо оцінювати на основі 718 незаповнених контейнерів та 2218 СП, сформованих з різним ступенем наповненості, за трьома критеріями:

- критерій 1 – відокремлення незаповнених контейнерів від СП;
- критерій 2 – виявлення колірної матриці СП у випадку вбудови ДІ в одну колірну складову ЦЗ;
- критерій 3 – визначення формату оригінального контейнеру.

У разі критеріїв 1 і 2 під помилками першого роду будемо розуміти пропуск СП, під помилками другого роду – визначення незаповненого контейнера як СП. У разі критерію 3 під помилками першого роду будемо розуміти пропуск факту стиску за його наявності, під помилками другого роду – визначення цифрового контенту як такого, що не піддалося стиску, у той час коли стиск ЦЗ відбувався. Результати обчислювального експерименту наведені в табл. 1.

ТАБЛИЦЯ 1. ЕФЕКТИВНІСТЬ ВИЯВЛЕННЯ СП І ВИЗНАЧЕННЯ ФОРМАТУ ОРИГІНАЛЬНИХ КОНТЕЙНЕРІВ ДЛЯ ЦЗ, %

	Критерій 1	Критерій 2	Критерій 3
<i>ППС 0.2 біт/піксель</i>			
Помилки 1-го роду	1.8485	0.3035	0.1022
Помилки 2-го роду	2.3677	0.3549	16.9184
<i>ППС 0.1 біт/піксель</i>			
Помилки 1-го роду	9.2425	0.8346	0.0341
Помилки 2-го роду	2.3677	0.3549	16.9184

Як видно з табл. 1, проведення удосконалення стеганоаналітичного методу забезпечує високу ефективність виявлення СП, сформованих вбудовою ДІ в контейнери в форматі з втратами, про що свідчать помилку першого та другого роду за критеріями 1 і 2. Однак при визначенні формату оригінального контейнеру помилки другого роду складають 16.9184%, що вимагає додаткових досліджень для забезпечення більш точного відокремлення форматів оригінальних ЦЗ, які використовуються в процесі СП.

ВИСНОВКИ

В роботі запропоновано удосконалення стеганоаналітичного методу виявлення вкладень ДІ, заснованого на врахуванні кількості послідовних тріад триплетів в матриці унікальних кольорів ЦЗ, шляхом розширення його області застосування, а саме можливості виявляти СП, сформовані з різним ступенем наповненості цифрових контентів.

REFERENCES ЛІТЕРАТУРА

- [1] R. Bohme, Advanced statistical steganalysis, Springer, 2010.
- [2] А.В. Ахметьєва, “Стеганоанализ цифровых изображений, хранящихся в формате с потерями”, Захист інформації, вип. 23, с.135-145, 2016.
- [3] V. Holub, J. Fridrich, T. Denemark, Universal Distortion Function for Steganography in an Arbitrary Domain, EURASIP Journal on Information Security, (Section:SI: Revised Selected Papers of ACM IH and MMS 2013), 2014

Аналіз параметрів алгоритму цифрового підпису на основі гешування W-OTS

Олександр Марухненко
Харківський національний
університет радіоелектроніки
Харків, Україна
oleksandr.marukhnenko@nure.ua

Геннадій Халімов
Харківський національний
університет радіоелектроніки
Харків, Україна
hennadii.khalimov@nure.ua

Михайло Наконечний
Харківський національний
університет радіоелектроніки
Харків, Україна
mykhailo.nakonechnyi@nure.ua

Analysis of the parameters of the hash-based digital signature algorithm W-OTS

Oleksandr Marukhnenko
Kharkiv National University of
Radioelectronics
Kharkiv, Ukraine
oleksandr.marukhnenko@nure.ua

Gennady Khalimov
Kharkiv National University of
Radioelectronics
Kharkiv, Ukraine
hennadii.khalimov@nure.ua

Mikhailo Nakonechnyi
Kharkiv National University of
Radioelectronics
Kharkiv, Ukraine
mykhailo.nakonechnyi@nure.ua

Анотація - В цій роботі надається опис алгоритму одноразового цифрового підпису Вінтерніца, що базується на гешуванні. Проводиться аналіз основного системного параметру та його вплив на розмір підпису та складність обчислень.

Ключові слова - Підпис на основі гешування, підпис Вінтерніца, одноразовий підпис, постквантова криптографія.

Abstract— In this paper we give a description of the algorithm of the Winternitz one-time digital signature, based on the hexification, is given. The analysis of the main system parameter and its influence on signature size and complexity of calculations are carried out.

Keywords— Hash-based signature, Winternitz signature, onetime signature, postquantum cryptography.

I. ВСТУП

Постійний прогрес у створенні квантових комп'ютерів у поєднанні з ефективно працюючим на них алгоритмом факторизації Шора представляє серйозну небезпеку найбільш поширеним асиметричним криптосистемам (RSA, DSA, ECDSA та їм подібні). У зв'язку з цим в останні роки ведеться активний пошук альтернативних криптоалгоритмів, стійких до квантових атак, так званих "постквантових алгоритмів".

Одним з існуючих варіантів постквантового цифрового підпису є сімейство ЦП на основі геш-функцій. Вперше подібний алгоритм був запропонований Лампортом ще в 1979 році [1], однак, він мав ряд недоліків і не набув популярності. Діффі незначно його модифікував, уточнивши вимоги до геш-функцій, що використовуються. Алгоритм ЦП Вінтерніца [2] відрізняється від попередників наявністю ітеративної геш-функції, дана модифікація дозволила зменшити розмір ключів і підпису. Серйозним недоліком даних алгоритмів є їх "одноразовість", тобто кожна пара ключей могла бути використана для підпису лише

одного повідомлення. Рішенням цієї проблеми стала ідея Мерклі [3] про використання так званих геш-дерев або дерев Мерклі, іншим варіантом багаторазового підпису є HORS [4]. В даний час найбільш перспективними є схеми XMSSMT [5], SPHINCS[6] та його модифікації Gravity-SPHINCS[7] та SPHINCS+[8], які побудовані на основі композиції вказаних алгоритмів.

В даній роботі надається опис цифрового підпису Вінтерніца, наведено залежність між значенням параметру Вінтерніца, розміром підпису та обчислювальною складністю генерації та перевірки підпису.

II. ОПИС АЛГОРИТМУ

Розглянемо алгоритм одноразового ЦП Вінтерніца (W-OTS). В подальшому під повідомленням будемо мати на увазі геш-значення від фактичного повідомлення, отриманого з використанням криптографічної геш-функції

$$H(M) : \{0,1\}^* \rightarrow \{0,1\}^m.$$

Загальносистемні параметри:

- одностороння функція $f : \{0,1\}^n \rightarrow \{0,1\}^n$;

- параметр Вінтерніца $w > 1$ - кількість біт повідомлення, які будуть одночасно підписані, в деяких джерелах параметром Вінтерніца називають значення 2^w , ми будемо дотримуватися першого варіанту.

Під ітеративною геш-функцією розуміють

$$f^k(x) = \underbrace{f(f(\dots f(x)\dots))}_k. \quad (1)$$

Вводять наступні додаткові параметри:

$t_1 = \lceil \frac{m}{w} \rceil$ - кількість бітових блоків довжини w в повідомленні.

$$t_2 = \left\lceil \frac{\log_2(t_1(2^w - 1))}{w} \right\rceil + 1 \quad \text{блокова довжина}$$

контрольної суми.

$t = t_1 + t_2$ - блокова довжина повідомлення з контрольною сумою.

Приватний ключ..

$$SK = (sk_0, sk_1, \dots, sk_{t-1}), \quad sk_i = \{0, 1\}^n \quad (2)$$

випадкові бітові послідовності

Для економії пам'яті в якості секретного ключа можна використовувати одну випадкову послідовність, яка при необхідності буде розгорнута в SK за допомогою генератора псевдовипадкових послідовностей.

Публічний ключ:

$$PK = (pk_0, pk_1, \dots, pk_{t-1}), \quad pk_i = f^{2^i - 1}(sk_i) \quad (3)$$

Створення підпису.

Вхідні дані: M, H, SK .

Вихідні дані: σ .

1) Повідомлення розбивається на t_1 блоків довжини w біт (при необхідності в кінці додаються нулі)

$$H(M) = (h_0, h_1, \dots, h_{t_1-1}). \quad (4)$$

2) Визначається контрольна сума

$$S = \sum_{i=0}^{t_1} 2^w - 1 - h_i = (s_0, s_1, \dots, s_{t_2}). \quad (5)$$

3) Контрольна сума додається до повідомлення

$$B = H(M) \| S = (b_0, b_1, \dots, b_{t-1}), \quad (6)$$

де $\|$ означає конкатенацію.

4) Обчислюється підпис

$$\sigma = (\sigma_1, \sigma_0, \dots, \sigma_{t-1}) = (f^{h_1}(sk_0), f^{h_0}(sk_1), \dots, f^{h_{t-1}}(sk_{t-1})) \quad (7)$$

Перевірка підпису.

Вхідні дані: M, H, σ, PK .

Вихідні дані: так/ні.

1) Аналогічно попередньому пунктам 1-3 алгоритму генерації підпису обчислюється

$$B' = (b'_0, b'_1, \dots, b'_{t-1}). \quad (8)$$

2) Обчислюється

$$Z = (z_0, z_1, \dots, z_{t-1}) = (f^{2^w - b_1}(\sigma_0), f^{2^w - b_0}(\sigma_1), \dots, f^{2^w - b_{t-1}}(\sigma_{t-1})) \quad (9)$$

3) Якщо $Z = PK$, підпис коректний.

Використання контрольної суми дозволяє захистити підпис від підробки: якщо зловмисник має правильний підпис, то через певну кількість обчислень він зможе створити повідомлення, з геш-значенням $H(M') = (h'_0, h'_1, \dots, h'_{t-1})$, таким що $h'_i \geq h_i$, таким чином

після додаткових ітерацій вихідний підпис стане правильним для повідомлення M' . Однак при збільшенні h_i значення контрольної суми зменшується, отже, елементи ключа, що використовуються для підписання контрольної суми, повинні бути прогешовані меншу кількість разів, ніж у вихідному повідомленні, що зловмисник не може зробити через односторонній характер функції.

Існують різні модифікації ЕЦП Вінтерніца, зміни, як правило, стосуються ітеративної функції, що використовується, серед них:

- перед черговим гешуванням на число операцією XOR накладається бітова маска, що дозволяє знизити вимоги до геш-функції (алгоритм WOTS+, який використовується в підписі SPHINCS та його модифікаціях) [9];

- використання ключової геш-функції, де результат попередньої ітерації використовується як ключ; число, що гешується генерується випадково та передається разом з підписом[2];

- зменшення розмірів відкритого ключа у t разів за рахунок використання модифікованого дерева Мерклі – L-дерева [6].

III. РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

Параметр Вінтерніца визначає розмір підпису, обчислювальну складність генерації відкритого ключа, створення та перевірки підпису. Розмір підпису дорівнює

$$SignSize = t * n \quad (10)$$

біт, для створення та перевірки підпису необхідно виконати

$$N = t * (2^w - 1) \quad (11)$$

операцій гешування. Розрахуємо дані параметри для конкретного прикладу. Нехай, ітеративна геш-функція обчислює значення довжиною 256 біт. У таблицях I та II наведені результати для підпису повідомлень довжиною 256 та 512 біт.

Таблиця I. РЕЗУЛЬТАТИ РОЗРАХУНКІВ ДЛЯ ПОВІДОМЛЕННЯ ДОВЖИНОЮ 256 БІТ

w	t	Розмір підпису (Кбайт)	Кількість ітерацій
2	133		399
3	90	2.81	630
4	67	2.09	1005
5	55	1.72	1705
6	45	1.41	2835
7	39	1.22	4953
8	34	1.06	8670
9	31	0.97	15841
10	28	0.88	28644
...
16	18	0.56	$1.18 * 10^6$

ТАБЛИЦЯ II. РЕЗУЛЬТАТИ РОЗРАХУНКІВ ДЛЯ ПОВІДОМЛЕННЯ ДОВЖИНОЮ 512 БІТ

w	t	Розмір підпису (Кбайт)	Кількість ітерацій
2	261	8.16	783
3	175	5.47	1225
4	131	4.09	1965
5	106	3.31	3286
6	89	2.78	5607
7	76	2.38	9652
8	66	2.06	16830
9	59	1.84	30149
10	54	1.69	55242
...
16	34	1.06	$2.23 \cdot 10^6$

З таблиць видно, що на практиці можна використовувати лише невеликі значення параметра Вінтерніца, оскільки кількість ітерацій для генерації відкритого ключа, створення та перевірки підпису експоненційно залежить від даного параметру, що значно уповільнює ці процеси. Оптимальним, на наш погляд, є значення $w = 4$, розмір підпису в такому випадку складає 2 кілобайти та потребує близько 1000 операцій гешування, що є прийнятною величиною. Розбиття даних, що підписуються, на 4-бітні блоки легко реалізується як програмно, так і апаратно.

REFERENCES ЛІТЕРАТУРА

- [1] Lamport L. Constructing digital signatures from a one-way function. — SRI International Computer Science Laboratory, 1979.
- [2] Johannes Buchmann, Erik Dahmen, Sarah Ereth, Andreas Hülsing, and Markus Rückert. On the security of the Winternitz one-time signature scheme. In A. Nitaj and D. Pointcheval, editors, *Africacrypt 2011*, volume 6737 of *Lecture Notes in Computer Science*, pages 363–378. Springer Berlin / Heidelberg, 2011
- [3] Ralph Merkle. A certified digital signature. In Gilles Brassard, editor, *Advances in Cryptology – CRYPTO '89*, volume 435 of *LNCS*, pages 218–238. Springer, 1989.
- [4] Leonid Reyzin and Natan Reyzin. Better than BiBa: Short one-time signatures with fast signing and verifying. In Lynn Batten and Jennifer Seberry, editors, *Information Security and Privacy 2002*, volume 2384 of *LNCS*, pages 1–47. Springer, 2002.
- [5] Johannes Buchmann, Erik Dahmen, and Andreas Hülsing. XMSS — A Practical Forward Secure Signature Scheme Based on Minimal Security Assumptions. In Bo-Yin Yang, editor, *Post-Quantum Cryptography 2011*, volume 7071 of *Lecture Notes in Computer Science*, pages 117–129. Springer Berlin / Heidelberg, 2011
- [6] Daniel J. Bernstein, Daira Hopwood, Andreas Hülsing, Tanja Lange, Ruben Niederhagen, Louiza Papachristodoulou, Peter Schwabe, and Zooko Wilcox O’Hearn. Sphincs: practical stateless hash-based signatures. *Cryptology ePrint Archive*, Report 2014/795, 2014.
- [7] Jean-Phillippe Aumasson and Guillaume Endignoux. Gravity-SPHINCS – Submission to the NIST’s post-quantum cryptography standardization process, 2017. Available at <https://github.com/gravity-postquantum/gravity-sphincs>.
- [8] Daniel J. Bernstein, Christoph Dobraunig, Maria Eichlseder, Scott Fluhrer, Stefan-Lukas Gazdag, Andreas Hülsing, Panos Kampanakis, Stefan Kolbl, Tanja Lange, Martin M. Lauridsen, Florian Mendel, Ruben Niederhagen, Christian Rechberger, Joost Rijneveld, and Peter Schwabe. SPHINCS+ – Submission to the NIST’s post-quantum cryptography standardization process, 2017. Available at <https://sphincs.org/>.
- [9] Andreas Hülsing. W-OTS+ — shorter signatures for hashbased signature schemes. In Amr Youssef, Abderrahmane Nitaj, and AboulElla Hassanien, editors, *Progress in Cryptology – AFRICACRYPT 2013*, volume 7918 of *Lecture Notes in Computer Science*, pages 173–188. Springer Berlin Heidelberg, 2013.

Trading in real time using an expert Advisor using Parabolic SAR, MA, CCI technical analysis methods

Oleg Motuzov
Grodno state University
Grodno, Belarus
motusov12@mail.ru

Abstract— The purpose of this work is to test the expert Advisor in real time using different methods of analysis. The object of the research is Forex trading. The subject of the study is a market Advisor.

Keywords— *Forex, market, Advisor, method, analysis.*

Анотація— Мета роботи - це тестування радника в реальному часі, що використовує різні методи аналізу. Об'єктом дослідження є торгівля на ринку Форекс. Предметом дослідження є радник ринку.

Ключові слова-*Форекс, ринок, радник, метод, аналіз.*

I. INTRODUCTION

Currently, information technology has reached such a level of development that it is possible to manage funds from your personal computer. You can, without leaving your own apartment, around the clock to buy or sell currency, making transactions within a few seconds through the information and trading terminal. Mechanical trading systems – expert advisors are used to automate trading. Advisors, independent work that perform all kinds of operations, from market analysis, calculation of available capital, to the opening and closing of transactions. Develop expert advisors based on trading strategies that show positive results for a certain period of time.

II. BENEFITS ADVISOR

Using expert advisors provides the following advantages:

- does not require a permanent stay at the computer;
- excludes the human factor;
- accelerates the process of making a profit;
- reduces errors when opening and closing trades;
- round the clock trade.

III. DEVELOPMENT ADVISER

MACCISAR expert Advisor was developed for trading, which uses Parabolic SAR, MA, CCI methods.

In order to start trading in real time, you first need to test the EA in the previous period. For testing, the settings of time interval, Deposit, currency pair, time frame, leverage are set. First, the EA is tested without

optimization, and then the parameters and optimization of the EA are set. The result should be a Sharpe ratio greater than one. The higher the profit when using a trading strategy, the higher the risk. And at some point, the risk of losing outweighs the probability of making a profit. The Sharpe ratio is a parameter that shows how much the income from the strategy relates to the potential risk. What should be the Sharpe ratio:

- "1 and above" - the optimal value of the coefficient, indicating a good strategy or high performance portfolio management;
- "0-1" - we can not say that the strategy is very good, because the risks are too high, but its application is possible;
- "0 and below" - on Forex strategy is better not to use, when investing in stock it is better to choose a different portfolio.

IV. TRADING RESULT

For more relevant results of the EA, we will trade in real time. The initial balance is \$ 2000. Trading was performed on EUR/USD, GBP/USD, USD/JPY currency pairs. The trading period with 07.05.2018 for 17.05.2018. The trade balance amounted to 2817,71\$. In total, 90 trades were made. Net profit amounted to 817,71\$, total profit 848,15\$, the total loss -60,42\$. Of the 90 perfect trades 80 trades were profitable. The biggest profit was \$ 54.78, the biggest loss - \$5.60. The Sharpe ratio was 1.13 and this suggests that the EA is now fully stable, the risks of its use are minimal.

REFERENCES ЛІТЕРАТУРА

- [1] Shabaker R. Technical analysis and stock market profit. Moscow, 2005, - pp.112.
- [2] Motuzov O. A. expert Advisor, developed on the basis of technical analysis methods of Parabolic SAR, CCI, MA, Economics and regional management: collection of articles of the international scientific-practical conference. Bryansk: BSU, 2017. - pp.699-703.
- [3] Шабакер Р. Технічний аналіз і прибуток фондового ринку. – Москва., 2005. – 112 с.
- [4] Мотузів О. А. Тестування торгового радника, розробленого на основі методів технічного аналізу Parabolic SAR, CCI. Економіка та регіональне управління: збірник статей міжнародної науково-практичної конференції / упоряд. І. Г. Чернишова-Брянськ: БГУ, 2017. - С. 699-703.

Using VGG-Type Models for Image Classification with Pytorch

Dmytro Myroniuk
Department of Radiophysics and Computer Technologies
Lviv National University named after Ivan Franko
Lviv, Ukraine
myronyukdmytro@gmail.com

Використання моделей VGG-типу для класифікації зображень із застосуванням бібліотеки PyTorch

Дмитро Миронюк
Кафедра радіофізики та комп'ютерних технологій
Львівський національний університет імені Івана Франка
Львів, Україна
myronyukdmytro@gmail.com

Abstract—There using open models for image classification (VGG (Visual Graphics Group) - type models of Oxford University) and their custom modifications to solve image classification task are considered. The models are trained on the open CINIC 10 data set and are tested using the CINIC-10 test example. PyTorch framework is used for model construction and training. The different training methods are analyzed and described techniques to prepare datasets for different type of tasks.

Keywords—Deep Learning, Image Classification, PyTorch, Python

I. INTRODUCTION

The widely used methods and solution graphs for popular image processing task – image classification are analyzed. This type of tasks can be used in different tasks, include business and scientific tasks. Today, image classification tasks solve with convolution neural networks that can be used to solve large specter of image processing tasks with impressive results. This networks become some “golden standard” of solutions for image recognition tasks (image classification, object detection, instance segmentation etc.), text recognition tasks (use Recurrent Neural Networks (RNN) as base solution) etc. From 2013 up to 2017 proposed different convolution models with pretrained weights and open datasets (with objective challenges), that should help to increase model accuracy.

II. USED FRAMEWORK

PyTorch framework is used for model construction. This is one of the most popular deep learning framework, that used by outstanding companies like Facebook. The main advantage of that library over other is that models in PyTorch have dynamic computational graph. So, one can change our computational graph without recompilation. Moreover, PyTorch have many optimized pretrained finetuning models that can be used as base for custom

powerful models. Computational graph and learning process plots can be visualized via compatible tensorboard interface. Main type that used in computations is tensor (optimized for GPU using multi-dimensional array).

III. USED MODEL

A. Convolutional classifier

Every convolutional classifier consist of two main parts: feature extractor (convolutional layers, that transform our RGB image array to maps that represent some features, which found on our image. As classifier network use fully-connected layers with special architecture: 2-3 layers with classic sigmoid activation functions on first layers and softmax function on last layer. The typical convolution classifier architecture is presented on figure 1.

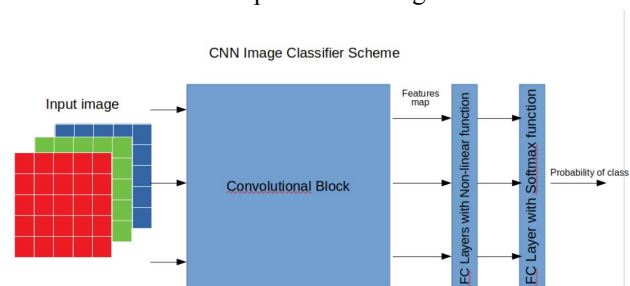


Fig. 1 Typical architecture of convolutional classifier

In 2018, there are many benchmarks for popular models (that can be included in model embedded libraries or not), which can be used by scientists to choose optimal model by many parameters (model accuracy, model computational complexity, hardware requirements etc). For 2018, one used MS COCO (Microsoft Common Objects in Context) [3] and PASCAL VOC 2012 [4] as top list of used models. In addition, the benchmark [2] is analyzed.

Base architecture use simple sequential type of model with one main flow and without any flops in convolutional block

(as in ResNet architectures). Convolution block in model VGG-16 consist of 5 subblocks with several separate convolution operations layers that used nonlinear ReLU functions to trim negative and separate main features. VGG-net (VGG-16) architecture is presented on figure 2.

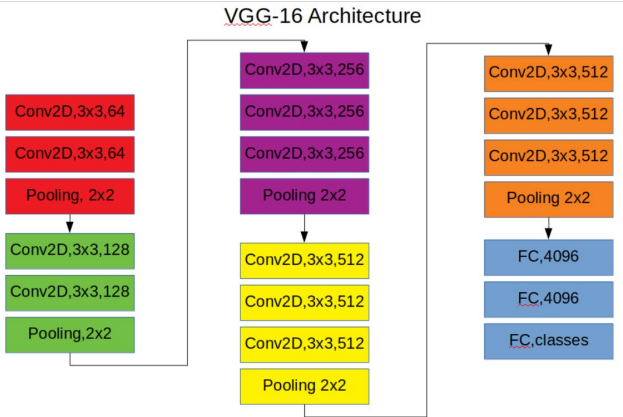


Fig. 2 Visual representation of VGG-16 architecture
B. Model overview

As one can see in picture 2, VGG-type architecture can be separated into some blocks, each of them consist of 2 types of layer – Convolution operation with using non-linear function and pooling operation.

- Convolution operation. Typical convolution block consist of two main parts: entire convolution, where our input image in RGB matrix representation multiply with list of filters–matrixes, which represent classic weights and improve during learning. After that, we have an output feature maps, which highlights some important pixels – features, which can be used as identifier for certain classes of images. Convolution operation realizes 3 main principles, different from ANN:

- 1) Sparse interactions. Convolution kernels, which used in this type of networks, help to reduce dimension of weights compare to input tensor. This feature reduces computational complexity of a task and reduces volume of memory in RAM or GPU memory.
 - 2) Parameters sharing – in classic networks every parameter used only ones to compute neuron output. In convolution, one kernel uses to compute output in every region of input image. This feature helps to reduce computational complexity of network too.
 - 3) Eqvariance to broadcast – if input function transform, then output function transform with the same method. Convolution operation can be computed by following formula:
- Pooling. This operation apply function of maximum function (Max Pooling), Average (average pooling), l2 – norm or weighted average on feature maps. Functions of this layer:
- 1) Reduce size of input tensor (list of feature maps after convolution can use large volumes of memory)
 - 2) Release independence from minor changes of input tensor.
 - 3) Distinguish entire entity rather than location of entity.
 - 4) Help to disassociate local signs in favor of global signs.

$$S(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i-m, j-n) K(m, n)$$

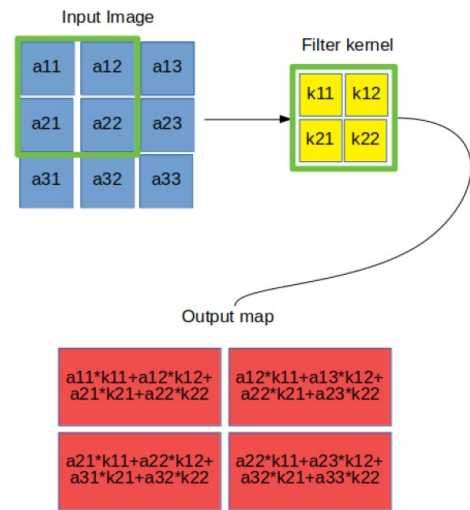


Fig. 3 Visualization of convolution

IV. LEARNING

A. Convolutional neural classifier learning process

To learn CNN there are special modifications of back propagation method that allow learn a kernel matrixes to highlights important features. On classifier stage one use standard back propagation function for weights correction:

$$w = w + \Delta w$$

$$\Delta w = \frac{-\eta \partial E}{\partial w}$$

For convolutional layers one used kernel correction based on partial gradients of errors in feature map cells. Visual representation of process is presented in Figure 4..

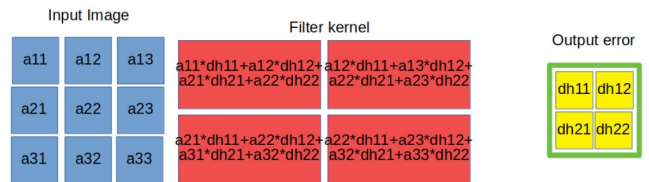


Fig. 4 Back propagation function for convolutional layer.

There are 2 main techniques that used to learn deep learning system, based on convolutional neural networks:

- Training from scratch. Model weights with random values and train model are initialized. This method help to train best and most specific model for concrete task, but one should use extremely large dataset (about 1000 images per one class) with powerful hardware for training (what can be expensive).
- Transfer learning. One can used pretrained model with weights, which oriented on recognition of some features. This method can be used for adjacent task. That model is proposed for using and retraining only classifier layers without training convolution kernels from scratch. That feature can be used to improve general accuracy of this model, reduce computational complexity of solutions and reduce number of labeled images for training. There are 3 techniques of transfer learning method:

1) Using completed models from library. We just use model from pretrained model list and replace standard classifier

with custom (based on our number of output classes of images)

2) Design own model based on described architecture. One can design model with specialized architecture and use pretrained weights from popular competitions like MS COCO, ImageNet, Pascal VOC, etc., which can be found on official site of the competition.

3) Combining several models and use several archives of weights.

V. HARDWARE RESOURCES

To design and train models one can use notebook Dell Inspiron 7567. Characteristics of this computer presented in table 1.

To improve speed of that model training one can use GPU training with Nvidia CUDA software. This method widely used in deep learning to improve training and recognition speed.

TABLE 1 . USED HARDWARE RESOURCES

Parameter	Value
Processor	Intel Core i5 7300 HQ 2.5-3.4 GHz
RAM	8 GB
ROM	SSHD TOSHIBA MQ02ABD1 1 Tb
GPU	Nvidia GTX 1050 4 GB
CUDA	v. 9.0.76 with CuDNN v.7.0
Driver GPU	v.390.77
OS	Ubuntu 18.04
PyTorch	0.4.1.post2
Python	3.6.5

VI. COMPUTATIONAL EXPERIMENTS

For experiments we use new CINIC-10 dataset with 270000 images, separated into training set (90000), validation (90000) and test (90000). For the first experiment we use our custom architecture with two standard blocks of VGG with 2 convolutional layers/ Kernel size – 2, count of filters – 64 in all layers. To avoid overfitting of model we use Dropout layers between convolutional blocks. We train this model from scratch. Model hyperparameters:

- 1) Criterion: cross-entropy loss;
- 2) Learning rate: 0.001;
- 3) Momentum:0.9
- 4) Batch size: 4

In the second experiment use VGG-16 as the base classifier. Here we apply transfer learning technique to provide solution of this task.

Model hyperparameters:

- 1) Criterion: cross-entropy loss;
- 2) Learning rate: 0.001;
- 3) Momentum=0.9,
- 4) Weight decay=5e-4
- 5) Batch size: 128

Learning results presented in table 1.2

Method	Model loss	Validation accuracy (90000 samples)
Custom model	0.864	63%
VGG-16	0.293	65.56%

Training plots for our custom model is presented in Figure 5.1 (Model Loss) and 5.2 (Model Training Accuracy)

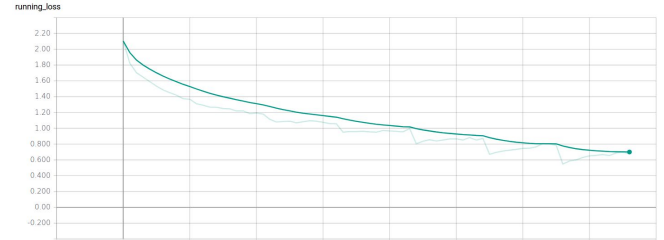


Fig. 5.1 Training plot for Model Training Loss (custom model)

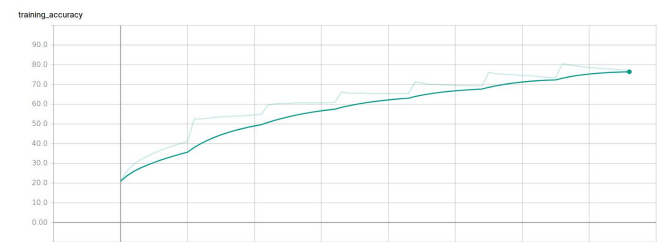


Fig. 5.2 Training plot for Model Training Accuracy (custom model).

CONCLUSION

Trained models demonstrate good results on large validation dataset. Our model with pretrained weights on ImageNet dataset and VGG-16 features extractor has higher accuracy than custom model, but custom model has lower computational complexity.

REFERENCES:

- [1] Hui, J. (2018, March 28) Object detection: speed and accuracy comparison (Faster R-CNN, R-FCN, SSD, FPN, RetinaNet and YOLOv3). Retrieved from: jonathan_hui/object-detection-speed-and-accuracy-comparison-faster-r-cnn-r-fcn-ssd-and-yolo-5425656ae359
- [2] Alfredo Canziani, Adam Paszke, Eugenio Culurciello (2017, April 14) An Analysis of Deep Neural Network Models for Practical Applications. Retrieved from: <https://arxiv.org/abs/1605.07678>
- [3] COCO – Common Objects in Context. Retrieved from: <http://cocodataset.org/#home>
- [4] Ian Goodfellow, Yoshua Bengio, Aaron Courville Deep Learning. A MIT Press Book //Ian Goodfellow, Yoshua Bengio, Aaron Courville . - MIT Press, 2016. - 716 p.
- [5] Brownlee, J. (2017, December 20) A Gentle Introduction to Transfer Learning for Deep Learning Retrieved from: <https://machinelearningmastery.com/transfer-learning-for-deep-learning/>
- [6] Prateek Joshi OpenCV with Python By Example//Prateek Joshi. - Packt Publishing, 2015. - 298p.
- [7] Siddharth, D. (2017, November 16) CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet and more... Retrieved from: <https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5>
- [8] PyTorch documentation Retrieved from: <https://pytorch.org/docs/stable/index.html>
- [9] CINIC-10 dataset. Retrieved from: <https://github.com/BayesWatch/cinic-10#data-loading>

Система Виявлення Атак на Локальну Комп'ютерну Мережу

Майк Максимов
кафедра математичного забезпечення
комп'ютерних систем
Одеський національний університет ім. І.І.Мечникова
Одеса, Україна
grendmike@gmail.com

Ірина Шпінарева
кафедра математичного забезпечення
комп'ютерних систем
Одеський національний університет ім. І.І.Мечникова
Одеса, Україна
ishpinareva@gmail.com

Attack Detection System on Local Computer Network

Mike Maksimov
Dept. of Mathematical Support of Computer Systems
Odessa I.I. Mechnikov National University
Odessa, Ukraine
grendmike@gmail.com

Irina Shpinareva
Dept. of Mathematical Support of Computer Systems
Odessa I.I. Mechnikov National University
Odessa, Ukraine
ishpinareva@gmail.com

Анотація — Інструменти виявлення атак мають дуже широкий спектр рішень, від застосування правил, отриманих досвідом, до використання методів машинного навчання, включаючи кілька методів біологічного натхнення. Задачі забезпечення безпеки корпоративних інформаційних систем традиційно вирішуються шляхом побудови системи інформаційної безпеки, де одним з найважливіших аспектів безпеки є безпека локальної комп'ютерної мережі. Станом на сьогоднішній день, кількість нових атак на корпоративні комп'ютерні мережі зростає кожного дня шляхом вторгнення у них за допомогою підключення до інтернету. У зв'язку з чим існуючі системи не завжди здатні виявити наявність атаки і вимагають застосування нових методів і підходів до виявлення атак.

Метою роботи є створення системи виявлення атак на локальну комп'ютерну мережу за допомогою пошуку аномальної поведінки мережі.

Abstract — Tools for detecting attacks have a very wide range of solutions, from application of rules to experience, to the use of machine learning techniques, including several methods of biological inspiration. The tasks of security of corporate information systems are traditionally solved by building an information security system where one of the most important aspects of security is the security of a local computer network. As of today, the number of new attacks on corporate computer networks is increasing every day by invading them by connecting to the Internet. In connection with which existing systems are not always able to detect the presence of an attack and require the application of new methods and approaches to detect attacks.

The purpose of the work is to create a system for detecting attacks on the local computer network by searching for an abnormal behavior of the network.

KDD, машинне навчання, TCP/IP.

Keywords — information security, intrusion detection, artificial neural networks, KDD, machine learning, TCP/IP.

І. ВСТУП

З самого початку передачі інформації за допомогою комп'ютерних мережевих ресурсів були розглянуті питання про загрози безпеки інформаційної системи за допомогою різних методів, найбільш часто використовуваних у системах виявлення вторгнень - IDS [1].

Аналітичні дані компаній [2], що спеціалізуються в сфері захисту інформації, таких як McAfee, Symantec, Trustware і Kaspersky Labs, показують, що протягом 2017 року стабільно зростала кількість інцидентів, пов'язаних зі спробами порушення безпеки інформаційних систем. Крім цього, спостерігається відносно стабільний ріст кількості нових зразків атакуючих дій. Згідно з міжнародним дослідженням ЕУ в даній області [3], у період 2016-2017 рр. кількість атак різних типів зросла у середньому на дев'ять відсотків. Ці дані показують, що існуючі IDS не можуть виявляти нові атаки і нові різновиди атак, та вимагають розробку нових методів для покращення результатів виявлення.

Станом на сьогоднішній день, існують два класи методів виявлення атак – метод зловживання та метод аномалій. Виявлення атак методами зловживання виконується за допомогою правил отриманих на основі досвіду або з використанням алгоритмів машинного навчання відповідно до наборів даних сигнатур, отриманих з мережевих сценаріїв зі звичайним трафіком і за допомогою ін'єкційних атак до одного і того ж сценарію. Цей клас методів використовують сучасні IDS доповнюючи його аналізом мережі у якій вони застосовуються.

Метою роботи є розробка та реалізація системи виявлення вторгнень яка може задовольнити наступні вимоги:

- система повинна виявляти погрозу застосовуючи методи аномалій;
- мати низьку ймовірність помилково-позитивних та помилково-негативних результатів;
- вміти виявляти нові різновидності однієї атаки.

II. ОГЛЯД МЕТОДІВ АНОМАЛІЙ

A. Аналіз методів аномалій

Методи аномалій засновані на використанні інформації про "нормальну" поведінку системи та її порівнянні з параметрами спостережуваної поведінки. Вони орієнтовані на побудову моделі штатного або нормального функціонування системи або користувача. До таких методів відносять:

1) *поведінкові методи*: Вейвлет-аналіз, статистичний аналіз, аналіз ентропії, спектральний аналіз, фронтальний аналіз та кластерний аналіз.

Переваги цих методів полягають у можливості визначити розподілені атаки, в тому числі і в часі, визначити взаємозв'язок між різними подіями, а також кореляція подій дозволяє визначити значущі події серед досліджуваних.

До недоліків відносять проблему чутливості методів, яка залежить від заданої величини відхилення, точності моделі штучної мережі та від функції подібності.

2) *методи машинного навчання*: Дерево рішень, Баєсова мережа, MAP-сплайни, алгоритми кластеризації та алгоритми регресії.

До переваги цих методів відносять: легкість реалізації, низькі вимоги до підготовки даних, малий час обробки великої кількості інформації. Використовують модель білого ящика, що допомагає перевірити модель з використанням статистичних тестів.

Недоліками методів є проблема навчання, що вимагає ретельного підбору коефіцієнтів та проблема побудови оптимальних моделей.

3) *методи обчислювального інтелекту*: Нейронні мережі, генетичні алгоритми, нечітка логіка, імунні системи, метод опорних векторів та ройові алгоритми.

Перевагами цих алгоритмів є можливість швидкої обробки великої розмірності даних. Також у паралельному режимі легкі для реалізації, добре працюють навіть на поганому наборі даних, здатні вирішувати нелінійні задачі.

Недоліками методів є необхідність як у позитивних, так і негативних прикладах. Вони вимагають багато пам'яті і процесорного часу, потребують виборі оптимальної функції ядра або активації.

Враховуючи ці особливості, було вирішено використовувати нейронні мережі типу – багатосаровий перцептрон Румельхарта та навчати їх алгоритмом зворотного поширення помилки.

III. БАЗА ДАНИХ ДЛЯ АНАЛІЗУ АТАК

NSL-KDD – це набір даних, запропонований Тавалае та др.[4]. Набір даних NSL-KDD є скороченою версією оригінального набору даних KDD 99. Він складається з тих самих функцій, що і KDD 99, 41 змінна та одного атрибуту класу. Атрибут класу має 21 значень, що потрапляють під чотири типи нападу: Probe attack, User to Root (U2R), Remote to Local(R2L) та атаки Denial of Service (DoS). Цей набір даних має атрибут дворядкового класу. Крім того, в ньому є достатня кількість навчальних та тестових прикладів, що робить можливим проведення експериментів.

A. Опис бази даних виявлення вторгнень

Записи з NSL-KDD

B. Атаки без контенту

У деяких атаках надсилається інформація вище 4-го рівня моделі OSI, щоб скористатися вразливістю програми, тому довжина корисного навантаження четвертого рівня, а саме поле сегмента даних (tcp.segment/tcp.payload) не порожнє. Такі типи нападів зазвичай називають "атаками на основі контенту" [5].

З іншого боку, є деякі атаки, яким не потрібно надсилати інформацію на рівень сеансу або вище, тобто поле сегмента даних пусте. Ці напади будуть називатися атаками без контенту.

Оскільки аналізується трафік локальної комп'ютерної мережі, отримати інформацію про контент атаки, а точніше, контент переданого пакета по мережі немає можливості, в слідстві чого, потрібно проаналізувати існуючий набір атак і вибрати характеристики, які не пов'язані із контентом атаки. З 41 доступних змінних 13 на основі контенту, це означає, що такі дані не є необхідними для виявлення атак не пов'язаних з контентом. Найважливішою особливістю при виявленні атак без контенту є прапор статусу з'єднання, тобто 4-а особливість в наборі даних. Найважливіші можливі стани прапору статусу наведені в Таблиці I.

ТАБЛИЦЯ I. Флаг Стану З'єднання

Стан	Значення
SF	Нормальне SYN/FIN завершення
REJ	З'єднання відхилено, початковий SYN викликав відповідь RST
S0	Стан 0: відправив SYN, відповіді нема
S1	Стан 1: з'єднання встановлено (обмінялись SYN), більше ніякої активності
S2	Стан 2: з'єднання встановлено, ініціатор закрив свою сторону
S3	Стан 3: з'єднання встановлено, відповідач закрив свою сторону
RSTO	Ініціатор оновлює з'єднання
RSTR	Відповідач оновлює з'єднання
OTH	Друге, цей стан не підтримується.

Історичні події можуть допомогти в процесі виявлення атак, бо деякі атаки приймають різні кроки, щоб бути успішними. Аналізуючи кроки, можна дізнатись, що в мережевому трафіку трапляється щось не так. Наприклад, в атаці з портів можливо знайти запити, які надходять до не запропонованої служби, викликаючи з'єднання з режимом REJ. Це не обов'язково означає, що присутня атака на порти portsweep, а можливо, хтось намагається отримати інформацію про послуги, що пропонуються в мережі.

Через характер нападів на комп'ютерну мережу були обрані усі характеристики, крім характеристик 10-22, які відносяться до атак з контентом. Описи кожної характеристики можна знайти в [6].

IV. МОДУЛІ СИСТЕМИ

Створена система здатна збирати інформацію з мережевого інтерфейсу. Архітектура системи представлена на Рис.1.

Модуль сенсору складається з застосування утиліти Tshark. На момент початку роботи системи запускається модуль сенсору і починається процес Tshark. Кожні 30 секунд система зберігає трафік мережі в обраній папці. При запуску цього модулю, запускається робота модуля аналізатора пакетів, який інтерпретує збережені пакети у

форматі XML, після чого збирає інформацію про пакети та передає аналізатору з'єднань.

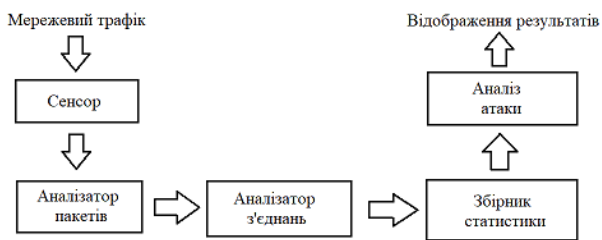


Рис. 1. Діаграма модулів системи

Модуль аналізатору з'єднань отримує пакети і в залежності від прапорів, встановлених в пакеті, вирішує чи встановити нове з'єднання, оновити інформацію вже існуючого або треба видалити з'єднання. Модуль аналізатору з'єднань відстежує стан з'єднання по описаним прапорам в таблиці I. А далі передає інформацію про з'єднання модулю збору статистики, а також, слідкує за часом з'єднання та виконує роботу двох секундного вікна. Якщо з'єднання було оновлено понад дві секунди тому, то з'єднання перевіряється на атаку і після чого видаляється зі списку з'єднань. Збірник статистики збирає статистичні дані протягом двох секунд з'єднання та утворює вектор з 28-ми змінними, який передається модулю аналізу атаки. Модуль аналізу атаки містить у собі каскад з двох нейронних мереж. Перша мережа аналізує на наявність атаки, друга мережа класифікує тип атаки.

Система працює у двох режимах: перший режим – режим налаштування, який навчає кожну нейронну мережу, другий режим – аналіз трафіку у реальному часі.

V. ТЕСТУВАННЯ

A. Метрика для оцінки якості

Для порівняння продуктивності і ефективності методів виявлення вторгнень в мережі використовуються такі метрики [7]: ймовірності істинно-позитивних (True Positive Rate, TPR) та помилково-позитивних результатів (False Positive Rate, FPR). FPR є ймовірністю отримати оповіщення у разі перевірки нормальної поведінки. До того, ймовірність помилково-негативного рішення (False Negative Rate, FNR) означає, що система не видає оповіщення у разі наявності зловмисної поведінки. Рівняння (1) та (2) відображують FPR та FNR:

$$FPR = \frac{\text{number of false positive}}{\text{number of negative}} \quad (1)$$

$$FNR = \frac{\text{number of false negative}}{\text{number of positive}} \quad (2)$$

Звідси, ймовірність TPR та істинно-негативних результатів (True Negative Rate, TNR) дорівнюють:

$$TPR = 1 - FNR \text{ та } TNR = 1 - FPR \quad (3)$$

Також, до цих показників можна додати ймовірність правильної класифікації (True Rate, TR) і ймовірність помилкової класифікації (False Rate, FR).

$$TR = \frac{\text{number of true positive} + \text{number of true negative}}{\text{number of positives} + \text{number of negatives}} \quad (4)$$

$$FR = 1 - TR.$$

Рівняння (5) відображує міру чутливості (sensitivity), яка визначається як частка нормальної поведінки:

$$Sensitivity = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}} \quad (5)$$

Але, ця міра недостатньо змістовна, оскільки вона може бути тривіально досягнута шляхом класифікації поведінки як зловмисної.

Ще одна метрика, міра специфіки (specificity). Це частка справжніх негативів всієї розглянутої негативної поведінки:

$$Specificity = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}} \quad (6)$$

При класифікації всього трафіку як нормального, міра специфіки досягається повністю.

F-міра є показником, що поєднає міри чутливості та специфіки:

$$F - \text{measure} = \frac{2 \times \text{sensitivity} \times \text{specificity}}{\text{sensitivity} + \text{specificity}} \quad (7)$$

B. Тестування різних нейронних мереж

Тестування і вибір розміру, функції активації та показника навчання проводились у ОС Windows® 10-64 з процесором Intel Core i7 2600k (3.4 GHz), 12,0 Гб ОЗУ. Реалізація нейронних мереж була створена за допомогою Microsoft CNTK і реалізована на мові C#. Система була навчена на 133504 векторах і протестована на 15040. Сама вибірка векторів мала 95% варіацій і кожен вектор унікальний, тобто вибірка навчальна і вибірка для тестування не мають однакових векторів.

Каскад складається з двох нейронних мереж. Перша нейронна мережа повинна встановлювати наявність атаки, а друга повинна встановлювати клас атаки. Тобто, перша нейронна мережа на виході у якості результату надає вектор з двох значень – є атака чи ні, друга нейронна мережа у якості відповіді надає вектор з п'ятьма значеннями – Probe, R2L, U2R, DoS та «не є атакою». Перші чотири значення відповідають класам атак, а п'яте значення застосовується для подвійної перевірки, якщо перша нейронна мережа помилково вирішила, що зв'язок небезпечний.

Для вибору структури нейронних мереж були протестовані п'ять функцій активації з різними коефіцієнтами навчання та кількості шарів. Таблиця II відповідає трьох шаровим нейронним мережам з коефіцієнтом навчання 0.02. Вхідний шар відповідає 28 змінним, другий шар містить кількість нейронів кратних трьом(84), або п'яти(140) та останній шар містить два, якщо це перша нейронна мережа та п'ять, якщо друга. Кількість нейронів у другому шарі було вибрано емпіричним шляхом у ході тестування. Кількість шарів нейронних мереж дорівнює трьом, при спробі збільшення їх кількості, показники точності значно падали.

Перша нейронна мережа показала найкращий результат за показниками Tr, Fr, TPR, FNR, Recall та F-measure з функцією активації ReLU та середнім шаром розміром 84 (табл. II). Ця нейронна мережа була обрана для початку каскаду.

Друга нейронна мережа показала більш розподілений результат. Нейронна мережа з функцією активації ReLU та з середнім шаром розміром 140 видає кращий результат за показниками TPR, FPR, Recall та F-measure, а нейронна мережа з середнім шаром розміром 84 видає кращий результат за показниками Tr та Fr. Оскільки головними для нас є показники Tr та Fr, у якості другої нейронної мережі була обрана мережа з функцією активації ReLU, середнім шаром розміром 84 та коефіцієнтом навчання $\eta=0.02$ (табл.III).

Також кожна нейронна мережа була протестована з коефіцієнтом навчання 0.04 або 0.08 та з більшою кількістю шарів. При використанні трьох шарів та коефіцієнту навчання $\eta=0.04$ показники Tr та Fr мають

гірший результат у середньому на 0.3 відсотка, а при збільшенні коефіцієнту навчання цей показник значно зростає, проте, при збільшенні кількості шарів та використанні коефіцієнту навчання 0.02, показники Tr та Fr відрізняються вже на 10,3 відсотки у середньому.

Таблиця IV демонструє порівняння між обраними нейронними мережами першого і другого типу окремо, та створений з них каскад. ReLU_1 – нейронна мережа першого типу, відповідно ReLU_2 – нейронна мережа другого типу. Як можна бачити з таблиці, каскад з цих мереж показав кращий результат за показниками Tr, Fr, TNR, та Precision. Це показує, що створений каскад краще працює, а також має нижчу ймовірність прийняти безпечний трафік за небезпечний.

ТАБЛИЦЯ II. Оцінка Якості Першої Нейронної Мережі (28/84/2) З Різними Функціями Активації та $\eta=0,02$

	Tr	Fr	TPR	TNR	FPR	FNR	Recall	Precision	F-measure
Tanh	96.85	3.15	93.51	99.78	6.49	0.22	93.51	99.78	96.54
ReLU	98.22	1.78	97.45	98.89	2.55	1.11	97.45	98.89	98.17
Sigmoid	97.00	3.00	94.63	99.08	5.37	0.92	94.63	99.08	96.80
ELU	96.66	3.34	93.02	99.85	6.98	0.15	93.02	99.85	96.32
SELU	96.29	3.71	92.08	99.98	7.92	0.02	92.08	99.98	95.87

ТАБЛИЦЯ III. Оцінка Якості Другої Нейронної Мережі (28/84/5) З Різними Функціями Активації та $\eta=0,02$

	Tr	Fr	TPR	TNR	FPR	FNR	Recall	Precision	F-measure
Tanh	97.73	2.27	95.91	99.38	4.09	0.62	95.91	99.38	97.61
ReLU	97.83	2.17	96.64	98.96	3.36	1.04	96.64	98.96	97.79
Sigmoid	97.29	2.71	94.70	99.61	5.30	0.39	94.70	99.61	97.10
ELU	97.41	2.59	95.64	99.16	4.36	0.84	95.64	99.16	97.37
SELU	97.46	2.54	95.29	99.46	4.71	0.54	95.29	99.46	97.33

ТАБЛИЦЯ IV. Порівняння Якості Каскаду

	Tr	Fr	TPR	TNR	FPR	FNR	Recall	Precision	F-measure
ReLU_1	98.22	1.78	97.45	98.89	2.55	1.11	97.45	98.89	98.17
ReLU_2	97.83	2.17	96.64	98.96	3.36	1.04	96.64	98.96	97.79
Cascade	98.46	1.54	96.48	99.34	3.52	0.66	96.48	99.34	97.89

ВИСНОВКИ

Метою цієї роботи було створити систему аналізу інтернет трафіку для виявлення атаки на локальну комп'ютерну мережу. В роботі для виявлення інформативних ознак атак були розглянуті нейронні мережі прямого поширення з різними функціями активації та кількості шарів, на підставі яких був створен каскад нейронних мереж у якості нового засобу аналізу даних трафіку. Запропонований каскад показав кращий результат – 98.46 відсотків, ніж використання кожної нейронної мережі окремо. Була створена система аналізу мережевого трафіку для виявлення атак. Така точність виявлення атак, а також гнучкість у налаштуванні під потреби конкретної системи демонструє перспективність нейромереж в порівняно із заздалегідь запрограмованими алгоритмами виявлення вторгнень.

REFERENCES ЛІТЕРАТУРА

- [1] Nikishova A., Churilina A. Distributed intrusion into information system of enterprise detection Izvestiya UFU, Technical science. 2013. №12 (149).
- [2] Information Security Threats, www.anti-malware.ru/threats/information-security-threats.
- [3] Garuba, M.; Chunmei Liu; Fraités, D.; "Intrusion Techniques: Comparative Study of Network Intrusion Detection Systems," Information Technology: New Generations, 2008. ITNG 2008. Fifth International Conference on , vol., no., pp.592-598, 7-9 April 2008 doi: 10.1109/ITNG.2008.231
- [4] Tavallaee, M.; Bagheri, E.; Wei Lu; and Ghorbani, A. (2009). A detailed analysis of the KDD CUP 99 data set. Proceedings of the 2009 IEEE Symposium on Computational Intelligence in Security and Defense Applications (CISDA 2009), 1-6. /
- [5] Pasupulati, A.; Coit, j.; Levitt, K; Wu, S.F.; Li, S.H.; Kuo, J.C.; Fan, KP.; "Buttercup: on network-based detection of polymorphic buffer overflow vulnerabilities," Network Operations and Management Symposium, 2004. NOMS 2004. IEEE/IFIP , vol.1, no., pp.235-248 Vol.1, 23-23 April 2004
- [6] Dhanabal, L. and Dr. S. P. Shantharajah. "A Study on NSL-KDD Dataset for Intrusion Detection System Based on Classification Algorithms." (2015).
- [7] Holz T. 13 security measurements and metrics for networks, Dependability Metrics, 2008, pp. 157–165. DOI: 10.1007/978-3-540-68947-8_13

Параметричний Синтез Моделі Багатокритеріального Оцінювання на Основі Нейронних Мереж

Януш Вероніка
кафедра системотехніки
Харківський національний університет
радіоелектроніки
Харків, Україна
nikayanush16@gmail.com

Володимир Безкоровайний
кафедра системотехніки
Харківський національний університет
радіоелектроніки
Харків, Україна
vladimir.beskorovainyi@nure.ua

Parametric Synthesis of Multicriteria Estimation Models based on Neural Networks

Veronika Yanush
Department of System Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
nikayanush16@gmail.com

Volodymyr Beskorovainyi
Department of System Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
vladimir.beskorovainyi@nure.ua

Анотація—Розглядається підхід до розв'язання задачі параметричного синтезу моделей багатокритеріального оцінювання, заснований на застосуванні штучних нейронних мереж. Запропоновано процедуру визначення вагових коефіцієнтів часткових критеріїв з використанням технології компараторної ідентифікації.

Abstract—An approach to solving the task of parametric synthesis of multi-factor estimation models based on the application of artificial neural networks is considered. The procedure of determination of weight coefficients of partial criteria using the technology of comparator identification is proposed.

Keywords—the task of multicriteria assessment, model, utility function, parametric synthesis, neural networks.

Ключові слова—задача багатокритеріального оцінювання, модель, функція корисності, параметричний синтез, нейронні мережі.

I. ВСТУП

Однією з актуальних проблем теорії прийняття рішень є формалізація процесів оцінювання альтернатив. Конструктивне вирішення цієї проблеми потребує синтезу моделі для формування скалярних оцінок якості рішень за множиною часткових критеріїв [1–2]. Практична значимість таких моделей обумовлена необхідністю підвищення ступеня ефективності процедур вибору рішень в системах планування, управління, проектування

соціально-економічних і технічних об'єктів, вирішення задач управління поведінкою соціальних груп [3]. При цьому як вхідні дані використовується інформація щодо цінності окремих формалізованих властивостей (часткових критеріїв) і їх значень. Це обумовлює актуальність науково-прикладних задач синтезу ефективних за показниками адекватності та економічності математичних моделей для кількісного оцінювання та вибору рішень за множиною показників для автоматизованих систем проектування й управління [4].

II. АНАЛІЗ СУЧАСНОГО СТАНУ ПРОБЛЕМИ

В основі сучасних моделей прийняття рішень лежить парадигма максимізації корисності. Вважається, що особа, яка приймає рішення (ОПР), при виборі варіантів з множини допустимих $x \in X$ приписує їм деяку корисність (цінність) $P(x)$, значення якої і визначають її вибір $x^o \in X$ [4]:

$$\forall x, y \in X : x \sim y \leftrightarrow P(x) = P(y); x \succ y \leftrightarrow P(x) > P(y); \\ x \sqsupseteq y \leftrightarrow P(x) \geq P(y); x^o = \arg \max_{x \in X} P(x).$$

При цьому розрізняють підходи оцінювання, побудовані на основі ординалістичної (порядкової) і кардиналістичної (кількісної) теорії корисності. В рамках обох теорій характеристики альтернатив x із допустимої

множини X можуть бути виміряні або приведені шляхом експертного оцінювання до чисельного вигляду:

$$P(x) = F[\lambda_i, k_i(x)], \quad (1)$$

де $k_i(x)$, $i = \overline{1, m}$ - значення часткових критеріїв (характеристик), однозначно встановлені для кожного $x \in X$; λ_i - вагові коефіцієнти часткових критеріїв, що враховують різну їхню соціальну значимість (вагу) в узагальненій оцінці $P(x)$.

Для практичного використання моделей (1) необхідно визначити їх структуру і параметри. Такі моделі синтезуються на основі адитивних, мультиплікативних або змішаних функцій загальної корисності. Найбільш універсальною вважається функція загальної корисності $P(x)$, побудована на основі поліному Колмогорова-Габора [4]:

$$P(x) = \sum_{i=1}^m \lambda_i \cdot \xi_i(x) + \sum_{i=1}^m \sum_{j=1}^m \lambda_{ij} \cdot \xi_i(x) \cdot \xi_j(x) + \dots, \quad (2)$$

де m - кількість часткових критеріїв; λ_i, λ_{ij} - вагові коефіцієнти часткових критеріїв $k_i(x)$ та їх добутоків $\lambda_i \geq 0, \lambda_{ij} \geq 0, i = \overline{1, m}$; $\xi_i(x)$ - функція корисності часткового критерію $k_i(x)$, $i = \overline{1, m}$;

Однією з найбільш поширених функцій корисності часткових критеріїв є функція:

$$\xi_i(x) = \left(\frac{k_i(x) - k_i^-}{k_i^+ - k_i^-} \right)^{\alpha_i}, \quad (3)$$

де k_i^+, k_i^- - відповідно найкраще та найгірше значення критерію; α_i - параметр, що визначає вид залежності (при $\alpha_i = 1$ реалізується лінійна, при $0 < \alpha_i < 1$ - опукла, при $\alpha_i > 1$ - увігнута залежності).

Шляхом введення позначень:

$$\begin{aligned} \xi_1(x) \cdot \xi_1(x) &= \xi_{m+1}(x), \quad \lambda_{1,1} = \lambda_{m+1}, \\ \xi_1(x) \cdot \xi_2(x) &= \xi_{m+2}(x), \quad \lambda_{1,2} = \lambda_{m+2}, \dots, \end{aligned} \quad (4)$$

модель (2) може бути подана в традиційній адитивній формі:

$$P(q, x) = \sum_{i=1}^N \lambda_i \xi_i(x), \quad (5)$$

де N - загальна кількість доданків в співвідношеннях (4).

Задача визначення вектора параметрів $\lambda = [\lambda_i, \lambda_{i,j}]$, $i = \overline{1, m}, j = \overline{1, m}$ з області допустимих є некоректною за Адамаром. Її регуляризація може бути здійснена, зокрема

для задач з встановленими параметрами функцій корисності часткових критеріїв (3) шляхом пошуку чебишовської точки [4]. Це дозволяло звести розв'язувану задачу до задачі лінійного програмування [5, 6].

Все більш широке застосування в сучасних системах підтримки прийняття рішень знаходять інтелектуальні технології на основі штучних нейронних мереж. До переваг штучних нейронних мереж перед традиційними методами розв'язання подібних задач відносять: здатність до навчання, що дозволяє знаходити невідомі взаємозв'язки між вхідними та вихідними даними; стійкість до шумів у вхідних даних; можливість реалізації шляхом паралельно виконуваних операцій [7].

III. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Сучасною альтернативою експертному оцінюванню вагових коефіцієнтів часткових критеріїв λ_i, λ_{ij} , $i = \overline{1, m}$ є технологія компараторної ідентифікації [4]. Вона передбачає формування ОПР бінарного відношення на множині альтернатив $R(X) \subseteq X \times X = \{ \langle x, y \rangle, x, y \in X \}$. Прикладом може слугувати бінарне відношення строгої переваги

$$R_S(X) = \{ \langle x^o, y \rangle, x^o, y \in X : x^o \succ y, \forall y \in X \}, \quad (6)$$

яке може бути встановлене шляхом вибору ОПР лише кращої за множиною часткових критеріїв альтернативи $x^o \in X$. Для відношення (6) отримуємо системи лінійних нерівностей і нормуючих умов:

$$\begin{cases} \eta_j(\lambda) \equiv \sum_{i=1}^N \lambda_i \xi_i(x^o) > \sum_{i=1}^N \lambda_i \xi_i(y); \\ j = \overline{1, n_S}, \langle x^o, y \rangle \in R_S(X); \\ \eta_{m_S+1}(\lambda) \equiv \sum_{i=1}^N \lambda_i = 1, \lambda_i \geq 0, i = \overline{1, N}, \end{cases} \quad (7)$$

де $\lambda = [\lambda_i, \lambda_{i,j}]$, $i = \overline{1, m}, j = \overline{1, m}$ - шуканий вектор параметрів функції (2), n_S - потужність бінарного відношення (6), $n_S = \text{Card } R_S(X)$.

Задача пошуку вектора параметрів λ , що задовольняє системі обмежень (7) є некоректною. Для її регуляризації найкращим будемо вважати такий вектор вагових коефіцієнтів λ , що максимізує мінімальну відстань $\Delta P_j(\lambda, x^o, y) = P(\lambda, x^o) - P(\lambda, y)$, $\langle x^o, y \rangle \in R_S(X)$ між рішеннями у бінарному відношенні строгої переваги (6).

Для розв'язання задачі використаємо апарат штучних нейронних мереж (рис. 1). Навчання мережі будемо здійснювати із застосуванням алгоритму зворотного розповсюдження помилки.

Алгоритм навчання реалізує градієнтний метод мінімізації опуклого (звичайного квадратичного) функціоналу помилки [7]. Для забезпечення коректного

навчання та функціонування мережі використаємо сигмоїдальну функцію активації, що є монотонно зростаючою та має відмінні від нуля похідні на всій області визначення.

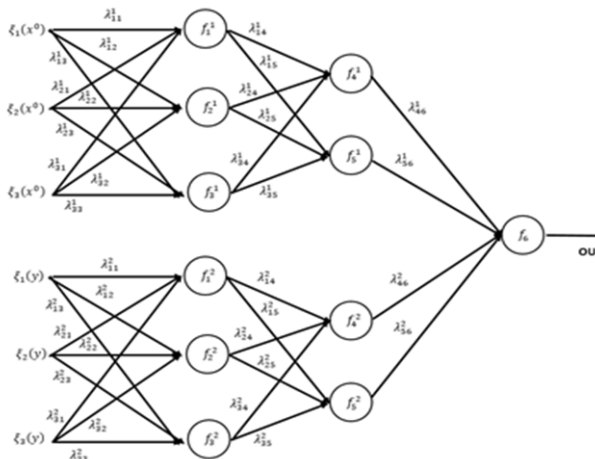


Рис. 1. Багатошарова нейронна мережа для трехчасткових критеріїв: f_j^i – функції активації нейронів.

Процес навчання полягатиме у подачі на вхід мережі інформації щодо порівнюваних альтернатив у вигляді значень функцій корисності їх часткових критеріїв $[\xi_1(x^o), \xi_2(x^o), \dots, \xi_m(x^o)]$ та $[\xi_1(y), \xi_2(y), \dots, \xi_m(y)]$, вивченні реакції на них мережі і корекції відповідно до реакції вагових параметрів. Перед початком навчання координатам вагового вектора присвоюються випадкові значення з діапазону від 0 до 1. Істотною умовою застосування методу є диференційованість активаційної функції. Як функція активації використовується сигмоїд.

Для обмеження простору пошуку у процесі навчання ставиться завдання мінімізації цільової функції помилки нейронної мережі, яка знаходиться за методом найменших квадратів:

$$\varepsilon = \frac{1}{2} \sum_k^n (t_k^n - o_k^n)^2, \quad (8)$$

де o_k^n - значення k -го виходу мережі; t_k^n - цільове значення k -го виходу; n - номер останнього шару.

Для оновлення значень вектора вагових коефіцієнтів використовується функція методу зворотного розповсюдження помилки, за допомогою якої обчислюється величина $\Delta\lambda$, на яку потрібно змінювати їх значення в поточному циклі $\lambda := \lambda + \Delta\lambda$ [7]. В результаті отримуємо оцінки параметрів вектора вагових коефіцієнтів часткових критеріїв, що задовольняють умовам (7) та максимізують мінімальну різницю $\Delta P_j(\lambda, x^o, y) = P(\lambda, x^o) - P(\lambda, y)$, $\langle x^o, y \rangle \in R_S(X)$ між

альтернативами у бінарному відношенні строгої переваги (6), яке сформоване особою, що приймає рішення, експертом або групою експертів.

ВИСНОВКИ

Одержав подальший розвиток підхід компараторної ідентифікації в частині його застосування до розв'язання задачі параметричного синтезу адитивних моделей багатофакторного оцінювання з використанням штучних нейронних мереж. Використання штучних нейронних мереж дозволяє інтегрувати сучасні методи штучного інтелекту в автоматизовані системи підтримки прийняття рішень. Отримані результати можуть бути використані в у системах планування, управління, проектування соціально-економічних і організаційно-технічних об'єктів, вирішення завдань управління поведінкою соціальних груп. Напрямами подальших досліджень можуть бути реалізація інтелектуальних технологій для параметричного синтезу більш ефективних за показником «точність-складність» функцій корисності часткових критеріїв [8–11].

REFERENCES ЛІТЕРАТУРА

1. S. Greco, M. Ehrgott, J.R. Figueira, Multiple Criteria Decision Analysis – State of the Art Surveys, New York: Springer, 1346 p., 2016.
2. М.М. Маляр, Моделі і методи багатокритеріального обмежено-раціонального вибору: монографія, Ужгород: РА АУТДОР-ШАРК, 222 с., 2016.
3. В.В. Крючковский, Э.Г. Петров, Н.А. Соколова, В.Е. Ходаков, Введение в нормативную теорию принятия решений, Херсон: Гринь Д.С., 284 с., 2013.
4. О.А. Овезгельдыев, Э.Г. Петров, К.Э. Петров, Синтез и идентификация моделей многофакторного оценивания и оптимизации, К.: Наукова думка, 161 с., 2002.
5. В.В. Бескоровайный, И.В. Трофименко, “Структурно-параметрична ідентифікація моделей багатофакторного оцінювання”, *Системи озброєння і військова техніка*, № 3 (7), с. 56–59, 2006.
6. В.В. Бескоровайный, И.В. Трофименко, “Метод решения общей задачи компаративной идентификации моделей многофакторного оценивания”, *Бионика интеллекта*, № 2 (65), с. 3–7, 2006.
7. О.Г. Руденко Е.В. Бодянский, Искусственные нейронные сети, Харьков: ООО «Компания СМИТ», 408 с., 2005.
8. Э.Г. Петров, В.В. Бескоровайный, В.П. Пискалова, “Формирование функций полезности частных критериев в задачах многокритериального оценивания”, *Радиоэлектроника и информатика*, №1, с. 71-73, 1997.
9. В.В. Бескоровайный, Е.В. Соболева, “Идентификация частной полезности многофакторных альтернатив с помощью S-образных функций”, *Бионика интеллекта*, №1, с. 50–54, 2010.
10. V. Beskorovainyi, H. Berezovskyi, “Identification of preferences in decision support systems”, *ECONTECHMOD*, vol. 06, №4, p. 15-20, 2017.
11. V. Beskorovainyi, H. Berezovskyi, “Parametric synthesis of models for multicriteria estimation of technological systems”, *Сучасний стан наукових досліджень та технологій в промисловості*, №2 (2), с. 5–11, 2017.

Big Data Technologies in Big Political Game

Kovalchuk Olha Yaroslavivna
Department of Applied Mathematics
Ternopil National Economic University
Ternopil, Ukraine
olhakov@gmail.com

Masonkova Mariia Mykhaylivna
Department of international relations,
public communications and regional studios
Ternopil National Economic University
Ternopil, Ukraine
masyonkova@gmail.com

Технології Big Data у Великій Політичній Грі

Ковальчук Ольга Ярославівна
к. ф.-м. н., доцент кафедри прикладної математики
Тернопільський національний економічний університет
olhakov@gmail.com

Масьонкова Марія Михайлівна
Магістрантка спеціальності «Міжнародні відносини,
суспільні комунікації та регіональні студії»
Тернопільський національний економічний університет
Тернопіль, Україна
masyonkova@gmail.com

Abstract— This work examines possibilities of application of big data information technologies for realization of effective political digital campaigns on the basis of analysis of information from virtual networks and hidden influence on acceptance of irresponsible decisions by publicity.

Keywords — big data; social networks; psychological impact; political game.

I. INTRODUCTION

In wide sense modern society can be defined as clever (smart), or servitive, society in that digital technologies are widely used on all management levels. The internet and social networks made the world more open and life more comfortable. However, there is another one, dark side of the newest technologies. Increasing of amount of the Internet users and modern methods of analysis of users behavior in the network gives an opportunity to influence on society with stunning efficiency.

Big data technologies acquire the special actuality, technologies that cardinaly change ways of information analyses and methods of making decision in all spheres of human activity, in particular in economic and political. Today it is possible to talk bravely not only about a digital economy but also about digital politics.

II. BIG DATA AS NEW FORM OF DIGITAL POLITICS

Big data is a term that is used for denotation of data analysis in prognostic analytics, intellectual data analysis (data mining), in customer relationships management(CRM - Customer Relationship Management) [1].

In wide sense "large data" is the socio-economic phenomenon related to appearance of technological possibilities to analyse the enormous arrays of data. The concept of big data envisages work with data of large amounts and various composition, that are got from different sources, in order to increase competitiveness and efficiency of acceptance of administrative decisions.

Strong people of this world actively form the so-called servitive society, in that they offer comfort life, attractive loans, unreal possibility to get all at once. Big data technologies open for large business and politics new instruments of secret influence on an audience, with the purpose of management of their behavior, acts and even choice. The use of the newest targeting methods assists forming of consumers psychological dependence. Primary purpose - to retain in the power a man that will not be able to provide the further existence independently. Such ideology of modern political forces and business is sent to the receipt of maximal amount of supporters and reckless superprofits [2].

Unlimited access to the array of digital tracks, consists of consumers advantages and necessities, what they left everyday in a network, carrying out searching queries, bank payments, using GPS- and passing psychological tests that contain many personal questions, opens unlimited possibilities to influence and manipulate them. Well-known world political leaders also use internet-marketing technologies based on big data .As experience testifies, victory is got by that, who uses the newest methods of analytics and information technologies.

Growing popularity of social networks is not just a change of method of communication, it is a change of doing business method, way of life in society. It is a change of the rhythm of modern life.

Connection to the network opens unlimited access to personal users data. For example, the motion sensor in smartphone shows, it is a smartphone swung by a hand, how far it moves (these actions correlate with emotional instability). A smartphone is an original free questionnaire that is willy-nilly filled by its proprietor. After its help it is possible not only to form a psychological portrait from the obtained data but also to choose necessary among these portraits. It is the original searching system on people, that is capable to find out trustful ones, apt to the spontaneous actions, criminal intents and others like that.

At the use of big data technologies for social researches a basic problem is collection of large arrays of the personal information. However, with appearance of the Internet, especially Facebook, Instagram, VKontakte, Odnoklassniki and other social networks, the capture of data takes place automatically - users do information about them popular themselves.

Facebook is said to be the symbol of new generation, epoch without physical borders. Most public conversations passed to on-line services of social medias and other digital platforms. Qualificatory role in forming of public opinion is played by forums and chats of the Internet. However, this powerful searching system on people can be used also in criminal aims. Big data technologies for social networks can carry a threat to prosperity, freedom and even life of people.

So once again Facebook became the source of psychological statistical data, by means of what Cambridge Analytica realized manipulation technologies with the aim of influence on the results of USA presidential elections in 2016 and referendum about the exit of Great Britain from EU [3].

Technically companies that grant political services use software of making decision strategy, that executes statistical processing of large data, got or renewed by means of different facilities for the digital profiling of electorat. The aim of these instruments application for end-user, for example, electoral campaign team, is forming of the most exact and simple elements, that can be effectively used for providing of communicative charts with electors. On the basis of individual data analysis computer programs create personalizing or reports, having special purpose - electronic reports, phone calls or personal visits. Reverse information that is used for further targeting has also an important value.

Thus, knowledge, got by campaigns team, give an opportunity to maximally adapt electoral discourse to the necessities, beliefs, age and social status of respondents with the aim of the hidden target influence.

A new form of the digital mediated politics is the key component of fourth industrial revolution : the Internet is already used for establishment of agenda from an ascending level, grant to the citizens of possibilities to express the opinions in network public sphere and organize collective actions. In particular, social medias changed character of political agitation and will continue to play an important role in future elections and political campaigns in the whole world.

However this technology also can be a platform for conflicts and ill-intentioned agitation of right populists that does not have functional possibilities for healthy democratic discourse, while modern control system is vulnerable to emotional splashes and populist motions that arise up in social networks. Brexit witnessed that this speed technology is open for all and can be used for affecting on public opinion and referendums or elections results.

Big politics is one actor grandstanding, not only races between parties. Boats with plenty of supporters that distribute false news in the closed social networks groups are ideal channels for disinformation.

If before nobody could in earnest perceive information that a post on Facebook or report in Twitter can entail oscillation on elections, today the propagandist strategies aims in social networks are sent to the real problems, such as immigration, economic crisis, terrorism and social inequality, in order to hiddenly influence on ideas and behavior of millions of people and induce to the choice of certain political procedure or point of view.

Big data technologies, that assisted Donald Trump's victory on USA presidential elections in 2016, also can be used during realization of electoral campaigns in other countries.

In recent year big data analysis technologies became popular or quite open. The hardware and softwear support can be bought very cheaply. A competitive edge is experience and knowledge, that is necessary for work with big data. Almost all services, accessible today in the Internet, are based on the big data analysis. Exactly today the algorithms of machine studies, worked out in 1960-70th of XXth century, are used for treatment of large data arrays. New big data technologies find application in the most various industries - from searching services, translations and navigation to the use of the conformities, educed in internet-environment, retail off-line sales and targeting advertisement.

Experts forecast, that in a current decade this area will make progress, generating new questions, in particular how to solve the problems of personal data safety, who does is the subject of these data disposing, how possible separate methods of obtained data application can be and others like that.

CONCLUSIONS

In the world, guided by data, the problems of data economy (from the mass attempts of the unauthorized use of data to intentional manipulation by on-line platforms) grew as never before. Privacy, intimacy, arcana, personal space - all these concepts lost sense in the epoch of digital technologies.

Modern servitive society can by easily managed by means of information. Data are constantly gathered, accumulated and kept, and that, who has an access to these data, owns the future. It is a group of multimillionaires that use big data technologies for realization of global shadow operations against democracy.

An only correct decision to prevent the threats of the new digital world is protection of data and use and development of effective strategies of information enciphering and safety.

Big data technologies and artificial intelligence is a way to the new innovations, extended access to possibilities and solusion of the most actual problems of society.

REFERENCES

- [1] Dehaye P.-O. The Data That Turned the World Upside Down. Public Policy Program. – Stanford, California: Stanford University, 2017. Retrieved from: <https://publicpolicy.stanford.edu>.
- [2] Bail C.A. Taming Big Data. 2017, 46(2), pp. 189-217. Retrieved from: <http://smr.sagepub.com>.
- [3] Weforum. – Retrieved from: <https://www.weforum.org>.

Digital Watermark Resistant to JPEG

Mariia Kozina

*Department of Informatics and
Information Security Management
Odessa National Polytechnic
University
Odessa, Ukraine
mashaK1989@rambler.ru*

Alla Kobozeva

*Department of Informatics and
Information Security Management
Odessa National Polytechnic
University
Odessa, Ukraine
alla_kobozeva@ukr.net*

Alexander Kozin

*Department of Information
Technologies
National University "Odessa Academy
of Law"
Odessa, Ukraine
kozindre@rambler.ru*

Natalia Loginova

*Department of Information Technologies
National University "Odessa Academy of Law
Odessa, Ukraine
kozindre@rambler.ru*

Elena Trofimenko

*Department of Information Technologies
National University "Odessa Academy of Law
Odessa, Ukraine
kozindre@rambler.ru*

Abstract—The popularity of compression attacks in steganography today is associated with the widespread embedding of compression algorithms for lossy graphics files for storing and forwarding information that does not attract attention to these destination files, not only with the use of hidden but also open data transmission. It was developed the new effective stegoalgorithm that based on the SVD of the matrix container. Efficiency is confirmed by computational results PSNR, MSE, correlation coefficient, computational complexity.

Keywords—*stegoalgorithm, watermark, SVD, lossy format, JPEG.*

I. INTRODUCTION

The standing of information technology has given impetus to the development of new modern methods aimed to organize the security of data transmission by telecommunication channels, which for today has become widely used for paid purposes. A promising direction in the organization of information security in advanced systems and networks is the digital steganography.

The popularity of compression attacks in steganography today is associated with the widespread embedding of compression algorithms for lossy graphics files for storing and forwarding information that does not attract attention to these destination files, not only with the use of hidden but also open data transmission. For example, when spreading of information in public networks, protecting information and ensuring its reliability are unresolved tasks today. Because of this, to modern stegoalgorithms and methods a mandatory requirement of resistance to violations, especially JPEG-compression to exposure in the communication channel is appointed.

II. MAIN PART

A. Aim and tasks of the research

The aim of this work is the development of new stegoalgorithms based on the singular value decomposition (SVD) of the matrix container and a comparative analysis based on a practical experiment to determine the reliability of the perception and stability of the stegoalgorithms in terms using compress (JPEG-compress).

Methods of data hiding in the image spatial area are not resistant to most of the known types of violations, especially, to lossy compression format. So, the biggest interest in the

area of digital imaging is the methods of embedding information in image where there is loss of quality (popular JPEG format). The methods used to hide the cover information to frequency domain are more resistant to various possible external influences on the image container.

Tasks according to the described aim:

- to define the effectiveness of the developed stegoalgorithm by visual assessment of quality;
- to define the effectiveness of the developed stegoalgorithm by degree of reliability the perception of stegomessage by calculation MSE, PSNR;
- to define the effectiveness of the developed stegoalgorithm by degree of reliability of the perception of stegomessage after compression with varying degrees of compression, which does not lead to a breach of reliability of perception due to JPEG-compression by calculation MSE, PSNR;
- to define the effectiveness of the developed stegoalgorithm by calculation of the Correlation Coefficient of a stegomessage by the proposed stegoalgorithms without JPEG-compression (Q);
- to define the effectiveness of the developed stegoalgorithm by calculation computational complexity by the proposed stegoalgorithms.

B. Analise of litreture

In [2], sufficient conditions for the stability of quilted methods and stegoalgorithms for compression are presented, including significant coefficients that do not depend on which domain of the digital image of the container (the spatial domain or frequency domain) is used for the integration of additional information.

Obtaining these conditions was the impetus for the further development of a general approach to the analysis of the state and technology of the functioning of information systems, based on matrix analysis and perturbation theory, and its adaptation for solving the problems of steganography [3].

The new common approach for integrity checking of digital images is developed [4]. The new features of formal parameters defining image are revealed, theoretically grounded and practically tested. The characteristics of the mutual arrangement of left and right singular vectors corresponding to the largest singular value of the image's matrix (block of matrix) and the vector composed of singular numbers is obtained.

The results of scientific works are presented in [2,3] are the basis for the theoretical developments of the new stegomethod. There is proposed a stegoalgorithm, relying on the developed stegomethod [3,5], the effectiveness will be computing and analyzing in this work.

C. Theoretical part

Let the A is $M \times N$ matrix of the container. The general steps of the proposed steganographic algorithm consists of three ones.

1) First, we pre-break blue matrix of the image matrix A into disjoint blocks f of a size 8×8 . It is the standard size of a block.

2) Further for each block we get singular value decomposition:

$$f = U \Sigma V,$$

where U, V are orthogonal matrices of size $n \times n$, columns u_1, \dots, u_8 of the matrix U - left-handed eigenvectors,

lexicographically positive (the columns v_1, \dots, v_8 of the matrix V are called the right-handed eigenvectors of the matrix f);

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_8), \sigma_1 \geq \dots \geq \sigma_8 \geq 0 - \text{eigenvalues.}$$

3) The maximum singular numbers get perturbation in the form of cover information, depending on the value of the bit of cover information, which does not lead to a breach of the reliability the perception of the formed stegomessage.

The process of compression of the digital image will make a certain way the perturbation of the eigenvalues of block matrices. Since the sets of eigenvalues of matrices of the digital image in the spatial and frequency domains coincide, any perturbations of the eigenvalues are identical for the of block matrices of digital image in the spatial and frequency domains. Therefore, the formalization of the stegotransformation process in the form of a set of perturbations of eigenvalues blocks does not depend on the analyzed region of the image (spatial, frequency), which is a significant advantage of the chosen method of formal representation of the stegotransformation.

D. Practical part

Eigenvalue matrices are well-conditioned [5]. The most vivid features of their perturbations are manifested during compression for the smallest eigenvalues of matrix f . The value of the smallest eigenvalue blocks of an image stored in a lossy format can be compared with the error of their rounding, which is not typical for lossless digital image-blocks. In addition, the behavior of the smallest eigenvalues ($\sigma_6, \sigma_7, \sigma_8$) of image blocks, if we talk about standart block splitting $s = 8$, saved with losses is qualitatively different from the character of the eigenvalues with the same numbers for image blocks saved without loss: the rate of their change in the first case is much less than similar to the second case.

The practical experiment was carried out in the following way: the cover information was embeded out by the proposed stegoalgorithm in the selected random image, after which the secret message was stored in a lossless format (TIF) or with loss format (JPEG) with different

compression coefficients Q. Compression coefficients was chosen with considering the saving of reliability perception $Q > 70\%$.

Developed stegoalgorithm implement block-by-block stegotext transformation/decoding. The number of operations for processing one-block of the digital image matrix in each of the stegoalgorithm does not depend on the size of the whole matrix, it is determined a specific value - a constant relative to the size of the digital image. Thus, the computational complexity of algorithm will be determined by the number of blocks 8×8 in the $M \times N$ digital image matrix:

$$\left[\frac{M}{s} \right] \times \left[\frac{N}{s} \right] = \underline{O}(MN)$$

CONCLUSION

Practical experiment has shown high reliability of the perception of the formed stegomessage in the proposed algorithm. Computing of the reliability the perception coefficient of the proposed algorithm for the shaped stegomessage without compression and after compression with different compression degrees are carried out. As a result of the calculations for the PSNR, the following result obtained: PSNR > 40 Db. Numerical evaluation of MSE also gave good results. The estimation of the correlation coefficient NC, while saving the stegomessage in a lossless format NC = 1. Summary of all the evaluations confirmed the efficiency of the developed stegoalgorithm. The computational complexity of the developed algorithm is defined as $\underline{O}(N^2)$, which means that all algorithms are polynomial of degree 2.

REFERENCES

- [1] A. Kozin, O. Papkovskaya, M. Kozina "Steganography method using Hartley transform", 3th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), 2016, pp. 473-475.
- [2] A.A. Kobozeva, M.A. Melnyk, "Nechuvstvitelnost steganosoobscheniya k szhatiyu i formalnyie dostatochnyie usloviya ee obespecheniya" Zblrnik naukovih prats Vlyskovogo Institutu KiYivskogo natslionalnogo unIversitetu Im. T. Shevchenka, 2012, Vol.38, pp.193–203.
- [3] M.O. Kozina, S.M. Njike Amougou "Steganography method of embedding information with singular value decomposition", Legal, regulatory and metrological support information security system in Ukraine, 2016, No2, pp. 56-61.
- [4] A.A. Kobozeva, I.I. Bobok, A.I. Garbuz "General Principles of Integrity Checking of Digital Images and Application for Steganalysis" Transport and Telecommunication, 2016, Vol.17(2), 128-137.
- [5] M.A. Kozina, A.B.Kozin, O.B Papkovskaya "Stegoalgorithm resistant to compression" The First International Conference on Computer Science, Engineering and Education Applications (ICCSEEA2018), 2018, pp. 419-428.
- [6] A.A.Kobozeva, M.A. Kozina "The steganographic method with a two-stage decoding which provides authentication the container" Informatics and Mathematical Methods in Simulation, 2013. Vol.3(2), pp.169-178.

Модифікація Засобів Open-Source для Підвищення Ефективності Розробки Програмного Забезпечення

Залізецький Василь Володимирович
кафедра обчислювальної техніки
Вінницький національний технічний університет
м. Вінниця, Україна
zwww@i.ua

The Open-Source Tools Modification to Improve Software Development Efficiency

Vasyl Zalizetskyi
department of computer technology
Vinnytsia National Technical University
Vinnytsia, Ukraine
zwww@i.ua

Анотація—У публікації розглянуто питання модифікації таких засобів розробки програмного забезпечення, як Bootstrap, Swagger та Handlebars. Проаналізовано особливості роботи та функціональні можливості цих засобів, а також вказано на їх недоліки. Запропоновано шляхи ліквідації недоліків окремо для кожного з цих засобів, що дозволило підвищити ефективність їх використання як окремо так і в комплексі.

Abstract—The publication addresses the modifying software development tools such as Bootstrap, Swagger, and Handlebars. Analyzed the features of work and functionality and the shortcomings of the mentioned software are identified. The ways of eliminating these disadvantages separately for each of these funds have been proposed, which made it possible to increase the efficiency of their use both separately and in a complex.

Ключові слова—підвищення ефективності, розробка програмного забезпечення, open-source, фреймворки, Bootstrap, Swagger, Handlebars

Keywords— increase efficiency, software development, open source, frameworks

I. ВСТУП

Ефективність розробки програмного забезпечення визначається прийняттям рішень на різних етапах, починаючи з вибору інструментів і засобів розробки. Від цього залежить архітектура програмного забезпечення, структура і представлення даних, функціональні можливості та швидкість розробки.

За останнє десятиліття з'явилося багато різноманітних засобів та інструментів розробки програмного забезпечення. Значне місце серед таких засобів займають фреймворки, що диктують правила побудови архітектури додатку, задають поведінку різних компонентів за промовчуванням та формують каркас. Фреймворки можуть включати допоміжні програми, бібліотеки, мови сценаріїв та інші компоненти, що полегшують розробку програмного забезпечення.

Проте при використанні фреймворків часто виникає потреба розширювати і змінювати їх функціональність відповідно до вимог певного проекту. В публікації розглядаються такі фреймворки, як Bootstrap, Swagger і Handlebars та пропонуються загальні модифікації для підвищення ефективності, що можуть використовуватись при розробці широкого кола проектів.

II. АНАЛІЗ ЗАСОБІВ OPEN-SOURCE

A. Реалізація адаптивності з використанням Bootstrap

Bootstrap — це фреймворк з відкритим кодом, призначений для створення веб-сайтів та веб-додатків, містить шаблони CSS та HTML для типографіки, форм, кнопок, навігації та інших компонентів інтерфейсу, а також додаткові розширення JavaScript [1].

Для того, щоб зрозуміти, як працює адаптивність, потрібно розібратись з основними будівельними блоками та властивостями Bootstrap. Для створення, наприклад трьох однакових по-ширині колонок, які будуть ставати в один ряд на широких та середніх моніторах, потрібно використати класи вказані на рисунку 1. На мобільних пристроях і планшетах ці колонки будуть автоматично складатись.

.col-md-4	.col-md-4	.col-md-4
-----------	-----------	-----------

Рис. 1. Три однакові колонки Bootstrap сітки

Створення двох колонок з вкладеними колонками, які будуть ставати в один ряд на широких та середніх моніторах зображено на рисунку 2. На мобільних пристроях ці колонки (в тому числі і вкладені) будуть складатись.

.col-md-8		.col-md-4
.col-md-6	.col-md-6	

Рис. 2. Дві колонки, які вміщують ще дві колонки

Система сітки Bootstrap має чотири рівні класів: xs (телефони), sm (планшети), md (настільні комп'ютери), та lg (великі екрани). Можна поєднувати майже будь-яку комбінацію цих класів, для створення більш динамічних та гнучких макетів. При цьому сума значень колонок не повинна перевищувати число дванадцять.

ТАБЛИЦЯ 1. СИСТЕМА СІТКИ BOOTSTRAP ДЛЯ РІЗНИХ ПРИСТРОЇВ

Пристрої	Телефони (<768px)	Планшети (≥768px)	Середні монітори (≥992px)	Широкі монітори (≥1200px)
Префікс класу	.col-xs-	.col-sm-	.col-md-	.col-lg-
Поведінка сітки	Горизонтальна завжди	Складена спочатку, але горизонтальна, коли ширша за контрольні точки		
Кількість колонок	12			
Ширина колонки	Автоматична	~62px	~81px	~97px
Ширина проміжку	30px (15px з кожної сторони колонки)			
Можуть вкладатись	Так			
Можуть зміщуватись	Так			
Порядок колонок	Так			

Сітка Bootstrap передбачає, лише один клас для мобільних телефонів - col-xs, але більшість сучасних телефонів обладнанні акселерометром, що дозволяє змінювати орієнтацію зображення в залежності від положення телефону у просторі. А це створює певні труднощі, оскільки те зображення, яке адекватно виглядає при альбомному перегляді на телефоні, виглядає дуже стиснутим при перегляді в портретному режимі та текст накладається з одного блоку на інший.

Окрім того при використанні мобільної версії браузеру Chrome, розмір вікна браузеру часто становить лише 640x360 у альбомному режимі та 360x640 у портретному режимі відповідно. При цьому розширення самого екрану може бути 1920x1680.

У зв'язку з цим, автором публікації пропонується рішення для кращої підтримки мобільних пристроїв, що дозволить скоротити часові затрати на реалізацію адаптивності.

В. Реалізація інтерактивної документації з використанням Swagger

Swagger являє собою набір скриптів, які генерують інтерактивну документацію для web-додатків з REST API. Інтерактивність проявляється в тім, що зі сторінки з документацією можна робити HTTP-запити, а сама документація оновлюється з коментарів, що написані у кодї програми [2].

Під час використання цього інструменту знайдено значний недолік, а саме, те, що обов'язково потрібно описувати варіанти відповідей від серверу, включаючи коди помилок та самі моделі об'єктів, що може повертати сервер. Приклад опису варіантів відповідей від серверу наведено нижче:

```
# responses:
# 200:
# description: Token response
```

```
# schema:
# type: object
# properties:
# token:
# type: string
# description: Authorization token that has to be placed into `X-Auth-Token` header
# 400:
# description: Wrong request format
# schema:
# $ref: '#/definitions/errorResponse'
# 403:
# description: Authorization data is invalid or user not confirmed
# schema:
# $ref: '#/definitions/errorResponse'
# 404:
# description: Specified user is not found
# schema:
# $ref: '#/definitions/errorResponse'
```

Навіть якщо з самими запитом проблем немає, але варіанти відповідей від серверу не описані то при запиті одержуємо помилку “Could not render Operation” Автором пропонується зробити опціональним опис варіантів відповідей від серверу. Це дозволить значно скоротити час на реалізацію інтерактивної документації.

С. Побудова шаблонів сторінок з використанням Handlebars

Handlebars є одним з найбільш популярних, швидких і багатofункціональних шаблонізаторів для JavaScript [3]. Handlebars разом з Ajax є чудовою альтернативою застарілим технологіям, таким, як: JSP, JSTL і org.springframework.web.servlet.ModelAndView.

Після одержання відповіді від серверу у форматі JSON, Handlebars дозволяє подібно до JSTL працювати з різними об'єктами, наприклад опрацьовувати колекції даних з використанням тегу `{{#each ...}}`. В середині тегу `{{#each ...}}` можна одержати індекс елемента з використанням `{{@index}}`. Фрагмент коду:

```
{{#each this}}
<tr>
<td>{{@index}}</td>
<td>{{objectName}}</td>
<td><a id="editObject" data-id="{{id}}"
class="material-icons icon-styled">edit</i></a></td>
<td><a id="deleteObject" data-id="{{id}}"
class="modal-trigger">delete_forever</a></td>
</tr>
{{/each}}
```

Але при відображенні списку елементів люди звикли бачити не id елементів і не індекс масиву, що починається з 0, а звичайну нумерацію починаючи з 1. На момент написання цієї публікації Handlebars не надавав таку можливість, тому автором пропонується така функціональність, що дозволить в подальшому скоротити час на пошук шляхів реалізації подібної функціональної можливості.

III. ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ЗАСОБІВ OPEN-SOURCE

А. Модифікація Bootstrap

Для подолання описаної проблеми, реалізовано додатковий клас сітки для Bootstrap.

Новий клас отримав назву .col-lx. Оскільки col-xs-означає extra small devices, то клас .col-lx означає less

than extra small, тобто менше ніж стандартний мобільний альбомний режим.

Реалізовано це наступним чином:

```
.col-lx-1, .col-lx-2, .col-lx-3, .col-lx-4, .col-lx-5, .col-lx-6, .col-lx-7, .col-lx-8, .col-lx-9, .col-lx-10, .col-lx-11, .col-lx-12 .col-xs-1 {
  position: relative;
  padding-right: 15px;
  padding-left: 15px;
}
.col-lx-1, .col-lx-2, .col-lx-3, .col-lx-4, .col-lx-5, .col-lx-6, .col-lx-7, .col-lx-8, .col-lx-9, .col-lx-10, .col-lx-11, .col-lx-12 {
  float: left;
}
.col-lx-12 {
  width: 100%;
}
.col-lx-11 {
  width: 91.66666667%;
}
.col-lx-10 {
  width: 83.33333333%;
}
.....
.col-lx-3 {
  width: 25%;
}
.col-lx-2 {
  width: 16.66666667%;
}
.col-lx-1 {
  width: 8.33333333%;
}
```

Адаптивність працює не тільки завдяки відсотковому вказанню ширини для різних блоків, основний рушій адаптивності це використання медіа-селекторів @media.

Чому в поданому вище CSS-коді, немає медіа-селектору, можна пояснити тим, що CSS-файл умовно ділиться на медіа-блоки, наступним чином:

```
@media (min-width: 768px) { .col-sm-... }
@media (min-width: 992px) { .col-md-... }
@media (min-width: 1200px) { .col-lg-... }
```

Як, видно із наведено прикладу, тут немає медіа-селектору для найменшого стандартного блоку .col-xs, тобто він використовується за замовчуванням, а при більших екранах спрацьовує відповідний медіа-селектор.

За таким же принципом реалізовано новий найменший блок .col-lx, із заданням медіа-селектору для старого найменшого блоку .col-xs:

```
@media (min-width: 640px) { .col-xs-... }
```

В. Модифікація Swagger

Зміни стосуються файлу src/core/components/operation.jsx в репозиторії swagger-ui.

```
import { OrderedMap } from "immutable"
```

До функції render() потрібно додати

```
if(responses === undefined) responses = new OrderedMap
if(response && response.size > 0) {
  let notDocumented = (responses !== undefined)?
!responses.get(String(response.get("status"))): true
```

```
response = response.set("notDocumented", notDocumented)
}
```

Запропоновані правки роблять опис варіантів відповідей від серверу опціональним, що дозволяє значно прискорити процес налаштування інтерактивної документації для своїх проектів. Після внесення змін і перевірки ресурсів засобами npm помилка більше не виникатиме.

С. Модифікація Handlebars

Для вирішення описаної проблеми, можна зареєструвати свій власний helper:

```
Handlebars.registerHelper({
  'math': function(lvalue, operator, rvalue, opts) {
    lvalue = parseFloat(lvalue);
    rvalue = parseFloat(rvalue);
    return {
      "+": lvalue + rvalue
    }[operator];
  });
});
```

І далі робити так `{{math @index "+" 1}}`, але цей варіант спрацьовує не завжди і при оновленні сторінок, деколи ми отримуємо помилку Missing helper: math. Тому автором тез пропонується модифікація handlebars.js, а саме функції execution:

```
function execution(field, index, last) {
  if (data) {
    data.key = field;
    data.index = index;
    data.number = index + 1;
    data.first = index === 0;
    data.last = !!last;
    if (contextPath) {
    }
  }
  ret = ret + fn(context[field], {
    data: data,
    blockParams: _utils.blockParams([context[field], field], [contextPath + field, null])
  });
}
```

Після такої модифікації можна використовувати `{{@number}}`, де потрібно вивести номер елемента в списку.

REFERENCES ЛІТЕРАТУРА

- [1] Bootstrap - The most popular HTML, CSS, and JS library in the world. [Електронний ресурс]. —Режим доступу: <https://getbootstrap.com>, вільний (дата звернення: 16.11.2018). — Назва з екрана.
- [2] The Best APIs are Built with Swagger Tools [Електронний ресурс]. —Режим доступу: <https://swagger.io>, вільний (дата звернення: 16.11.2018). —Назва з екрана.
- [3] Handlebars.js: Minimal Templating on Steroids [Електронний ресурс]. —Режим доступу: <https://handlebarsjs.com>, вільний (дата звернення: 16.11.2018). —Назва з екрана.
- [4] О. Азаров, О. Черняк, і Л. Савицька, Аспекти критичного підходу до викладання поняття поліморфізму в об'єктно-орієнтованому програмуванні, ІТКІ, vol 39, № 2, с. 31-34, Жов 2017.

Виявлення Штучних Змінень в Зображеннях За допомогою Нейронних Мереж

Глайборода Марина Віталіївна
Факультет інформаційно-аналітичних технологій та менеджменту
Харківський національний університет радіоелектроніки
Харків, Україна
m.hlaiboroda@gmail.com

Detection of Artificial Modifications of Images Using Neural Networks

Hlaiboroda Maryna Vitaliivna
Faculty of Information and Analytical Technologies and Management
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
m.hlaiboroda@gmail.com

Анотація—Виявлення штучних змінень в зображеннях відрізняється від традиційного семантичного виявлення об'єктів, оскільки воно приділяє більше уваги спотворенням, ніж вмісту зображення. У роботі описана двопоточна мережа R-CNN з end-to-end навчанням, щоб виявляти регіони з артефактами на підроблених зображеннях. Один з двох потоків – це потік RGB, метою якого є витягування функцій з вхідного кольорового зображення для виявлення областей спотворення. Інший – це потік шумів, який використовує шумові ознаки, витягнуті з фільтруючого шару моделі стеганалізу, щоб виявити зашумовану несумісність між автентичними та підробленими регіонами. Дослідження з трьома стандартними наборами даних зображень зі штучними зміненнями демонструють, що створена двопоточна архітектура ефективніша за кожний окремих потік, а також забезпечує високу продуктивність.

Ключові слова—штучні змінення, двопоточні нейронні мережі, потік RGB, шумовий потік, машинне навчання, глибоке навчання.

Abstract—Image manipulation detection is different from traditional semantic object detection because it pays more attention to tampering artifacts than to image content. In work described a two-stream Faster R-CNN network trained end-to-end to detect regions with artifacts given a manipulated image. One of the two streams is an RGB stream whose purpose is to extract features from the RGB image input to find tampering regions. The other is a noise stream that leverages the noise features extracted from a steganalysis rich model filter layer to discover the noise inconsistency between authentic and tampered regions. Experiments on three standard image manipulation datasets demonstrate that created two-stream architecture outperforms each individual stream, and also achieves high performances.

Keywords—image manipulation, two-stream neural networks, RGB stream, noise stream, machine learning, deep learning.

I. ВСТУП

Завдяки досягненням технологій редагування зображень і зручному у використанні програмному забезпеченню для редагування стали широко доступні дешеві підробки або маніпуляційні процеси генерації зображень. Серед методів фальсифікації найбільш поширеними є склеювання, копіювання, переміщення та

видалення. Склеювання зображень полягає в копіюванні регіонів з автентичного зображення та вставленні їх на інші зображення. Копіювання-переміщення полягає в копіюванні регіонів зображень та вставленні їх в межах одного зображення, а видалення – в видаленні регіонів з автентичного зображення і замалюванні утворених прогалів. Приклади цих маніпуляцій показані на рис.1. Навіть при ретельній перевірці людям важко розпізнати зловмисні підроблені регіони.

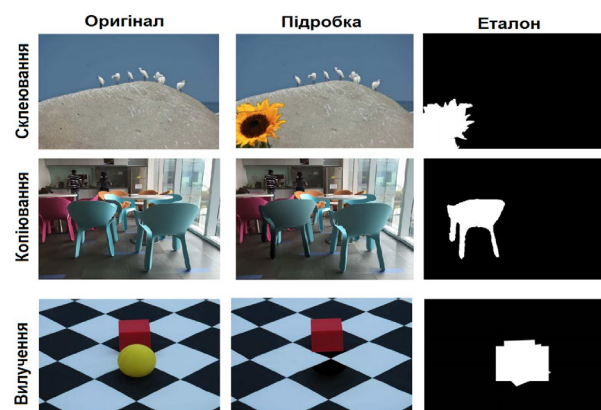


Рис.1. Приклади підроблених зображень, що зазнали різноманітних методів фальсифікації. Зверху вниз: склеювання, копіювання та видалення

В результаті, відрізнення автентичних зображень від підроблених стає дедалі складнішим. фальшиві зображення все частіше і частіше використовуються в політичних цілях і для несумлінного ведення справ. Ще сильніше ускладнює ситуацію те, що навіть при ретельній перевірці людині дуже важко визначити оброблене місце на зображенні. На відміну від сучасних мереж виявлення об'єктів [1, 2, 3], які спрямовані на виявлення всіх об'єктів різних категорій на зображенні, мережа для виявлення маніпулювання на зображеннях має мету виявляти лише підроблені області (як правило, об'єкти). В роботі досліджується як вибрати мережу виявлення об'єктів для виявлення маніпуляцій на зображеннях шляхом вивчення вмісту зображення RGB та характеристик шумів зображень.

Нещодавно запропонована архітектура [4], заснована на сегментах Long Short Term Memory Network (LSTM), підробила виправлення, демонструючи стійкість до

кількох методів втручання, навчаючись виявленню підроблених країв. В роботі пропонується нова структура нейронної мережі для виявлення змін з двома потоками, яка не тільки моделює артефакти візуального спотворення, але також фіксує невідповідності в локальних шумових особливостях.

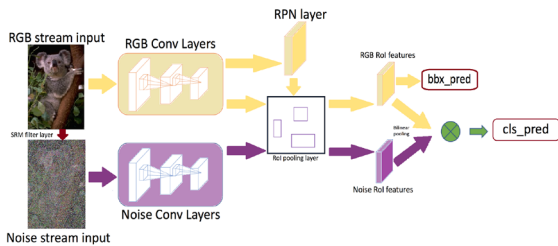


Рис.2. Ілюстрація розробленої двопоточної мережі R-CNN

На малюнку 2 RGB-потік моделює візуальні спотворюючі артефакти, такі як незвично високий контраст по краях об'єкта, і регресує обмежуючі прямокутники до істини. Потік шуму спочатку отримує карту характеристик шуму, передаючи вхідне зображення RGB через шар фільтра SRM, і використовує функції шуму, щоб надати додаткові докази для класифікації маніпуляцій. RGB і шумові потоки використовують одні і ті ж пропозиції регіону з мережі RPN, які використовують тільки функції RGB в якості вхідних даних. Рядок згрупування RoI [5, 6] вибирає просторові функції як з RGB, так і шумових потоків. Передбачені обмежувальні коробки (позначені як 'bbx pred') генеруються з функцій RGB RoI. Білінійний пул [7, 8, 9] після об'єднання RoI дозволяє мережі об'єднати просторові функції спільного виникнення з двох потоків. Нарешті, передаючи результати за допомогою повністю підключеного шару та шару softmax [10, 11], мережа виробляє передбачувану мітку (позначається як 'cls pred') і визначає, чи були оброблені передбачені області чи ні.

Була досліджена можливість адаптації нейромереж, які використовуються для детекції образів, до задачі визначення підроблених зображень. Пропонований підхід дозволяє виявити всі три способи зміни зображень: склеювання, копіювання-переміщення та вилучення.

II. ЗАПРОПОНОВАНЕ РІШЕННЯ

У цьому розділі описано обрану архітектуру двопоточної нейромережі, яка обробляє одночасно RGB-зображення і його шумові особливості. Також перероблено Faster R-CNN в двопоточну нейромережу з end-to-end навчанням. Перший потік обробляє особливості з RGB каналів для визначення візуальних невідповідностей на кордоні відредагованих областей, а також для реєстрації розбіжності контрасту оброблених і оригінальних частин зображення. Другий потік аналізує локальні шумові особливості. Два потоки доповнюють один одного в пошуку фальсифікації на зображеннях.

Розроблена нейромережа складається з трьох головних блоків.

A. RGB-потік

Відповідальний за семантичне визначення відредагованих областей. Особливості вихідного RGB

зображення подаються на вхід нейромережі ResNet 101, потім вони використовуються для визначення методу редагування за допомогою техніки bounding box regression. Нейромережа RPN [12] в RGB потоці також використовує ці особливості для визначення можливого RoI (region of interest), який буде використаний в рамках цієї техніки. Експерименти показують, що RGB особливості визначаються нейромережею RPN краще, ніж шумові особливості.

Одного RGB потоку недостатньо для визначення тих областей редагування, де була застосована пост-обробка для згладжування кордонів вставленого елемента і вирівнювання його контрасту відповідно із зображенням. Тому був введений другий потік.

B. Шумовий потік

Введений для визначення шумових, а не семантичних особливостей вхідного зображення. Тут використовується досягнення стегоаналіза (steganalysis rich model (SRM)) [13, 14, 15, 16] і застосовується SRM-фільтр [17] для створення шумових особливостей на зображенні. Отримувані карти шумових особливостей можна побачити в третій колонці на малюнку 3.

Шум створюється як різниця між значенням пікселя і оцінкою його значення, отриманою за допомогою інтерполяції значень сусідніх пікселів. У шумовому потоці використовується той же об'єднуючий RoI шар (RoI pooling layer на схемі) [18], що і в RGB потоці.



Рис.3. Ілюстрація артефактів підробки

C. Шар білінійного об'єднання

Виробляє з'єднання RGB і шумового потоків в двопоточній СНМ зі збереженням просторової інформації для підвищення достовірності виявлення відредагованих областей. На вихід цього шару надходять RoI особливості обох потоків. Потім проводиться нормування по квадратичній нормі і нормі L2, і дані передаються в повністю з'єднаний шар. У техніці bounding box regression [19] використовується гладка функція втрат L1, а для класифікації методу редагування застосована функція втрат перехресної ентропії.

III. ЕКСПЕРИМЕНТИ ТА РЕЗУЛЬТАТИ

У цьому розділі приведені результати експериментів з використанням запропонованих методів та вибраної реалізації.

Представлений метод був приведений в кількісне порівняння з іншими сучасними методами на прикладі набору даних NIST16, COVER і CASIA [20, 21, 22].

- NIST16 – складний набір даних, який містить всі три методи підробки. Маніпуляції в цьому наборі даних генеруються після обробки, щоб приховати видимі сліди. Вони також надають маски для укріплення прав на землю для оцінки.
- CASIA надає склеєні і скопійовані зображення різних об'єктів. Ретельно підібрані області ретельно відбираються, і також застосовується деяка пост-обробка, така як фільтрація і розмиття. Маски для наземної істини отримують шляхом порога різниці між підробленими та оригінальними зображеннями. Я використовую CASIA 2.0 для навчання і CASIA 1.0 для тестування.
- COVER – відносно невеликий набір даних, орієнтований на копіювання. Він охоплює аналогічні об'єкти, такі як вкладені області, щоб приховати спотворювані артефакти (див. другий рядок на рис.1). Надаються маски із заземленням.

Порівняння було виконано із застосуванням оціночної метрики Area Under the receiver operating characteristic Curve (AUC).

Розроблена модель (RGB-N) була приведена в порівняння як з іншими методами (ELA, NOI1, CFA1), так і з одним шумовим потоком (Noise Net) і з моделлю, в якій реалізовано пряме об'єднання всіх зареєстрованих RoI з RGB Net і Noise Net (Late fusion). Результати даного порівняння представлені в таблиці 1.

ТАБЛИЦЯ 1. Порівняння по метриці AUC

	NIST16	COVER	CASIA
ELA [21]	0.429	0.583	0.613
NOI1 [24]	0.487	0.587	0.612
CFA1 [14]	0.501	0.485	0.522
Noise Net	0.881	0.753	0.693
Late Fusion	0.924	0.793	0.777
RGB-N	0.937	0.817	0.795

З таблиць видно, що модель RGB-N перевершує такі загальновідомі методи, як ELA, NOI1 і CFA1. Причиною цього може бути той факт, що всі ці методи сфокусовані на специфічних артефактах редагування, які містять тільки частину інформації про локалізацію обробленого регіону.

Вже згадана нейромережа також дозволяє визначити метод, використаний при редагуванні зображення. Використовуючи інформацію шумового і RGB потоків, вона може розрізнити склеювання, копіювання і видалення, вироблені при обробці зображень.

ВИСНОВОК

Був розглянутий новий підхід до визначення автентичності зображення, який показав досить високі результати. Такий результат досягається шляхом

поєднання двох різних потоків (RGB і шумового) для виявлення особливостей, що виникають при редагуванні. Як видно, два потоки доповнюють один одного при визначенні оброблених областей зображення. Шумові особливості, виявлені SRM фільтром, дозволяють нейромережі перехоплювати шумові невідповідності між відредагованим і оригінальним регіонами зображення.

Крім того, RGB-N здатна розрізнити методи редагування. Таким чином, вона не тільки вказує, яка область зображення була оброблена, а й говорить, як ця область була оброблена: чи був об'єкт вставлений, що знаходиться на відстані або копіювати. Тому, використовуючи таку нейромережу, можна без зусиль відрізнити оригінальне зображення від підробки.

REFERENCES ЛІТЕРАТУРА

- [1] Nist nimble 2016 datasets. <https://www.nist.gov/itl/iad/mig/nimble-challenge-2017-evaluation/> 2, 5.
- [2] J. H. Bappy, A. K. Roy-Chowdhury, J. Bunk, L. Nataraj, and B. Manjunath. Exploiting spatial structure for localizing manipulated image regions. In ICCV, 2017. 1, 3, 6.
- [3] B. Bayar and M. C. Stamm. A deep learning approach to universal image manipulation detection using a new convolutional layer. In IH&MMSec, 2016. 3
- [4] T. Bianchi, A. De Rosa, and A. Piva. Improved dct coefficient analysis for forgery localization in jpeg images. In ICASSP, 2011. 2, 3.
- [5] J. Chen, X. Kang, Y. Liu, and Z. J. Wang. Median filtering forensics based on convolutional neural networks. Signal Processing Letters, 2015. 3.
- [6] D. Cozzolino, D. Gragnaniello, and L. Verdoliva. Image forgery localization through the fusion of camera-based, feature-based and pixel-based techniques. In ICIP, 2014. 1.
- [7] D. Cozzolino, G. Poggi, and L. Verdoliva. Splicebuster: A new blind image splicing detector. In WIFS, 2015. 3, 4.
- [8] [8] D. Cozzolino, G. Poggi, and L. Verdoliva. Recasting residual-based local descriptors as convolutional neural networks: an application to image forgery detection. In IH&MMSec, 2017. 3.
- [9] D. Cozzolino and L. Verdoliva. Single-image splicing localization through autoencoder-based anomaly detection. In WIFS, 2016. 1, 3.
- [10] J. Dai, Y. Li, K. He, and J. Sun. R-fcn: Object detection via region-based fully convolutional networks. In NIPS, 2016. 1.
- [11] T. J. De Carvalho, C. Riess, E. Angelopoulou, H. Pedrini, and A. de Rezende Rocha. Exposing digital image forgeries by illumination color classification. TIFS, 2013. 1.
- [12] J. Dong, W. Wang, and T. Tan. Casia image tampering detection evaluation database 2010. <http://forensics.idealtest.org> 2, 5.
- [13] J. Dong, W. Wang, and T. Tan. Casia image tampering detection evaluation database. In ChinaSIP, 2013. 5.
- [14] P. Ferrara, T. Bianchi, A. De Rosa, and A. Piva. Image forgery localization via fine-grained analysis of cfa artifacts. TIFS, 2012. 6.
- [15] J. Fridrich and J. Kodovsky. Rich models for steganalysis of digital images. TIFS, 2012. 2, 3, 4.
- [16] M. Gao, R. Yu, A. Li, V. I. Morariu, and L. S. Davis. Dynamic zoom-in network for fast object detection in large images. In CVPR, 2018. 1.
- [17] Y. Gao, O. Beijbom, N. Zhang, and T. Darrell. Compact bilinear pooling. In CVPR, 2016. 2, 5.
- [18] R. Girshick. Fast r-cnn. In ICCV, 2015. 1.
- [19] M. Goljan and J. Fridrich. Cfa-aware features for steganalysis of color images. In SPIE/IS&T Electronic Imaging, 2015. 1, 3.
- [20] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016. 3.
- [21] N. Krawetz. A picture's worth... Hacker Factor Solutions, 2007. 6.
- [22] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick. Microsoft coco: Common objects in context. In ECCV, 2014. 2, 5.

Система Підтримки Роботи Ситуаційного Центру На Основі Інтелектуальних Хмарних Технологій

Леонід Куперштейн
Факультет інформаційних технологій
та комп'ютерної інженерії
Вінницький національний технічний
університет, Вінниця, Україна
kupershtein.lm@gmail.com

Михайло Кренцін
Факультет інформаційних технологій
та комп'ютерної інженерії
Вінницький національний технічний
університет, Вінниця, Україна
mishatron98@gmail.com

Алла Штокал, Олександр Восьмушко
Факультет інформаційних технологій
та комп'ютерної інженерії
Вінницький національний технічний
університет, Вінниця, Україна
alla.shtokal@gmail.com
sasha.vosmushko@gmail.com

The Support System For The Situational Center On The Intelligent Cloud Technologies Base

Leonid Kupershtein
Faculty of Information Technology and
Computer Engineering
Vinnytsia National Technical
University, Vinnytsia, Ukraine

Mykhailo Krentsin
Faculty of Information Technology and
Computer Engineering
Vinnytsia National Technical
University, Vinnytsia, Ukraine

Alla Shtokal, Oleksandr Vosmushko
Faculty of Information Technology and
Computer Engineering
Vinnytsia National Technical
University, Vinnytsia, Ukraine

Анотація — Запропоновано архітектуру системи підтримки автоматизації роботи ситуаційного центру міста по виявленню інцидентів. Реалізація системи базується на інтелектуальному хмарному сервісі Google Cloud Vision.

Abstract — The support system architecture for the automation of the city situational center for the incidents detection is proposed. The system implementation is based on the intelligent technology – Google Cloud Vision.

Ключові слова — безпека міста ситуаційний центр, хмарні технології, Google Cloud Vision.

Keywords — city security, situational center, cloud technologies, Google Cloud Vision.

I. ВСТУП

Питання безпеки на вулицях міст України і світу на сьогоднішній день є досить важливим. Концепція сучасного міста будується на забезпеченні комфортних умов життя населення за допомогою інноваційних технологій. Впровадження таких технологій дозволяє економічно та екологічно використовувати всі міські системи життєдіяльності.

Основною систем сучасного міста є ідеально працююча система безпеки. Щоб зберегти безпеку громадян все частіше почали впроваджувати ситуаційні та аналітичні центри [1]. Працівники центру моніторять місця найбільшого скупчення людей, об'єктів, які потребують посиленої уваги з погляду громадської безпеки. Зокрема, камери встановлюють на виїздах та в'їздах у місто, магістральних вулицях, перехрестях, площах, вокзалах, зупинках, навчальних закладах тощо.

Такі установи є відносно новими в Україні, а за кордоном це вже відома практика. Зазвичай ситуаційний центр являє собою простору залу, де в кілька рядів розміщені монітори. За ними спостерігають десятки операторів, кожен відповідає за певну ділянку, за якою потрібно дуже пильно стежити, а загальна картина по місту виводиться на велике спільне відеотабло. Це є досить відповідальна робота, при якій важлива будь-яка

деталь. Тому питання моніторингу об'єктів є досить актуальним.

II. ПОСТАНОВКА ЗАДАЧІ

Варто зазначити, що використання камер відеоспостереження є досить ефективним задля збереження безпеки громадян, адже це дозволяє оперативно реагувати на ті чи інші ситуації. Проте, слідкувати за всіма у режимі реального часу просто неможливо. Дуже часто кількість камер сягає кількох сотень, що дуже ускладнює роботу ситуаційних центрів, адже потребує більше людських ресурсів та значні обсяги технічного обладнання.

Якщо працівники центру помічають правопорушення чи екстрену ситуацію, повідомляють поліцію, «Цілодобову варту» або на гарячу лінію екстрених служб для подальшого вирішення. Також, згодом, аналітичний центр такі дані структурує у звіти або формує відкриті дані (Open Data) у зручному форматі. За необхідності ці дані можуть надаватися зацікавленим службам – правоохоронцям, силовикам і комунальникам.

На сьогоднішній день ситуаційний центр працює переважно в режимі виявлення правопорушень [2]. Але цього недостатньо, адже важливо також слідкувати за пожежами, виявленням скупчень людей та автомобілів в певній місцевості. На жаль при цьому негативно спрацьовує людський фактор, а саме цього ресурсу та відсутність можливості швидко перевіряти сотні камер та приймати оперативні рішення. Тому актуальним є розробка автоматизованої системи, яка б дозволила перекласти на комп'ютер частину кропіткої роботи та забезпечити своєчасне реагування на інциденти.

III. РОЗРОБКА СИСТЕМИ

Одним із підходів вирішення вищеприписаної задачі є використання хмарних технологій та штучного інтелекту. Це дозволить у багато разів збільшити

ефективність роботи ситуаційних центрів, не витрачаючи багато людських ресурсів і часу. Завдяки технології покадрового розпізнавання об'єктів на відео можна зібрати статистику, зробити прогнозування чи передбачити небезпеку (наприклад, розпізнати пожежу, аварію тощо) та, повідомити відповідну службу. Відтак, можна спростити контроль за громадським транспортом, трафіком на конкретних напрямках, роботою комунальників тощо. Наприклад, можна з точністю до 99% розпізнати пожежу на початковому етапі займання та одразу сповістити у пожежну службу.

Виходячи з цього, було розроблено систему, яка являє собою веб-сервіс. Система складається з трьох модулів: модуль збору даних з камер, аналізу даних, формування наборів відкритих даних. Архітектура системи підтримки роботи ситуаційного центру представлена на рис. 1.



Рис. 1. Архітектура системи

Сервер встановлює з'єднання з кожною камерою і в задані моменти часу отримує кадр з відео потоку. Цей кадр передається в модуль аналізу даних, що в свою чергу надсилає його для аналізу в сервіс Google Cloud Vision і отримує результат. Далі відбувається обробка отриманої інформації на наявність специфічних сутностей. У разі виявлення подібного формується звіт, інформація записується в базу даних та відправляється у ситуаційний центр для перевірки людиною-оператором.

Серверна частина розроблена мовою C# з використанням фреймворку ASP.NET Core, що дозволяє встановити його на сервері з будь-якою операційною системою [3]. Виділення кадру з потоку відео відбувається завдяки набору вільних бібліотек з відкритим сирцевим кодом FFmpeg [4]. Система дозволяє додавати нові камери для збору інформації.

Модуль аналізу даних використовує Google Cloud Vision API, що представляє собою добре навчену нейронну мережу [5]. Кожен кадр аналізується на наявність сутностей (наприклад, автомобіль, дорога, людина, колір, пожежа тощо). Точність такого аналізу зображень є досить висока. Проте за використання сервісу необхідно сплачувати певну суму: перші 1000 запитів в місяць є безкоштовні, наступні 1001-5000000 запитів коштують \$1,5 за кожен тисячу, далі \$1 за кожен тисячу запитів. Щомісячна вартість використання сервісу залежить від кількості камер, які треба

аналізувати та проміжку часу, з яким береться кадр відео. Наприклад, якщо в місті встановлено 500 камер та аналізувати кожен з періодом в 2 секунди, то отримаємо вартість \$972 000 за 30 днів, а за рік – 331 128 000 гривень. Якщо врахувати річний бюджет міста (бюджет м. Вінниці на 2018р. складає близько 5 млрд. гривень), то ця сума є незначною.

Модуль формування є досить важливою частиною системи. Тут дані нормалізуються та записуються у базу даних. Для кожного кадру зберігається час його створення, камера, з якої він взятий та масив рядків (виявлені сутності). При цьому, якщо виявлено якусь небезпеку чи надзвичайну ситуацію, формується звіт для відправлення в ситуаційний центр. Далі працівники зможуть побачити детальну інформацію, швидко перевірити її та оперативно зреагувати.

В якості системи керування базою даних (СКБД) було використано PostgreSQL

[6].

Створення таблиць відбувалось через EntityFramework, що дозволяє описати сутності класами і автоматично згенерувати відповідні таблиці. PostgreSQL не контролюється якоюсь однією компанією, її розробка можлива завдяки співпраці багатьох людей та компаній, які хочуть використовувати цю СКБД та впроваджувати у неї найновіші досягнення.

ВИСНОВКИ

Проаналізувавши актуальність і важливість досліджуваної теми, було розроблено систему для створення наборів відкритих даних з використанням штучного інтелекту. Система призначена для автоматизації збору та обробки інформації з камер відео спостереження для автоматизації роботи ситуаційного центру та його своєчасного реагування на виявлені інциденти. Практична реалізація виконана з використанням інтелектуального модуля Google Cloud Vision. Система з високою точністю може виявляти різноманітні небезпечні інциденти. Система реалізована у вигляді web-сервісу. Також система дозволяє формувати набори відкритих даних на основі інформації з камер спостереження, які можна використовувати різними дослідниками та аналітиками, наприклад, для побудови, проектування і навчання побічних інтелектуальних систем.

ЛІТЕРАТУРА REFERENCES

- [1] Ситуаційні центри. Теорія і практика / За ред. А.О. Морозова, Г.Є. Кузьменко, В.А. Литвинова. – Київ: СП «Інтертехнодрук», 2009. – 348 с.
- [2] Як працює ситуаційний центр [Електронний ресурс] – Режим доступу: <http://misto.vn.ua/news/item/id/11621>
- [3] ASP Core | Полное руководство [Електронний ресурс] – Режим доступу: <https://metanit.com/sharp/aspnet5/>
- [4] FFmpeg [Електронний ресурс] – Режим доступу: <https://www.ffmpeg.org/>
- [5] Cloud Vision [Електронний ресурс] – Режим доступу: <https://cloud.google.com/vision/overview/docs>
- [6] PostgreSQL - Вікіпедія [Електронний ресурс] – Режим доступу: <https://uk.wikipedia.org/wiki/Postgre>.

Задача Моделювання Мотивації Членів Інтелектуальної Команди

Пономарьова С.В.
кафедра системотехніки
Харківський національний університет радіоелектроніки
Харків, Україна
svitlana.ponomarova@nure.ua

Калита Н.І.
кафедра системотехніки
Харківський національний університет радіоелектроніки
Харків, Україна
nadiia.kalyta@nure.ua

The Task Of Modeling The Motivation Of Members Of The Intellectual Team

Ponomarova S.V.
Department of System Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
svitlana.ponomarova@nure.ua

Kalyta N.I.
Department of System Engineering
Kharkiv National University of Radio Electronics
Kharkiv, Ukraine
nadiia.kalyta@nure.ua

Анотація — Робота присвячена розробленню моделі мотивування інтелектуальної команди на основі типів особистості її членів. Змодельовані можливі рівні задоволеності співробітників в залежності від обраних факторів мотивації та функція зростання трудової мотивації індивідуума в залежності від вкладених ресурсів.

Ключові слова — інтелектуальна команда, моделювання, фактори мотивації, типи особистості.

Abstract — The work is devoted to the development of a motivation model of an intellectual team based on the types of personality of its members. The possible levels of employee satisfaction are modeled depending on the chosen motivation factors and the function of growth of labor motivation of the individual is modeled depending on the invested resources.

Keywords — intellectual team, modeling, motivation factors, personality types.

1. ВСТУП

Для управління будь-яким проектом на період його здійснення створюється специфічна тимчасова організаційна структура, очолювана керівником проекту. Найчастіше вона називається проектною командою. Команда являє собою групу людей, що прагнуть до досягнення результату. Під командою розуміється колектив (об'єднання людей, що здійснюють спільну діяльність і володіють спільними інтересами), здатний досягати мети автономно й узгоджено, при мінімальних керуючих впливах [1].

Раніше при управлінні проектами в організації співробітники розглядалися як елемент технічної системи, і завдання управління командою зводилася до задачі управління людськими ресурсами компанії. У сучасних умовах все більша увага приділяється особистості співробітника, і персонал компанії розглядається як її інтелектуальний капітал, що визначає актуальність проектного менеджменту та безпосередньо впливає на ефективність управління проектом і успішність його реалізації.

Основним інструментом управління інтелектуальною командою є мотивування її членів. Вирішення задачі мотивування членів інтелектуальної команди складається з наступних етапів:

- 1) визначення факторів мотивації усіх співробітників у команді;
- 2) ідентифікація ступеня задоволеності, тобто початкового рівня мотивації кожного члена команди;
- 3) пошук керуючих параметрів впливу на трудову мотивацію членів інтелектуальної команди;
- 4) безпосередньо вплив на трудову мотивацію з урахуванням індивідуальних особливостей працівника;
- 5) оцінка ефективності вибраних заходів впливу.

2. АНАЛІЗ ВІДГУКУ ІНДІВІДУУМА НА ЗОВНІШНІЙ ВПЛИВ

Успішність процесу мотивування члена інтелектуальної команди значною мірою залежить від його особистісних рис, переваг, цінностей і установок. Людина, як суб'єкт активної свідомої діяльності, має конкретні сукупності соціально-значущих рис. Ці риси формуються, з одного боку, під впливом його природних властивостей (психічних процесів, специфіки пам'яті, емоцій, почуттів, особливостей сприйняття). Ці якості утворюють незмінні, об'єктивні риси особистості. З іншого боку, індивідууму притаманні такі соціальні моменти, як досвід, знання, звички, ідеали, інтереси, які є набутими і змінними.

Всі риси особистості і їх окремі прояви тісно пов'язані між собою і визначають:

- 1) індивідуальність;
- 2) активність і цілеспрямованість;
- 3) стійкість поведінки, що дозволяє передбачити його в основному, а, отже, здійснювати управління ним, використовуючи різні методи впливу.

Ймовірність відгуку індивідуума на властивості оточення або на зміни цих властивостей є функцією інтенсивності впливу [2]. Цю функцію можна представити графічно (рис.1).

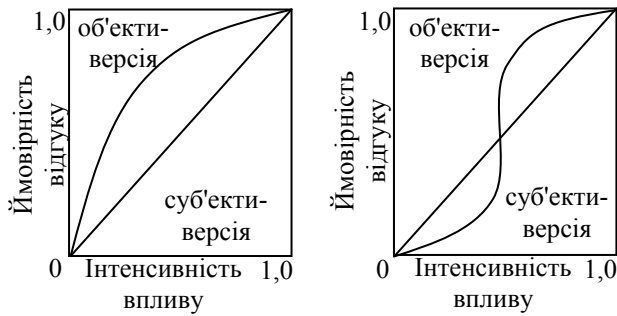


Рис. 1 – Функція відгуку на оточення

Простір відгуку на оточення можна розділити на дві рівні частини діагоналю, проведеною з нижнього лівого у верхній правий кут. Верхня частина (об'єктиверсія) відповідає порівняно високому ступеню чутливості до оточення, нижня (суб'єктиверсія) – порівняно низькому ступеню чутливості. Той самий простір можна було б розділити горизонтальною лінією, проведеною через середню точку ординати, але діагональ краща, так як, суб'єктам властиво відгукуватися частіше на більш інтенсивні впливи. Тому навіть дуже нечутливий до оточення індивідуум, швидше за все, відгукнеться на дуже сильний вплив. Ясно, що з точки зору входів (впливів) об'єктиверсія відповідає екстраверсії, а суб'єктиверсія – інтроверсії. Отже, об'єктиверт чутливий і сприйнятливий до того, що відбувається навколо нього, тому більш схильний підпадати під вплив того, що відбувається. Суб'єктиверт більше схильний до самоаналізу та впливу своїх власних думок і почуттів, ніж оточення на нього. Об'єктиверт об'єктивно спрямовується і мотивується, в той час як суб'єктиверт спрямовується і мотивується суб'єктивно.

На основі цих двох типів особистості в соціоніці було запропоновано більш детальну класифікацію, що складається з 16 психотипів та в якій враховані зазначені вище особливості індивідуумів, наприклад, «логіко-інтуїтивний екстраверт», «етико-сенсорний інтроверт», тощо.

3. МЕТОД ВИРІШЕННЯ ЗАДАЧІ МОДЕЛЮВАННЯ

У загальному вигляді залежність результативних показників функціонування систем від показників-факторів описується виробничою функцією, або кривою зростання. Специфіка виробничих функцій в тому, що в

якості незалежних змінних в них виступають в основному різні ресурси виробництва [3]:

$$y = f(x_1, x_2, \dots, x_n, a_1, a_2, \dots, a_k), \quad (1)$$

де y – результуючий фактор;

$x_i, i=1, n$ – ресурси;

$a_j, j=1, k$ – параметри функції.

Як правило, припущення про вид функції (1) робиться на основі аналізу особливостей предметної області: визначається змінна y , істотні фактори $x_i, i=1, n$ і область визначення функції. Виробнича функція має наступні властивості:

1) неперервна і неперервно диференційована, принаймні, двічі;

2) є функцією, яка не убуває, тобто $f(\lambda x) \geq f(x), \lambda > 1$;

3) друга похідна $\partial^2 y / \partial x_i^2$, починаючи з будь-якого значення x_i^0 не є позитивною, тобто $\partial^2 y / \partial x_i^2 \leq 0$, що означає зниження ефективності використання ресурсу, або логістичний закон.

Найчастіше використовуються поліноміальні, експоненціальні і S-образні виробничі функції.

Дослідження динаміки соціальних і економічних процесів виявило досить сильне поширення ефекту насичення, тобто виходу на асимптоту при досягненні певних значень результуючого показника.

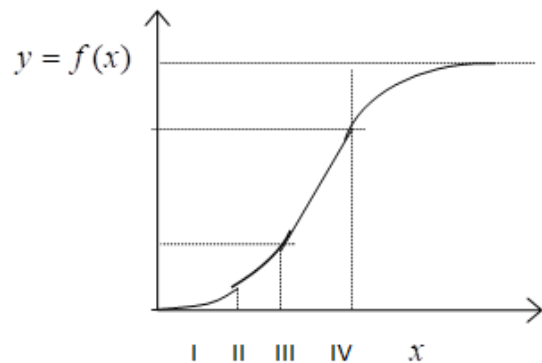


Рис. 2 – S-образна крива

При цьому показник спочатку зростає незначно (I), потім зростання прискорюється (II), стаючи практично постійним (III), а потім знову сповільнюється, прагнучи до якої-небудь межі (IV).

Для моделювання таких процесів використовуються логістичні криві декількох видів, наприклад [4]:

$$y = K / (1 + ab^{-x}),$$

$$y = K / (1 + 10^{a-bx})$$

абу крива Перла-Рида

$$y = K/(1+ae^{-bx}). \quad (2)$$

Логістична крива (2) має точку симетрії, що збігається з точкою перегину. Для моделювання процесу мотивування візьмемо саме цю функцію.

4. ОПИС ІМІТАЦІЙНОЇ МОДЕЛІ МОТИВУВАННЯ ЧЛЕНІВ ІНТЕЛЕКТУАЛЬНОЇ КОМАНДИ

Розглянемо задачу мотивування інтелектуальної команди в такій постановці. Відома множина груп факторів, що впливають на трудову мотивацію $Y=\{y_j\}$, $j=1,m$. Також відома множина співробітників $X=\{x_i\}$, $i=1,n$. Кожен співробітник в ході проходження анкетування або будь-яких інших психологічних тестів висловлює ступінь задоволеності по кожній з груп факторів, що відображає його початкову мотивацію U_{ij}^0 .

Після розрахунку ступеня задоволеності співробітників і виявлення факторів мотивації, які задоволені в найменшій мірі, необхідно виробити заходи впливу на трудову мотивацію співробітників для підвищення рівня їх задоволеності.

Очевидно, що будь-яка цілеспрямована зміна рівня задоволеності співробітника потребує вкладення ресурсів, причому отриманий ефект від витрат однієї і тієї ж їх кількості для різних факторів мотивації буде різним.

Тоді задачу мотивування інтелектуальної команди можна розглядати в двох постановках:

- 1) визначення кількості необхідних ресурсів r_{ij} для повного задоволення співробітника x_i ;
- 2) прогнозування зростання мотивації співробітника x_i з урахуванням вкладеної кількості ресурсів r_{ij} .

Функцію залежності рівня задоволеності i -го співробітника від витрат ресурсів на j -ий фактор мотивації $U_{ij}(r_{ij})$ можна представити логістичною функцією виду:

$$U_{ij}(r_{ij}) = 1/(1+c_i e^{-b_{ij} r_{ij}}), \quad (3)$$

де r_{ij} – кількість ресурсів, що використовується для корекції значення рівня задоволеності i -го співробітника по j -тому фактору мотивації;

c_i, b_{ij} – відомі кількісні значення параметрів, що відповідають типу особистості.

Функція мотивації $U_{ij}(r_{ij})$ була змодельована для різних кількісних показників параметра b_{ij} , тобто психотипів особистості членів інтелектуальної команди. Результат моделювання графічно представлено на рис.3

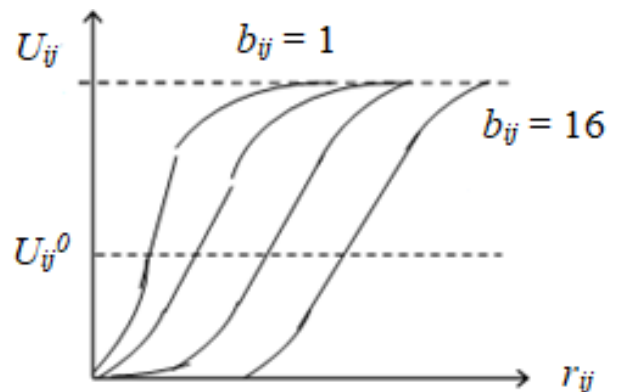


Рис. 3 – Графік залежності $U_{ij}(r_{ij})$ від кількості ресурсів r_{ij} при різних значеннях параметра b_{ij}

Судячи з наведеного графіку, в залежності від типу особистості для досягнення максимального рівня вмотивованості співробітників потрібне вкладання різної кількості ресурсу, але для деяких психотипів меншої, а для інших значно більшої.

Таким чином, маючи знання про психотипи членів інтелектуальної команди та виявивши фактори їх мотивації, можна змодельовати залежність рівня задоволеності співробітника від вкладеного ресурсу та спрогнозувати оптимальний рівень.

ЛІТЕРАТУРА REFERENCES

- [1] Петров Э.Г. Организационное управление городом и его подсистемами : (методы и алгоритмы) [Текст] / Э. Г. Петров. – Харьков: Вища шк., 1986. - 143 с.
- [2] Акофф Р., Эмери Ф. О целеустремленных системах [Текст] / Пер. с англ. под ред. И. А. Ушакова – М.: «Сов. радио», 1974. – 272 с.
- [3] Економіко-математичні методи та моделі (економетрика): навчальний посібник [Текст] / О.В. Козьменко, О.В. Кузьменко. – Суми: Університетська книга, 2014. – 406 с.
- [4] Математичні моделі та методи управління мотивацією персоналу: монографія [Текст] / А. О. Азарова, О. А. Ковальчук. – Вінниця: ВНТУ, 2014. – 140 с.

Інформаційно-Аналітична Служба як Елемент Кібербезпеки Держави

Миронюк Віталій
кафедра захисту інформації
Вінницький національний технічний університет
Вінниця, Україна
vitmir1001@gmail.com

Information-Analytical Service as an Element of Cybersecurity of the State

Mironyuk Vitaliy
dept. of Information protection
Vinnytsia national technical university
Vinnitsya, Ukraine
vitmir1001@gmail.com

Анотація – розглянуто основні особливості та функції інформаційно-аналітичної служби, запропоновано узагальнену структурну модель інформаційно-аналітичної системи

Ключові слова – кібербезпека держави, інформаційно-аналітична служба, інформаційно-аналітична система

Abstract—main specifics and functions of informational-analytical systems are reviewed. A generalized structural model of the information-analytical system is proposed.

Keywords— state cybersecurity, informational-analytical service, informational-analytical system

I. ВСТУП

Сучасні інформаційно-комунікаційні технології використовуються для реалізації технологій інформаційної війни та інформаційного протиборства, а кіберпростір використовується як специфічна арена для проведення спеціальних інформаційних операцій (СІО), які спрямовуються на різноманітні системи, зокрема соціоінформаційні (СІС), до яких відносяться, як локальні об'єкти захисту, наприклад, різні підприємства, установи і організації так і держава в цілому.

Забезпечення комплексної інформаційної безпеки держави передбачає створення систем захисту, які мають бути адаптовані і реалізовані на всіх рівнях управління інформаційною безпекою. Найважливішим компонентом системи оцінювання та забезпечення заданого рівня кібербезпеки є інформаційно-аналітична служба (ІАС) та відповідні інформаційно-аналітичні системи, які забезпечують користувачів своєчасною якісною інформацією по досліджуваному питанню.

З урахуванням головної мети функціонування інформаційно-аналітичної служби, а також задач, які покладені на координаційний центр з кібербезпеки, який входить у структуру РНБО України [1], зокрема таких, як:

- прогнозування та виявлення потенційних та реальних загроз у сфері кібербезпеки України;

- оперативне, інформаційно-аналітичне забезпечення Ради національної безпеки і оборони України з питань кібербезпеки;
- забезпечення розвитку інформаційної інфраструктури держави, узгодження і координації діяльності суб'єктів сектору безпеки і оборони, які забезпечують кібербезпеку України,

Актуальною задачею є розробка моделей інформаційно-аналітичної служби і відповідних інформаційно-аналітичних систем.

II. МЕТА ТА ГОЛОВНІ ФУНКЦІЇ ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СЛУЖБИ

Головною процедурою інформаційно-аналітичної роботи є проведення первинного аналізу та відбір важливої інформації. Ця процедура є своєрідним фільтром, який відкидає непотрібне, та захищає аналітика від інформаційної надмірності. Така процедура складається з, насамперед, встановлення суті, важливості, точності, повноти та значущості інформації на основі її розподілення та зіставлення [2].

Відповідно до стратегії кібербезпеки України, кіберзахисту державних електронних інформаційних ресурсів та інформаційної інфраструктури, призначеної для обробки інформації, вимога щодо захисту якої встановлена законодавством України, має полягати, зокрема, в [3]:

- упровадженні організаційно-технічної моделі національної системи кібербезпеки, оперативному реагуванні на кібератаки та кіберінциденти;
- розгортанні (відповідно до компетенції) єдиної системи ситуаційних центрів профільних органів державної влади сектору безпеки і оборони на базі захищеної інформаційної інфраструктури. Очевидно, що рішення вищенаведених задач можливо лише за умови організації і функціонування ефективної ІАС.

Можна виділити такі основні особливості та функції ІАС для виявлення СІО:

- єдиний спосіб надання інформації з різних джерел;
- пошук і навігація по всьому об'єму розподіленої бази даних СІО;
- комплексна аналітична обробка інформації.

Проведення інформаційно-аналітичної роботи передбачає виконання певних етапів, які можна представити у такій послідовності:[4]

1) *Загальне знайомство з СІО.* Цей етап передбачає вивчення поняття СІО з метою складання плану виконання робіт, основних джерел інформації, що можуть бути використані, а також виконавців.

2) *Визначення термінів і основних понять, що використовуються.* Виконання цього етапу дозволить ліквідувати різні толкування основних термінів і понять, що використовуються у процесі роботи.

3) *Збір фактів.* Підбір фактів може здійснюватись у таких режимах:

- неперервний режим або режим моніторингу, використовується тільки для збору найважливіших факторів і характеризується найбільшими затратами;
- фокусний режим - відповідає збору інформації по кожній окремій задачі;
- пакетний режим відповідає паралельній роботі з декількома задачами і потребує звернення до одних джерел інформації.

4) *Вивчення та обробка фактів.* На цьому етапі виконується оцінка, класифікація, аналіз та усвідомлення отриманої інформації.

5) *Побудова гіпотези.* Використовуючи результати виконання попереднього етапу пропонується робоча гіпотеза. Побудова робочої гіпотези пов'язана з формулюванням конкретних питань, відповіді на які власне дозволяють перевірити і саму гіпотезу.

6) *Висновки.* На цьому етапі проводиться робота, яка необхідна для доказу або спростування гіпотез висунутих на попередньому етапі. Тут формулюються кінцеві висновки, які є завершенням будь-якого аналітичного документа.

7) *Викладення.* Складання документа, який завершує роботу. Такий документ-звіт, може бути 3-х типів:

- тактичний або оперативний звіт, необхідний для термінового прийняття рішення;
- стратегічний звіт вміщує більш повну інформацію по відповідному питанню;
- періодичний звіт готується за певним графіком.

III. МОДЕЛЬ ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ

У даній роботі пропонується структурна модель інформаційно-аналітичної системи (рис. 1), що складається з трьох взаємопов'язаних елементів [5]:

1) Моделі структури метаданих, що забезпечує єдине відображення і опис всієї різномірної інформації;

2) Функціональної моделі, що описує поведінку та станів системи і способи переходу між ними.

3) Моделі розмежування доступу, що забезпечує управління доступом суб'єктів до інформаційних ресурсів.



Рис. 1. Структурна модель інформаційно-аналітичної системи

Будь-яка аналітична робота забезпечується комплексом технологій обробки інформації, які засновані на цілісній методологічній базі. Інформаційно-аналітичні технології можна визначити як систему знань, методів, операцій і правил, які дозволяють на основі залучення необхідних ресурсів забезпечити найбільшу ефективність того чи іншого виду діяльності.

Але для того, щоб інформаційно-аналітичні технології дійсно могли працювати в якості системи протидії інформаційній зброї, необхідно, щоб аналітик неперервно здійснював контроль за станом свого модельного світу.

ВИСНОВКИ

В протидію спеціальним деструктивним інформаційним операціям, які проводяться з метою зміни свідомості соціуму, дестабілізації та руйнування всієї СІС, існують засоби, які стабілізують та організують процеси мислення та цілеутворення. Саме таким засобом протидії СІО є інформаційно-аналітичні технології.

Запропонована структурна модель інформаційно-аналітичної системи може бути використана для побудови ефективної і надійної ІАС та системи управління кібербезпекою.

REFERENCES ЛІТЕРАТУРА

- [1] Положення про Національний координаційний центр кібербезпеки від 07.06.2016 № 242/2016 [Електронний ресурс]. Режим доступу: URL: <http://zakon.rada.gov.ua/laws/show/242/2016> - Назва з екрану.
- [2] В.И. Ярочкин, Я.В. Бузанова, «Корпоративная разведка», М.: Ось-89, 2004, с. 288.
- [3] Стратегія кібербезпеки України від 15.03.2016 № 96/2016 [Електронний ресурс]. Режим доступу: URL: - Назва з екрану.
- [4] Ю.В. Курносков, П.Ю. Конотопов, «Аналитика: методология, технология и организация аналитической работы», М.: РУСАКИ, 2004, с. 472.
- [5] В.А. Лужецький, О.П. Войтович, А.В. Дудатьєв, «Інформаційна безпека», навчальний посібник, Вінниця ВНТУ, 2009, с. 240.
- [6] Н. Г. Демурчев, Р. И. Касимов, «Математическая модель информационно-аналитической системы регионального управления», вестник Ставропольского государственного университета, 2009, № 63, с. 132.

Обробка Просторових Даних

Новожилова М.В.

кафедра ПМіІТ

Харківський національний університет міського

господарства

Харків, Україна

m.novozhilova04@gmail.com

Умріхін Ілля

кафедра ПМіІТ

Харківський національний університет міського

господарства

Харків, Україна

NeonQ@protonmail.ch

Prof. Marina Vladimirovna Novozhilova

Department of Applied Mathematics and Information

Technologies

Kharkiv National University

of Municipal Economy

Kharkiv, Ukraine

m.novozhilova04@gmail.com

Umrikhin Ilya

Department of Applied Mathematics and Information

Technologies

Kharkiv National University

of Municipal Economy

Kharkiv, Ukraine

NeonQ@protonmail.ch

Анотація - метою даної роботи є отримання нових знань в роботі з просторовими даними а також просторовими базами даних які дозволяють виконувати аналітичні запити, що містять просторові оператори для аналізу просторово-логічних відносин об'єктів.

Abstract- the purpose of this work is to gain new knowledge in working with spatial data, as well as spatial databases that allow you to perform analytical queries containing spatial operators for analyzing spatial and logical relations of objects.

Ключові слова – просторові дані; просторова база даних; геоінформаційні системи.

Keywords – spatial data; spatial database; geographic information systems

I. ВСТУП

З кожним днем збільшується загальний обсяг оброблюваної інформації, а з нею приходять і нові технології. З розвитком нових технологій взаємодіючих між собою за різними стандартами, які здійснюються на основі консенсусного процесу і доступні для всіх, вдосконалюються обмін всілякими даними в світі .

Просторові дані — це відомості про елементи включають в себе різні просторові набори. Цей термін є фундамент інформаційного супроводу геоінформаційних систем.

Безліч геоінформаційних систем, внесених, записаних різними способами, є просторовою базою даних. Актуальні бази даних створюються за допомогою професійного ПО, яке дозволяє обробляти збережені дані, включаючи в себе просторовий аналіз, все елементи якого управляються логічно загальної БД [1], [2], [3].

В даний час системи управління базами даних мають властивості і якості просторових розширень - геометричних типів даних і просторових індексів. На прикладі вільної реляційної системи управління базами даних - MySQL, який працює з різними типами даних, починаючи з точки, закінчуючи різними фігурами і їх поєднання в різних варіаціях між собою. Геометричні примітиви зберігаються в базі даних у вигляді бінарного коду, але також здатні бути показані в текстовому форматі WKT [4], [5].

```
POINT(15 20)
LINESTRING(0 0, 10 10, 20 25, 50 60)
POLYGON((0 0,10 0,10 10,0 10,0 0),(5 5,7 5,7 7,5 7, 5 5))
MULTILINESTRING((10 10, 20 20), (15 15, 30 15))
GEOMETRYCOLLECTION(POINT(10 10), POINT(30 30), LINESTRING(15 15, 20 20))
```

Рис. 1 — Геометричні примітиви в бінарному вигляді

Список систем, що взаємодіють з просторовими базами даних досить великий:

- Це всі продукти які задовольняють стандартам ГІС. ГІС — це міжнародна некомерційна організація, прихильна забезпеченню якісних відкритих стандартів для глобального геопространственного спільноти. У даній організації більше 500 чоловік з різних урядових, комерційних організацій, академічних та науково-дослідних інститутів.
- Відкриті API для роботи з просторових баз даних, частина з наявних задовольняє стандартам ГІС.

- Система керування базами даних - Microsoft SQL Server, з версії 2008.
- PostgreSQL в якому використовують просторове доповнення PostGIS для потрібного нам типу даних geometry.
- MySQL працює з типом даних geometry і декількома просторовими функціями, що не були зроблені відповідно до стандартів ГІС [6], [7].

Список систем, що працюють з просторовими БД



Рис. 2 — Список систем

ВИСНОВОК

В ході даної роботи ми отримали нові знання з просторовими даними. А зокрема роботу з різними базами даних, стандартами ГІС, відкритими API, системами

управління SQL Server, розглянули доповнення PostGIS в PostgreSQL, дізналися про роботу MySQL і типом даних geometry.

REFERENCES ЛІТЕРАТУРА

- [1] Модель пространственных данных для решения задач регионального управления [Электронный ресурс]// КБ ПАНОРАМА – Режим доступа: <https://bit.ly/2DNciHt>
- [2] Загальні відомості про системи автоматизованого проектування. [Електронний ресурс]// Підручники – Режим доступу: <https://bit.ly/2BBuGBK>
- [3] Пространственные данные [Электронный ресурс]// Вікіпедія. – Режим доступу: – <https://bit.ly/2FJXgoG>
- [4] Пространственная база данных [Электронный ресурс]// Вікіпедія. – Режим доступу: – <https://bit.ly/2RkaOID>
- [5] Orngeospatial [Электронный ресурс]// Вікіпедія – Режим доступу: <https://bit.ly/2A1leaZ>
- [6] Microsoft SQL Server [Электронный ресурс]// Вікіпедія – Режим доступу: <https://bit.ly/2OsKtuG>
- [7] Пространственные данные (SQL Server) [Электронный ресурс]// microsoft – Режим доступу: <https://bit.ly/2FHTTrjC>

Parallel Functional Data Processing in LISP

Orest Geiko

*Department of Computer Science
Vasyl Stefanyk Precarpathian
National University
Ivano-Frankivsk, Ukraine
ifgo69@gmail.com*

Artur Martsinkovskyi

*Department of Computer Science
Vasyl Stefanyk Precarpathian
National University
Ivano-Frankivsk, Ukraine
arthurmarz.learn@gmail.com*

Паралельна Функціональна Обробка Інформації на LISP

Орест Гейко

*кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет
імені Василя Стефаника
Івано-Франківськ, Україна
ifgo69@gmail.com*

Артур Марцінковський

*кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет
імені Василя Стефаника
Івано-Франківськ, Україна
arthurmarz.learn@gmail.com*

Abstract— This article is a review of some aspects of functional data processing schemes, that are used for parallel computation and creation of parallel and asynchronous applications based on referential transparency, pure functions, persistent data structures and data immutability using modern LISP dialect Clojure. This is also a consideration of features that functional data processing techniques such as Map/Reduce, and its applicability to the large datasets.

Keywords — *distributed programming, functional programming, execution in parallel.*

Анотація — В статті розглянуто деякі аспекти функціональних схем обробки даних, які використовуються для паралелізації обчислень з допомогою сучасного діалекту LISP - Clojure та формування асинхронних додатків на основі прозорості посилань, чистих функцій; персистентних структури даних та імутабельності даних. Розглянуто особливості функціональних алгоритмів для паралельної обробки гігантських наборів даних, таких як Map/Reduce.

Ключові слова— *дистрибутоване програмування, функціональне програмування, паралелізм.*

I. INTRODUCTION

From the 1980-s the orders of growth of hardware parameters like processor clock frequency, caches, and RAM size [1] of information systems have allowed us to use substandard software that lacked optimization. Software performance defects were neglected due to the fast growth in the computational power. The scale of data processed was not very huge. the nature of the tasks that computers had to perform was far from complex either permissible in the means of time of execution. Recent development and comprehensive nature of computer networks caused rapid change in the scale of data that needed to be transmitted, stored and processed. At the same time, orders of growth for capabilities of computers

are not as high as the need for them, that requires multithreading and clasterization efforts to be able to process data streams coming in [2]. This makes parallel execution a requirement for any modern system. Parallel execution can be regarded as a powerful tool for making data processing more efficient, but it also puts some additional requirements and brings up issues that have never been a problem in sequential programming. These problems mostly consist of modification and access issues that can come to life in any system that processes and stores data in parallel. Functional programming and algorithms are the efficient way to handle such issues.

II. RESEARCH RELEVANCE

Usual imperative paradigm, as well as object-oriented paradigm is not very suitable for the distributed data processing. Programs in these paradigms are sequential and have side effects, that can make an effect on the process of execution unexpectedly. On the other hand, functional programs when composed in the right fashion, can make parallelization much easier. There is a multitude of features that are unique to the functional paradigm: immutability of data, pure functions, referential transparency, recursive algorithms, functors, convolutions and higher order functions, persistent data structures. All these elements of functional programming let us to avoid some classes of errors that are usual for parallel execution, like deadlocks, data corruption on simultaneous write, race conditions. Clojure as a functional language with applicable data structures and inner structure applies well for the following techniques and algorithms on it benefit much from the way this LISP dialect is built. Also, functional algorithms are better at scaling and distribution amongst multiple computers. The main elements of functional programming that provide us with advanced parallelization techniques are [2]:

- Pure functions. Pure functions are functions that do not have IO side effects and memory (they are stateless, which means that they don't have any state and return same values for different calls with same arguments). They allow reinterpretations and optimizations on the compile (interpreter processing) stage. The result of pure function call can be cached in a hash table, allowing substantial speed-up for recursive algorithms. Also, such functions can be considered thread-safe.
- Referential transparency. Referential transparency is one of the fundamental principles of the functional programming; only referentially transparent functions can be memorized (transformed into the equivalent functions that use cached results). Some languages provide the programmer with tools to guarantee this structure. Some other on the other hand require referential transparency in all functions. Since these features require same output for same input in any time, the referentially transparent expression is determined by the definition. This allows the programmer to avoid some extra calculations that frequently appear in the flow of parallel programming.[6]
- Persistent data structures. Persistent data structures are kind of data structures that retain access to the previous versions. If we have the sequence of p elements and we must change the element n in it, we create the new version of the structure that is different from the previous only by the value of the nth element. As a result, we will have two versions of the sequence with access to each one of them. This technique lets the programmer keep the data structure fast and immutable at the same time sparing memory by reusing it.
- Functors. Any class or data type, that stores values, and implements method map is called a functor. Also, functor should return the collection with the same type to the collection passed to it [7]. For example, array is a functor, because an array is data structure that stores values and implements method map that allows us to apply the function against the values it stores. A benefit from such approach to the parallelization lies in the fact that transformation of the elements is independent which means that the sequence can be divided into ranges each of which would be transformed in parallel.
- Monads. The main application of monads in the functional programming is the isolation of the IO and stateful behavior. The principle of monad application is the fact that function that don't create the side effect itself, it can create and can be used later. But IO and presence of state are not the only use cases for monads. They are useful when the programmer wants to describe the calculations in purely functional manner, at the same time executing other calculations in parallel. [4] Thus, two main use cases of monads

are sequential calculations and handling of the inherently side-effect generated data.

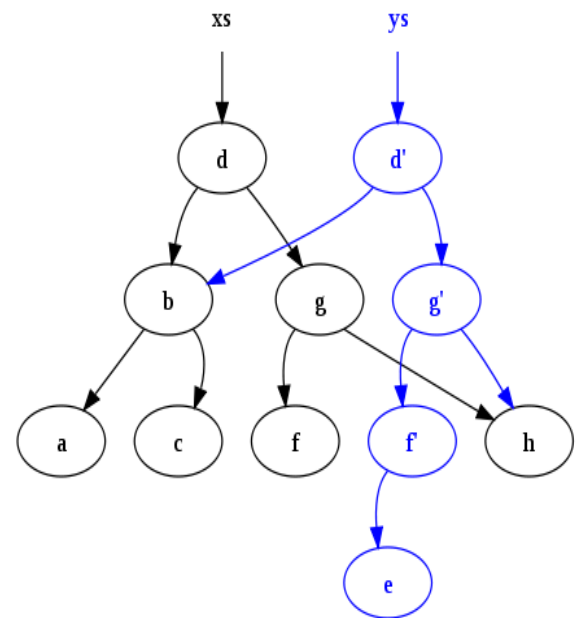


Fig 1. Representation of the work of persistent data structures (hash tables, vectors, lists, etc.)

III. PRACTICAL IMPLEMENTATION OF FUNCTIONAL PROGRAMMING IN PARALLELIZATION

The most famous use of the ideas of functional programming in parallel for relatively fast analysis and processing of huge amounts of data is the Google MapReduce algorithm.

This algorithm consists of two successive steps: the clustering of functors to an array of data and subsequent convolution, map steps and reduces accordingly (there can be more than one functor, for example, on Figure 2. there are three of them).

In the map step, a conversion is applied to each element of an array of input data that changes its contents but retains the number of elements. In the case of distributed architecture, for the map step, the host computer receives the input data array, marks it and distributes it between its nodes by specifying the transformation. This process can take place several times before the convolution.

In the reduce step, preprocessed data gets convoluted. The main node receives responses from the working units and, based on them, forms the result - a solution to the problem that was initially formed.

The advantage of this algorithm is the possibility of its almost unlimited scaling, the absence of a global state and the possibility of an iterative data flow with multiple transformations and reductions, of which even the hierarchical system can be built [4].

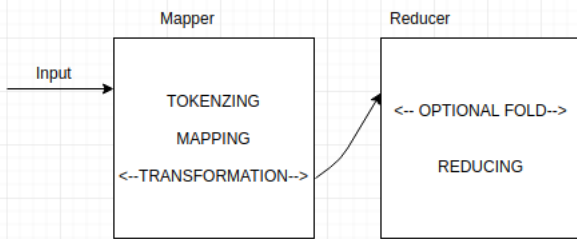


Fig. 2. Listing an application of the Map Reduce algorithm for counting words in incoming texts with the ability of parallel procession on Clojure

Algorithm 1

```
(require '[clojure.string :as str])
;; Function for processing nodes on the Map
step.
(defn word-filter[words]
  (filter #(re-matches #"\\w*" %) words))
;; Function that is used by processing nodes on
the Map step
;; for data transformation into indexed original
form
(defn sanitize-map[words]
  (map #(.toLowerCase %) words))
;; Function for Reduce step that is used for
data array convolution in the occurrence hash
map form.
(defn count-reduce[words]
  (reduce #(assoc %1 %2 (inc (%1 %2 0)))
    {}
    words))
;; Reduce step function for the final merge and
accumulation of values of the first-step
convolutions
(defn merge-count-reduce[occurrence-maps]
  (reduce #(merge-with + %1 %2)
    (hash-map)
    occurrence-maps))
(defn map-reduce[documents]
  (merge-count-reduce
    (map
      (comp
        count-reduce
        sanitize-map
        word-filter) documents)))
(def docs [["a" "." "c"] ["mine" "token" "is"
"here" "a"]])
(map-reduce docs)
```

CONCLUSION

Functional programming is an effective paradigm for software development, which is becoming more and more important with the gradual deprecation of the Moore's law and a significant increase in the number of processed information. Pipelining, which can be achieved with state-of-the-art functionality, the ability to cache and memorize the results of pure functions, the advantage of persistent data structures that occupy significantly less memory and the unchangeability of data with the isolation of the data in the individual elements of the program allows you to develop faster and less prone to program mistakes that allow you to avoid problems that are traditionally encountered during parallelization.[3] The functional approach allows you to create abstractions based on data and perception of functions in the form of data, which

contributes to abstraction and allows you to use the mathematical apparatus organically when solving practical problems.

However, on the other hand, functional programming is now faced with several problems that make it difficult to use in the industry from a commercial point of view. First, this is an inadequate qualification of most programmers who traditionally use the imperative paradigm of programming and its derivatives in the form of object-oriented languages. Also, functional programming requires a lot of training in connection with many fundamentally new concepts and theoretical constructions that take time to soak in the mind or a regular software engineer. Compilers and interceptors of such languages are considerably more complex because they require the implementation of different calculation procedures simultaneously, the implementation of high-speed garbage collector (GC) and the distribution of clean and "contaminated" data zones [8].

Another problem with functional programming is the complexity of simulating simulations and fast-changing systems that make lazy computing and immutability a burden rather than an advantage in connection with the constant change of contexts and the complex indivisible interconnections of such systems [5,6].

In general, the prospect of using functional programming widely in the industry and academic environment is likely to be the gradual incorporation of the mature and appropriate elements of functional programming in more common programming development languages, as well as the formation of the appropriate ecosystems and communities of purely functional programmers with respect to tasks that are most in line with the rules. and points of benefit from the use of this type of language.

REFERENCES ЛИТЕРАТУРА

- [1] C. Okasaki "Purely Functional Data Structures" Cambridge University Press 1998
- [2] G. Cousineau "The Functional Approach to Programming" Cambridge University Press 1998
- [3] T. VanDrunen "Discrete Mathematics and Functional Programming" Paperback – October 16, 2012
- [4] G. Michaelson "An Introduction to Functional Programming Through Lambda Calculus (Dover Books on Mathematics)" Copyright, 1989
- [5] B. Jay "Pattern Calculus: Computing with Functions and Structures" Brodway NSW 2007
- [6] G. E. Revesz "Lambda-calculus, Combinators and Functional Programming (Cambridge Tracts in Theoretical Computer Science)" Cambridge University Press 1998
- [7] Ben Vandgrift "Clojure Applied. From practice to practioneer" Pragmatic Programmer Bookshelf 2015
- [8] Michael Linn Functional programming patterns in Scala and Clojure Pragmatic Programmer Bookshelf 2013

Обробка Великих Даних за допомогою ELK-стеку

Ольга Заверуха
кафедра Прикладної математики
Харківський національний університет
радіоелектроніки
Харків, Україна
olha.zaverukha@nure.ua

Володимир Кобзєв
кафедра Прикладної математики
Харківський національний університет
радіоелектроніки
Харків, Україна
volodymyr.kobziev@nure.ua

Big Data Processing Using the ELK Stack

Olha Zaverukha
Department of Applied Mathematics
Kharkiv National University
of Radioelectronics
Kharkiv, Ukraine
olha.zaverukha@nure.ua

Volodymyr Kobziev
Department of Applied Mathematics
Kharkiv National University
of Radioelectronics
Kharkiv, Ukraine
volodymyr.kobziev@nure.ua

Анотація—Дана стаття присвячена системі швидкої та якісної обробки Великих даних - ELK-стек. У статті розглянуто переваги цього стеку та приклади зберігання та візуалізації даних за його допомогою.

Ключові слова—Великі дані, стек, обробка, зберігання, візуалізація

Abstract—This article is devoted to the system of fast and high-quality processing of Big data - ELK-stack. The article outlines the advantages of this stack and examples of storing and visualizing data using it.

Keywords—Big Data, stack, processing, storage, visualization

I. ВСТУП.

У сучасному світі широко розповсюджено збір та аналіз даних різної природи. За останні десятиліття складність та об'єм даних зростають з неймовірною швидкістю. Тому, збір та аналіз наявної інформації являють собою суттєву проблему на етапі зберігання великого об'єму даних. Поняття «Великі дані» для ІТ-індустрії з'явилося у 2011 році. З появою Великих (великих за об'ємом, неструктурованих і швидко зростаючих) даних виникла потреба обробки цих даних новими методами з використанням більш потужних технологій, тому створюються нові та вдосконалюються раніше відомі мови програмування. У багатьох публікаціях викладені описи проблем і моделей Big data та підходи до їх практичного вирішення.

Так у 2010 році створена потужна система гнучкого пошуку, зберігання та візуалізації даних ELK (Elasticsearch, Logstash, Kibana). Протягом останніх п'яти років вона стає дуже популярною серед аналітиків

Великих даних завдяки своїм можливостям, простоті освоєння і застосування [1, 2].

II. СКЛАД І ВЗАЄМОДІЯ ПІДСИСТЕМ ELK

Стек ELK складається з трьох головних компонентів, які наведені на Рис. 1.



Рис. 1. Структура підсистеми ELK

Підсистема Logstash має змогу передати дані, вилучити їх з бази даних, з файлу, або з встановленого клієнту, що збирає логи, а також може перетворювати їх на потрібний формат чи структуру та завантажити у будь-яке сховище даних. Крім того, вона дозволяє у будь-який час передати дані до підсистеми Elasticsearch.

За своїми функціями Logstash – це доволі простий інструмент, який передає дані з одного чи багатьох входів та виводить їх на один, чи декілька виходів. На входах можуть бути будь-які джерела даних, такі як файли журналів або логи. Після того як вхідні дані будуть зчитані, logstash аналізує їх за допомогою кодків, таких як JSON [2].

Коли дані переміщуються з джерела в сховище, фільтри Logstash аналізують кожну подію, ідентифікують назви полів для побудови структури та

перетворюють їх у загальний формат для полегшення, прискореного аналізу. Logstash динамічно трансформує та готує дані незалежно від формату або складності: створює структуру з неструктурованих даних, розшифровує географічні координати, анонімізує дані. При цьому характерною є простота обробки даних незалежно від джерела даних, формату чи схеми.

Можливий вигляд структури файлу конфігурації Logstash наведений на Рис.2.

```

1 input{
2   file {
3     path => "/tmp/access_log"
4     start_position => "beginning"
5   #Перша частина config файлу складається з команди, що зчитують інформацію з потрібного джерела даних
6   }
7 }
8
9 filter{
10  grok {
11    match => { "message" => "%{MESSAGE}" }
12  }
13  date {
14    match => [ "timestamp", "dd/MMM/yyyy:HH:mm:ss.SSS" ]
15  }
16  # Друга частина config файлу включає в себе фільтрування даних.
17 }
18
19 output{
20  elasticsearch {
21    hosts => ["localhost:9200"]
22  }
23  # Третя частина config файлу складається з виведення даних в необхідне сховище.
24 }
25

```

Рис. 2. Приклад файлу конфігурації Logstash.

Крім того, Logstash відстежує введені дані, які він обробляє. Тому його можна перезапустити без ризику дублювання даних. Хоча Logstash написано в Ruby, працює він дуже швидко. Пакетна версія запускається на JRuby, і вона використовує можливості JVM в потоці, викидаючи десятки потоків для паралелізації обробки даних.

Logstash може зберігати дані у блоці Elasticsearch, який також запущений на машині, де працює Logstash. Elasticsearch виконує роль сховища даних. Elasticsearch – це не база даних, яка може сприймати все, що користувач забажає помістити до неї, але він масштабує дані, коли це потрібно.

Стан Elasticsearch [3] - це коротке, однорядкове відображення інформації (/_cluster / health) (див. Рис.3) про статус, очікування інформації, зайнятий обсяг пам'яті, тощо.



Рис. 3. Відображення стану Elasticsearch у браузері.

За допомогою простої команди (/_cat/indexes) можна побачити наступну інформацію (рис. 4): стан індексу, статус, назву, об'єм, кількість рядків, розмір на диску, тощо.

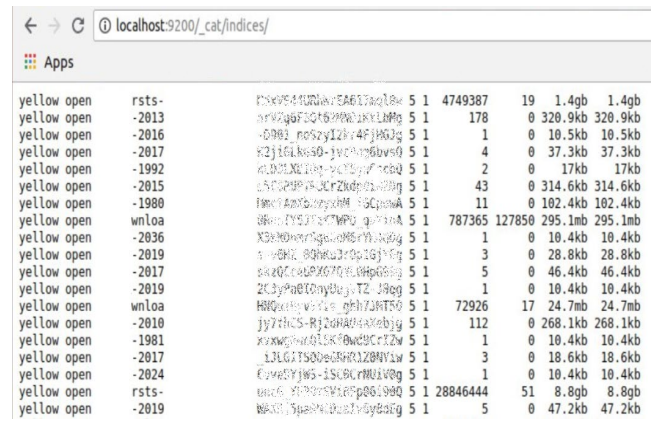


Рис. 4. Відображення індексів Elasticsearch та інформації про них у браузері.

Останнім кроком є відображення результатів аналізу. Kibana вже пов'язана з ElasticSearch, тому всі індекси вже знаходяться в ній, тож треба лише вибрати той, який у подальшій роботі необхідний для аналізу.

Kibana має добрий інтерфейс для візуалізації величезної кількості даних і дозволяє краще зрозуміти дані через аналіз та візуалізацію, що можна створити за лічені хвилини. Незважаючи на простоту даних інструментів, вони дуже потужні. Наявність шаблонів розгортання дозволяє легко вибрати апаратні та архітектурні профілі, які найкраще відповідають потребам користувача.

Kibana - єдиний компонент, який не резервується, так як вихід його з ладу не впливає на процеси збору і обробки інформації.

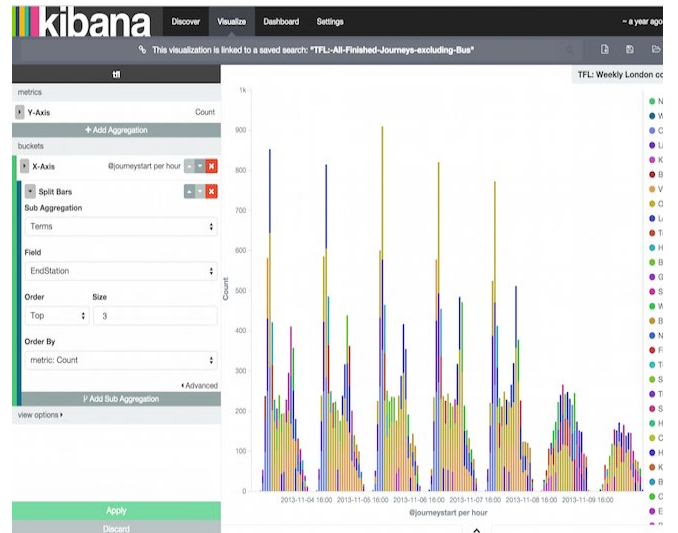


Рис. 5. Приклад однинчої візуалізації за допомогою Kibana.

Kibana дозволяє [4] будувати аналіз за такими напрямками, як геолокація, кількість, часові проміжки, тощо (рис.5). Окрім цього, всі зображення мають гарну графіку. Їх можна об'єднувати у інформаційні панелі dashboards, завдяки чому весь необхідний аналіз може бути зібраний на одній сторінці (див. рис.6).

ELK-стек призначений для вистежування інформації у реальному часі, тому після запуску усіх компонент, якщо джерело входу інформації буде поповнюватися новими даними, усі компоненти будуть відстежувати їх, додавати у сховище та відображати на візуалізаціях без перезавантаження їх самих. Це дуже зручно, якщо дані надходять постійно.

III. ЗАСТОСУВАННЯ ELK-СТЕКУ .

Для наглядної демонстрації роботи ELK-стеку обрані дані про країни, які люблять відвідувати

туристи. Рис. 7 свідчить, що кількість туристів, які прибули за один день до Індії, Таїланду, Австралії та Японії переважає показники інших країн.

Крім відвідуваності, також є статистика запитів в інтернеті про найпривабливіші країни для туристів. Кількість запитів на день переважає у Таїланду, про що свідчать дані наведені на рис.8. Також важливо враховувати не тільки запити, але і реальні дані про країни та їх міста, які популярні серед туристів, про що гарно свідчить подвійна кругова діаграма. На рис.9 країни та міста умовно позначені числами.



Рис. 6. Приклад панельної візуалізації за допомогою Kibana.



Рис. 7. Візуалізація кількості відвідувань країн за один день.

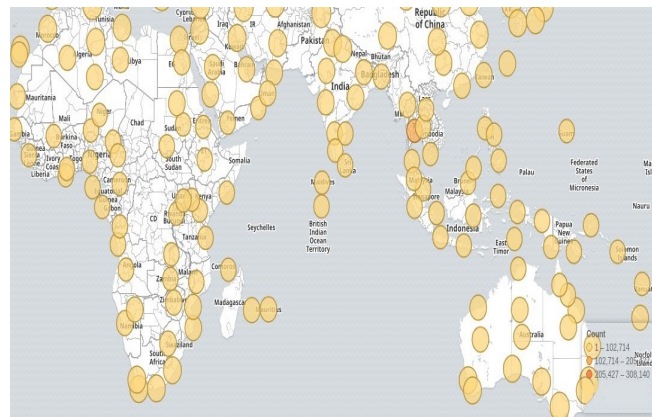


Рис. 8. Кількість запитів в інтернеті щодо подорожей до країн за один день

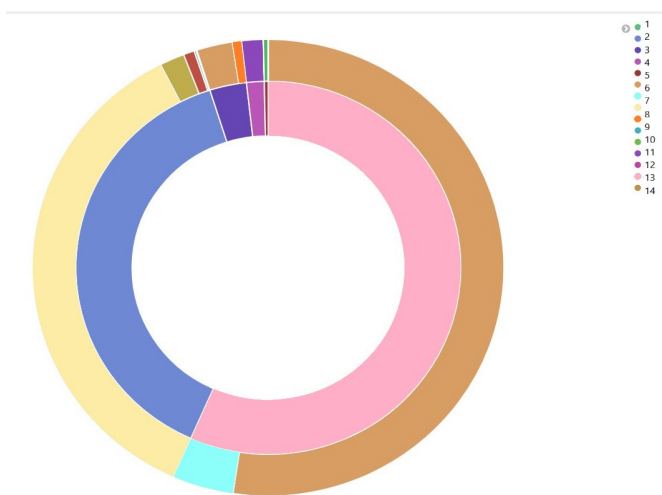


Рис. 9. Кругова діаграма популярності міст та країн.

Швидкість роботи та візуалізації за допомогою ELK-стеку дійсно вражає. Аби завантажити дані та побудувати зрозумілі графіки, знадобилося 2 години. Притому, що кількість даних сягала приблизно 30 мільйонів. ELK-стек дозволяє не тільки будувати

кругові або стовпчасті діаграми, але й дає змогу чітко візуалізувати дані з географічними полями на детальній мапі світу. Окрім яскравої та зрозумілої графіки, Kibana може виводити згруповані та відсортовані дані у таблицях, метриках або строках.

ВИСНОВОК.

ELK-стек є доступним, високоякісним, надійним і простим у обслуговуванні інструментом, який надає можливість за лічені хвилини доставити, впорядкувати та візуалізувати Великі дані. Дана система надає нові можливості для створення комерційних звітів (унікальні підрахунки, послідовності переходів тощо). Вона добре захищена та швидко працює, що дуже важливо для систем обробки Великих даних.

REFERENCES ЛІТЕРАТУРА

- [1] The Elastic Stack 6.5 [Online]. Available: <https://www.elastic.co/>
- [2] Big data in minutes with the ELK Stack [Online]. Available: <https://brewhouse.io/blog/2014/11/04/big-data-with-elk-stack.html>
- [3] Saurabh Chhajer "Learning ELK Stack", 2015, Packt Publishing.
- [4] Google Trends Analytics Data [Online]. Available: <https://trends.google.ru/trends/?geo=RU>

The Expert System for Monitoring and Preventing Natural Emergency Situations on Automobile Highways

Leonid Nefedov
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
nefedovli@i.ua

Kononuhin Aleksandr
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
makonon@i.ua

Fil Nataliya
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
fil_nu@i.ua

Binkovskaya Angela
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
ab.binkovskaya@gmail.com

Експертна Система Моніторингу і Попередження Природних Надзвичайних Ситуацій на Автомобільних Дорогах

Леонід Нефьодов,
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
Харківський національний автомобільно-дорожній
університет
Харків, Україна
nefedovli@i.ua

Кононіхін Олександр
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
Харківський національний автомобільно-дорожній
університет
Харків, Україна
makonon@i.ua

Філь Наталія
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
Харківський національний автомобільно-дорожній
університет
Харків, Україна
fil_nu@i.ua

Бінковська Анжела
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
Харківський національний автомобільно-дорожній
університет
Харків, Україна
ab.binkovskaya@gmail.com

Abstract—The article describes an expert system for monitoring and preventing natural emergency situations on automobile roads. A method for generating fuzzy SQL queries, based on standard language constructs, is developed. An example of constructing a fuzzy logic based selection of projects on preventing natural emergency situations on automobile roads is considered. The direction for further work is related to the development of an intelligent interface for the expert system for monitoring and preventing natural emergency situations on automobile highways.

Keywords—natural emergency situations, fuzzy logic, fuzzy set, membership function

I. INTRODUCTION

Modern automobile highways (AHs) are the driver of development of the Ukrainian economy. To increase the volume of transportation and improve the investment climate, it is necessary to create a modern network of AHs.

However, the condition of a vast majority of AHs with respect to their use for traffic is unsatisfactory. Moreover, 97% of all AHs are completely worn down and require capital repair or maintenance. Ukraine ranks 132nd in the

world in terms of the level of development of AHs. In addition, more than 9.6 thousand bridges in terms of technical parameters do not meet the requirements of modern standards and actual load, 1923 bridges require urgent repairs, and 86 bridges are in unsafe condition. Such a position of public roads and bridges does not ensure to the full extent a speedy, comfortable, economical, and safe transport of passengers and freight, development of transit traffic; creates social tension; reduces the competitiveness of the Ukrainian economy; hinders the further socio-economic development of Ukraine and its integration into the European community. To solve the problem, the Government of Ukraine approved the Concept of the state target economic program for the development of public roads for 2018-2022 [1].

In the last decade, there observed a steady trend of increasing anthropogenic pressures on the environment, land degradation, global climate change. All these causes activation and intensive development of dangerous natural processes and phenomena, increases the vulnerability of the population and territories to natural emergency situations (NES). The insufficient efficiency and development of

systems for monitoring potentially dangerous natural processes and phenomena, absence or poor condition of protective engineering structures are also important factors that affect material losses, and in some cases also human ones [2].

The implementation of preventive measures, projects, programs and strategies for protecting the population and territories reduces potentially negative consequences of emergency situations. The strategy for preventing NESs is aimed at solving the following tasks: prevention of causes of NESs; prevention of NESs and maximum mitigation of their consequences. The unstable economic situation in Ukraine makes it difficult to fully protect the population and territories from NESs. In addition, all existing projects on preventing NESs are very expensive and require significant financial, material, technical and human resources.

Particular attention should be paid to monitoring and preventing NESs. The problem of monitoring and predicting potentially dangerous sources of possible NESs can be solved by conducting multi-criteria analysis of various situations and making scientifically based effective decisions.

The main tasks of the system for monitoring the threat of NESs on AHs are [2]: conducting observations of sources of NESs and AHs; collecting, processing and analyzing information on sources of NESs and AHs; creating a data bank on sources of NESs and condition of AHs; predicting NESs on AHs; providing government bodies and other state organizations subordinated to the Government of Ukraine, local executive and regulatory bodies with information on the threat or occurrence of NESs on AHs [2].

II. PROBLEM STATEMENT

Based on the monitoring and expertise of sources of NESs on AHs and condition of the roads, there identified the AR sections and roadside facilities that require implementation of preventive measures and works considered as a set of possible projects on preventing NESs on AHs.

All of these projects define a strategy for preventing NESs on AHs. The strategy for preventing NESs is aimed at creating sustainable development conditions and ensuring a significant mitigation of risk and decrease of losses from NESs.

Due to the lack of financial resources for the implementation of the whole range of measures to minimize the risk of occurrence of NESs, the task of optimal distribution of the allocated financial, material, technical and human resources among all possible projects on preventing NESs on AHs comes to the fore.

For this purpose, the evaluation and ranking of possible projects on preventing NESs on AHs is performed, with regard to multi-criteria analysis of all possible projects on preventing NESs on AHs, based on expert information, i.e., their advantages and disadvantages are considered from the system point of view. The evaluation of possible projects on preventing NESs on AHs can be carried out according to the following criteria: frequency of occurrence of NESs on the given section of the AR; damage in the event of NES on the given section of the AR; capital costs for the implementation

of works to prevent NESs on the given section of the AR, etc.

For carrying out multi-criteria evaluation of possible projects on preventing NESs on AHs, it is suggested to use the hierarchy analysis method, which, based on expert data, allows ranking projects on preventing NES by their significance based on the given criteria.

The prediction and prevention of NESs is currently being implemented largely on the basis of poorly formalizable and completely non-formalizable knowledge, which are the result of years of observations, experience and intuition of various specialists [3]. Under these conditions, expert systems (ESs) are an important tool to solve problems of modeling and predicting dangerous events.

ESs for monitoring and preventing NESs on AHs are sophisticated software packages intended for collecting, processing and generalizing poorly formalizable knowledge of experts. Therefore, when developing a knowledge base for ESs, it is necessary to use the fuzzy logic technique that will allow formalizing the procedure for evaluating sources of NESs on AHs and make scientifically based decisions to prevent NESs on AHs. A fuzzy ES for monitoring and preventing NESs on AHs uses knowledge representation in the form of fuzzy products and linguistic variables [3].

Figure 1 presents a fuzzy logic structural model of an ES for monitoring and preventing NESs on ARs [3].

III. THE METHOD FOR FUZZY SELECTION OF PROJECTS ON PREVENTING NESs ON AHs

One of the elements of the system under consideration is the knowledge base. It is necessary to develop means to form a strategy, which contains projects on preventing NESs on AHs, based on fuzzy logic inference in the relational database. Fuzzy logic search and inference is based on the theory of fuzzy sets and is implemented through fuzzy SQL queries [4]–[7].

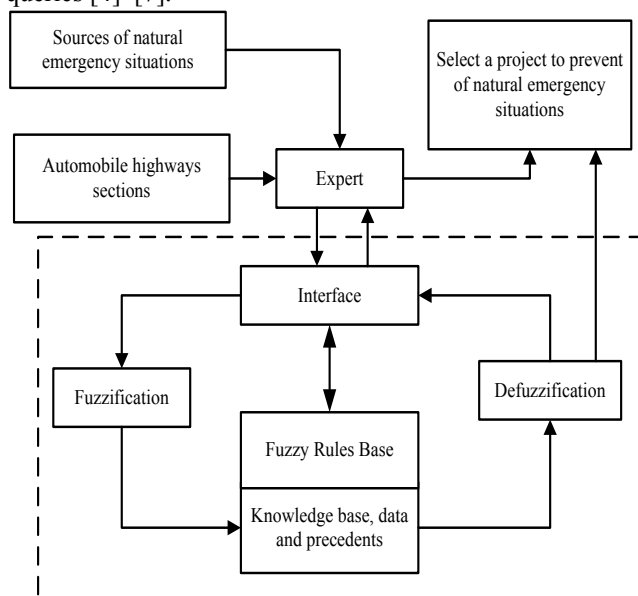


Fig. 1 The fuzzy logic structural model of an ES for monitoring and preventing NESs on ARs

The final decision-making takes into account the whole set of criteria for the implementation of measures, even, at

first glance, insignificant ones, which makes the decision most suitable to the reality.

After the synthesis of resulting (generalized) estimates and the ranking of projects, it is necessary to formulate a strategy for preventing NESs on AHs for the future, representing the portfolio of projects.

Based on the received portfolio of projects on preventing NESs on AHs, it is necessary to form a portfolio of projects on preventing NESs for the current period, which requires more detailed and accurate information.

In work [2] the information and reference support for the technique for monitoring and preventing NESs on AHs was developed. Information on sources of possible NESs, operating conditions is stored in relational databases. There is also a database of possible projects on preventing NESs on AHs. When developing a strategy for preventing NESs on AHs and a program for preventing NESs on AHs, it is necessary to select from the database of possible projects those that have the best value/cost ratio. This choice is supposed to be carried out using fuzzy queries.

The method of fuzzy data selection is based on the theory of fuzzy sets, which was first proposed by Lotfi Zadeh [9]-[11]. To implement fuzzy queries, it is necessary to formally describe the sets of values of terms in a logical expression. The model of such structures is the linguistic variable (LV) corresponding to the used fuzzy concepts of “low cost”, “high value” [4]. In this case, there should be defined universal sets of these variables, sets of basic terms, and membership functions for each term, taking into account the requirements imposed on them [4].

Let us consider an example of a fuzzy selection of projects on preventing NESs on AHs. Fuzzy data selection in the relational database is performed through SQL queries. All information about possible projects on preventing NESs on AHs is stored in the database in the table TProject (Tbl. 1), which has the following columns: id – the number of the projects on preventing NESs on.

AHs, TWorth – the value of the projects on preventing NESs on AHs, TValue - cost (UAH ths).

TABLE I. INFORMATION ON PROJECTS ON PREVENTING NESs ON AHs

id	TWorth	TValue
1	0,20	60
2	0,17	35
3	0,16	50
4	0,11	18
5	0,1	45
6	0,07	18
7	0,06	30
8	0,05	6
9	0,04	25
10	0,03	25

To implement fuzzy queries, we formalize two linguistic variables: “Cost” and “Value”. The linguistic variable “Value” has the following form:

To implement fuzzy queries, we formalize two linguistic variables: “Cost” and “Value”. The linguistic variable “Value” has the following form:

$$\langle \beta, \Psi, T, G, M \rangle, \quad (1)$$

where β - “Value”; $\Psi = [0.03; 0.20]$ - the scope of definition of the linguistic variable; $T = \{“high”, “medium”, “low”\}$; G - the syntactic procedure for the formation of new terms with the help of the connectives “AND”, “OR” and the modifiers “NO”, “VERY”, “MORE OR LESS”; M - the procedure for setting on the set Ψ the fuzzy sets $A = “low”, B = “medium”, C = “high”,$ and fuzzy sets for terms according to the rules for translating fuzzy connectives and modifiers and other operations on fuzzy sets.

Let us define trapezoidal membership functions for the variable “Value”:

$$\mu_A(x) = \begin{cases} 1, & x \in [0, 0,03] \\ (0,06 - x) / 0.03, & x \in (0,03, 0,06] \\ 0, & x > 0,06 \end{cases} \quad (2)$$

$$\mu_B(x) = \begin{cases} 0, & x < 0,05; x > 0,12 \\ (x - 0,05) / 0.03, & x \in [0,05, 0,08] \\ 1, & x \in [0,08, 0,1] \\ (0,12 - x) / 0.02, & x \in (0,1, 0,12] \end{cases} \quad (3)$$

$$\mu_C(x) = \begin{cases} 0, & x < 0,11 \\ (x - 0,11) / 0.04, & x \in [0,11, 0,15] \\ 1, & x \in [0,15, 0,2] \end{cases} \quad (4)$$

The graphs of the membership functions are shown in Figure 2.

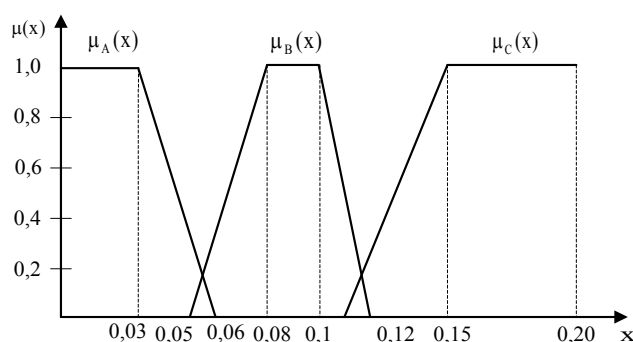


Fig. 2. The membership function of terms of the variable “Value”

The linguistic variable “Cost” has the form:

$$\langle \gamma, \Gamma, T, G, M \rangle, \quad (5)$$

where γ - “Cost”; $\Gamma = [6; 60]$; $T = \{“low”, “medium”, “high”\}$; G - the syntactic procedure for the formation of new terms with the help of the connectives “AND”, “OR” and the modifiers “NO”, “VERY”, “MORE OR LESS”; M - the procedure for setting on the set Γ the fuzzy sets $D = “low”, E = “medium”, F = “high”,$ and fuzzy sets for terms $G(T)$ according to the rules for translating fuzzy connectives and modifiers and other operations on fuzzy sets.

Let us define trapezoidal membership functions for the variable “Cost”:

$$\mu_D(y) = \begin{cases} 1, & y \in [6;15] \\ (20 - y)/5, & y \in (15;20); \\ 0, & x > 20 \end{cases} \quad (6)$$

$$\mu_E(y) = \begin{cases} 0, & y < 18; y > 40 \\ (y-18)/4, & y \in [18;22); \\ 1, & y \in [22;35] \\ (40 - y)/5, & y \in (35;40] \end{cases} \quad (7)$$

$$\mu_F(y) = \begin{cases} 0, & y < 38 \\ (y-38)/7, & y \in [38;45); \\ 1, & y \in [45;60] \end{cases} \quad (8)$$

The graphs of the membership functions are shown in Figure 3.

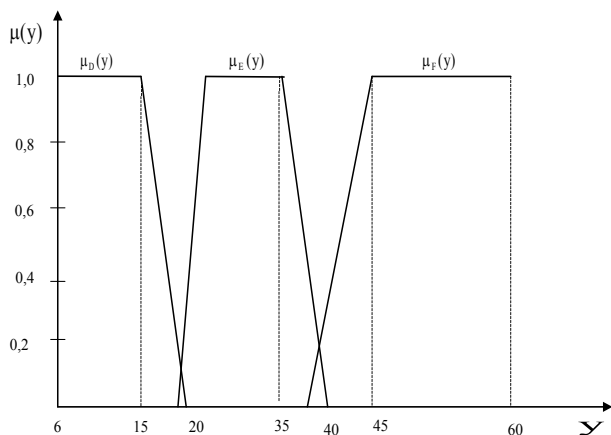


Figure 3 - The membership function of terms of the variable “Cost”

Table 2 presents the projects and the calculated degrees of membership.

TABLE II.
PROJECTS AND THE CALCULATED DEGREES OF MEMBERSHIP

id	TWorth	TValue	TWorth			TValue		
			μ_A	μ_B	μ_C	μ_D	μ_E	μ_F
1	0,20	60	0	0	1	0	0	1
2	0,17	35	0	0	1	0	1	0
3	0,16	50	0	0	1	0	0	1
4	0,11	18	0	0,5	0	0,4	0	0
5	0,1	45	0	1,0	0	0	0	1
6	0,07	18	0	0,67	0	0,4	0	0
7	0,06	30	0	0,33	0	0	1	0
8	0,05	6	0,33	0	0	1,0	0	0
9	0,04	25	0,67	0	0	0	1,0	0
10	0,03	25	1,00	0	0	0	1,0	0

Let us consider a fuzzy selection in the form of the SQL query “to find all projects of medium value and low cost”:

```
select * from TProject where (TWorth = "средняя" and
TValue = "низкая");
```

The fuzzy selection result is presented in Tabl. 3.

In this query, the two parameters are determined by the degree of membership and the operator “and” is defined as the Cartesian product of fuzzy sets.

Fuzzy queries can be used effectively to select information from relational databases using qualitative criteria and fuzzy conditions/

TABLE III THE RESULT OF A FUZZY SELECTION OF PROJECTS ON PREVENTING NESS ON AHs

id	TWorth	TValue	TWorth			TValue		
			μ_A	μ_B	μ_C	μ_D	μ_E	μ_F
4	0,11	18	0	0,5	0	0,4	0	0
6	0,07	18	0	0,67	0	0,4	0	0

CONCLUSION

In the article, a structural model of the expert system for monitoring and preventing natural emergency situations on automobile highways is developed, which, unlike known ones, is based on fuzzy logic. The method for fuzzy selection of projects in the expert system for monitoring and preventing natural emergency situations on automobile highways is elaborated, which, unlike known ones, is based on the value/cost fuzzy evaluation of possible projects on preventing natural emergency situations on automobile highways. This allows to adequately assess sources of natural emergency situations and condition of automobile highways and on this basis choose effective projects. An example of constructing a fuzzy query for the selection of projects on preventing natural emergency situations on automobile highways is considered. The processing of fuzzy queries results in an intelligent search, when the exact request attributes are fuzzy or undefined. This significantly simplifies the work with large volumes of various data for experts. The direction of further work is related to the development of an intelligent interface for the expert system.

REFERENCES ЛІТЕРАТУРА

- [1] Pro shkvalennia Kontseptsii Derzhavnoi tsilovoi ekonomichnoi prohramy rozvytku avtomobilnykh dorih zahalnoho korystuvannia derzhavnoho znachennia na 2018-2022 roky. Available at: <http://zakon3.rada.gov.ua/laws/show/34-2018-p>. Accessed: 2 March 2018.
- [2] Nefedov, L. I., Fil', N. Yu., Gubin, Yu. L., Mel'nichenko Ye. M. (2011). Modeli i metody upravleniya chrezvychnymi situatsiyami na magistral'nykh avtomobil'nykh dorogach. KhNADU, 136.
- [3] Drobnohod, N. I. (2001). Ustoychivoye ekologicheskii bezopasnoye razvitiye: ukrainskiy kontekst. *Zerkalo nedeli*, 21, 15. Accessed: 2 March 2018.
- [4] Shtovba, S. D. Vvedeniye v teoriyu nechetkikh mnozhestv i nechetkuyu logiku. [Introduction to the theory of fuzzy sets and fuzzy logic.] Available at: <http://matlab.exponenta.ru/fuzzylogic/book1>. Accessed: 2 March 2018.
- [5] Paklin, N. B. Nchetkiye zaprosy k relyatsionnym bazam dannykh. [Fuzzy demands to relational databases.] Available at: <http://www.basegroup.ru>. Accessed: 2 March 2018.
- [6] Chuykova, E. N., Molchanov, A. A. Formirovaniye nechetkikh zaprosov k baze dannykh. [Assignment of fuzzy demands to the database.] Sbornik konkursnykh rabot Vserossiyskogo smotrakonkursa nauchno-tekhnicheskogo tvorchestva studentov vysshikh uchebnykh zavedeniy «Eureka-2007». [Collection of entries of the All-Russian competition of sci and tech creativity of university students “Eureka-2007”.] Novocherkassk : Oniks+, 2007, pp. 38–40.
- [7] Moiseyenko, S. I., Levchenko, V. A. Translyatsiya operatorov relyatsionnoy algebry v SQLoperatoru. [Translation of relation algebra statements to SQL – statements.] *Vestnik of DSTU*, 2005, vol. 1, no. 1, pp. 691–696.
- [8] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning - I. *Information Sciences*, 1975, vol. 8, iss. 3, pp. 199–249.
- [9] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning - II. *Information Sciences*, 1975, vol. 8, iss. 4, pp. 301–357.
- [10] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning - III. *Information Sciences*, 1975, vol. 9, iss. 1, pp. 43–80.

Особливості Моделювання Процесів Оцінювання Поточного і Прогнозованого Технічного Стану Автотранспорту при Діагностуванні

О. В. Лапіна
Одеська державна академія
технічного регулювання та якості,
Одеса, Україна
elapina@ukr.net

Д. М. Андрица
Одеська державна академія
технічного регулювання та якості,
Одеса, Україна
andritsa@gmail.com

К. Р. Лапін
Одеська державна академія
технічного регулювання та якості,
Одеса, Україна
lapinkaren@gmail.com,

Я. Ю. Алексенко
Одеська державна академія
технічного регулювання та якості,
Одеса, Україна
yaroslav_barcafans@gmail.com

Features of Modeling Processes Estimation of Current and Forecast Technical Condition of Motor Transport at Diagnosis

Elena Lapina
Odessa State Academy of Technical Regulation and Quality,
Odessa, Ukraine
elapina@ukr.net

Dmitry Andritsa
Odessa State Academy of Technical Regulation and Quality,
Odessa, Ukraine Одеса, Україна
andritsa@gmail.com

Karen Lapin
Odessa State Academy of Technical Regulation and Quality,
Odessa, Ukraine
lapinkaren@gmail.com,

Yaroslav Aleksenko
Odessa State Academy of Technical Regulation and Quality,
Odessa, Ukraine
yaroslav_barcafans@gmail.com

Анотація — Розглянуті аспекти моделювання при оцінюванні поточного і прогнозованого технічного стану автомобіля при діагностуванні. Проаналізовано проведення контролю параметрів стану транспортних засобів в системі інформаційного забезпечення технічного діагностування.

Ключові слова — моделювання, обслуговування, модель, діагностування, транспортний засіб, технічний стан, працездатність, надійність, параметр, діагностична модель.

Abstract - The aspects of modeling are considered in the estimation of the current and predicted technical condition of the car during diagnosis. The control of parameters of the state of vehicles in the system of information provision of technical diagnostics has been analyzed.

Keywords - modeling, maintenance, model, diagnosis, vehicle, technical condition, performance, reliability, parameter, diagnostic model.

I. ВСТУП

Для ефективного оцінювання технічного стану автотранспорту необхідно регулярно проводити конструктивну модернізацію засобів технічного діагностування. Застосування нових технологічних та інформаційних методів вибору стратегії обслуговування

та ремонту автотранспорту сприяє удосконаленню технічного діагностування. Точність результатів діагностування залежить від вірності вибору перевіряємих параметрів, які зумовлюють методичну похибку, та ступеня досконалості засобів діагностування, які зумовлюють інструментальну похибку. Великі похибки вимірювань призводять до помилок при оцінюванні поточного і прогнозованого технічного стану автотранспорту.

II. ОСНОВНА ЧАСТИНА

Стрімкі темпи розвитку і впровадження в усі сфери діяльності сучасних технічних систем, безперервне зростання їх структурної складності і розмірності, спеціалізовані умови застосування і вимоги до безвідмовності виконуваних функцій визначають актуальність проблеми надійності, якості і безпеки експлуатації технічних засобів. Велике значення в успішному вирішенні цих завдань належить методам і засобам контролю і технічної діагностики.

При експлуатації слід враховувати вимоги технічного контролю - своєчасно визначати дійсний стан об'єкта (справний, допустимий, передаварійний, аварійний), і, в разі несправності, ефективно виявляти і усувати виниклі дефекти.

Інженерна практика все частіше стикається з проблемою вирішення завдань діагностики транспортних засобів з великим числом можливих пошкоджень, що вимагають швидкої локалізації для запобігання серйозних аварійних наслідків. Вирішення цих завдань вимагає розробки математичних моделей, методів і алгоритмів ефективного виявлення дефектів, які можна застосувати не тільки в конкретній прикладній області, але що володіють властивістю спільності для великого класу технічних систем.

Проблема оптимізації стратегії діагностування з метою оперативної локалізації та усунення пошкоджень істотно загострюється, з огляду на вимоги надійності. Вирішення цього завдання приймає першочергового значення для засобів, якість функціонування яких істотно впливає на екологічну обстановку, і несвочасне виявлення дефектів може призвести до небажаних наслідків. Оптимізація управлінських рішень, забезпечення оперативного контролю та діагностування технічного стану транспортних засобів є актуальною, що вимагає особливої уваги проблемою.

Особливістю будь-якої технічної системи як об'єкта діагностування є тісний взаємозв'язок електричних, електромеханічних і механічних пристроїв і елементів, що відрізняються функціональним призначенням і принципом дії. При описі їх технічного стану логічно застосовувати ті математичні формули і апарати, які б найкращим чином відповідали пошуку дефекту в цьому пристрої (диференціальні, різницеві, логічні рівняння, функціональні, структурні схеми, орієнтовані графи, кінцеві автомати і т.п.)

При моделюванні відбувається спрощений тією чи іншою мірою розгляд явища, яке досліджується. При цьому потрібно визначити основні закономірності об'єкту дослідження.

Моделюванням вважають будь-яку формалізовану сутність (наприклад, безліч параметрів і їх взаємозв'язків), що характеризує деякі певні властивості реального об'єкта (процесу), представлені в прийнятній або наочній формі.

З метою скорочення обсягу робіт в процесі діагностування об'єкта, встановлення взаємозв'язку між параметрами, розкриття фізичної сутності процесів, які відбуваються, ідентифікації вимірюваних величин і параметрів технічного стану, застосовують методи математичного моделювання.

Експлуатація транспортних засобів вимагає постійного моніторингу технічних параметрів фактичного стану, проведення необхідних технічних дій з обслуговування для забезпечення належного працездатного стану. Вищезазначене можливе лише за умов регулярного моніторингу технічного стану, який потребує постійної діагностики та прогнозування параметрів.

Справний або несправний об'єкт може бути представлений як динамічна система, стан якої в будь-який момент часу визначається вхідних, внутрішніх і вихідних параметрів.

Побудові діагностичної моделі повинні передувати різні дослідження, в результаті яких необхідно з'ясувати структуру об'єкта, функції, які вионуються блоками та

об'єктом в цілому, режим роботи, склад елементів і зв'язок між ними, наявність зворотних зв'язків і можливості їх розриву на час діагностування, ознаки та параметри нормального функціонування, робочі сигнали, діапазон виміру параметрів при нормальному функціонуванні, характерні відмови елементів та їх комбінації, наявність вузлів регулювання.

Математичну модель об'єкта діагностування можна подати в аналітичній, графічній, векторній або табличній формах. Позначимо символом X n -мірний вектор, компонентами якого є значення n вхідних змінних x_1, x_2, \dots, x_n (рис.1).

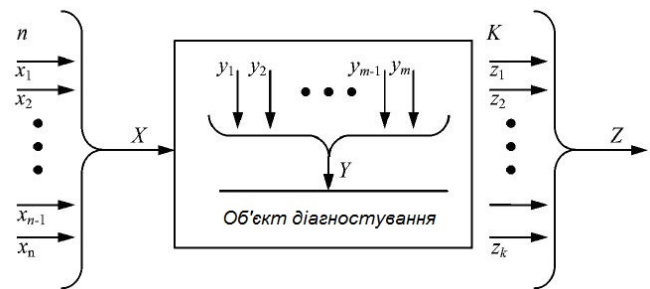


Рис. 1 – Математична модель об'єкта діагностування

Аналогічно Y є m -мірним вектором значень m внутрішніх змінних y_1, y_2, \dots, y_m , а Z - k -мірним вектором значень k вихідних функцій z_1, z_2, \dots, z_k . вихідна функція $Z = \varphi(X, Y_{поч}, t)$ є математичною моделлю справного об'єкта. При цьому $Y_{поч}$ відображає початкові значення внутрішніх змінних параметрів об'єкта, а t - фактор часу. В процесі роботи об'єкта відбувається зміна внутрішніх змінних y_1, y_2, \dots, y_m і можлива поява поодиноких або кратних відмов. Під одиночною несправністю розуміють елементарну відмову, яка не є сукупністю невеликих незначних несправностей. Кратна несправність є сукупністю одночасної появи двох та більше поодиноких несправностей.

Математичну модель об'єкта діагностування можна подати в табличній формі з використанням результатів елементарних перевірок. Відмова хоча б одного з елементів системи призвести її до неробочого стану, що можна визначити як за допомогою таблиці функцій несправностей, так й аналітичним способом. Но задля побудування оптимальних алгоритмів діагностування об'єкту необхідно враховувати параметри надійності об'єкта, такі як: ймовірність відмови, частота відмов, виходячи з відповідного закону розподілу, та інші параметри.

Визначимо ці величини для довільних законів надійності елементів об'єкта, для чого розглянемо невеликий відрізок часу, обмежений точками t та $t + \Delta t$. Введемо випадкову величину T - час виникнення відмови в об'єкті. Ймовірність попадання відмови в проміжок часу $[t, t + \Delta t]$ позначимо $P(t \leq T \leq t + \Delta t)$.

Для будь-якої випадкової величини з щільністю розподілу, яка дорівнює $F'(t)$, де $F'(t)$ - закон її розподілу, ймовірність попадання в невеликий відрізок Δt дорівнює

$$f(t)\Delta t = F'(t)\Delta t \quad (1)$$

Цей вираз дозволяє стверджувати, що оптимальний алгоритм перевірок системи, побудований із врахуванням ймовірностей, буде оптимальним під час всього періода експлуатації системи.

Швидкість розвитку несправностей і робочих процесів визначає необхідність врахування динамічних ефектів при моделюванні.

Використовуючи математичну модель об'єкта контролю, можна створити найбільш раціональні методи діагностування конкретної системи в цілому або окремих її функціональних вузлів та блоків.

В процесі побудови математичної моделі розглядаються суттєві властивості об'єкта, необхідні для постановки діагнозу. Заміна реальних об'єктів їх моделями, які ідеалізовані, дозволяють відволіктися від їх фізичної природи і за допомогою різних математичних методів розв'язати діагностичні задачі формальним чином.

Для вибору типу моделі потрібно проаналізувати ряд чинників: умови експлуатації машини, конструктивне виконання, частотний діапазон вібрацій, що генеруються нею, і тому подібне [8]. Для створення розрахункової моделі (схеми) необхідно врахувати найбільш істотні властивості об'єкту, що визначають його динамічну поведінку.

За допомогою розрахункової моделі можлива теоретична оцінка вібропараметрів об'єкту (коливальної потужності, віброшвидкості, вібропереміщення і так далі). Для цього необхідно визначити вимушені значення сил, параметри вимушених коливань, власні частоти вільних коливань віброактивних конструктивних вузлів і деталей.

Аналіз діагностичних моделей об'

Результат діагностичного експерименту завжди випадковий, так як якщо він передумовлений, то проводити його нема сенсу. Таким чином, будь-який процес діагностування складається з послідовності елементарних перевірок із даним набором контрольних точок та при однакових умовах. В межах структурного підходу поняття «елементарна перевірка» застосовують також до окремих частин об'єкта діагностування або їх сукупності. В цьому випадку передбачається можливість входів та виходів цих частин. Потужність множини

можливих елементарних перевірок характеризується граничною множиною параметрів.

При дослідженні реальних об'єктів діагностування не всі перевірки можуть бути технічно реалізовані, тому формальна постановка питання повинна бути доповнена конкретним дослідженням об'єкта діагностування.

В залежності від характеру послідовності елементарних перевірок розрізняють два основних способи пошуку дефектів: комбінаційний та послідовний. При застосуванні комбінаційного способу стан об'єкта діагностування визначається шляхом виконання заданої кількості елементарних перевірок, послідовність виконання яких не має значення. Виявлення дефектних блоків здійснюється після проведення всіх заданих елементарних перевірок. З цією метою виконується аналіз результатів проведених перевірок. Для даного способу характерні такі ситуації, коли не всі результати проведених перевірок не є необхідними для визначення стану об'єкта діагностування. При використанні послідовного способу перевірки, здійснення яких достатньо для визначення всіх наперед заданих розрізняємих станів об'єкта, виконуються в деякому порядку. Результат кожного елементарної перевірки аналізується безпосередньо після його отримання.

Система моніторингу технічного стану автотранспорту в реальних умовах експлуатації складається з великої кількості елементів, які характеризуються відповідними вихідними умовами, знаходяться на різних ієрархічних рівнях і мають свої особливості функціонування. Це є складна динамічна і інформаційна система із чітко впорядкованою ієрархічною структурою, розгалуженою мережею взаємозв'язків між її елементами, що розвивається в просторі і часі. Діяльність складної виробничої системи визначається технологічними процесами, які реалізують її суб'єкти відповідно до цільової спрямованості в рамках предметної області [4]. Кожен суб'єкт моніторингу має властиві процедури – алгоритми функціонування, розрахунку, результати яких використовуються в алгоритмах суб'єктів процесу моніторингу різних ієрархічних рівнів. Після завершення процедур управління та удосконалення відповідних технологічних процесів відбувається зміна стану предметної області, тобто зміна (підвищення) рівня оцінки працездатності об'єкта дослідження.

У цілому модель функціонування суб'єктів моніторингу ефективності функціонування може бути з виподана як сукупність таких компонентів: моделі об'єктів, що беруть участь у структурних зв'язках; моделі параметрів, що визначають результати функціонування суб'єктів; моделюючих алгоритмів, що встановлюють правила функціонування об'єктів і зміни значень їх параметрів тощо. Така модель є динамічною та інформаційною і відображає не лише поточний стан транспортних засобів і умов експлуатації, але й зміну їх стану в часі. Таким чином, у загальному розумінні процес моніторингу технічного стану автотранспорту у реальних умовах експлуатації є процесом трансформації інформації параметрів технічного стану і процесів функціонування транспортних засобів (ТЗ) та умов їх експлуатації. Формально це відображення має вигляд:

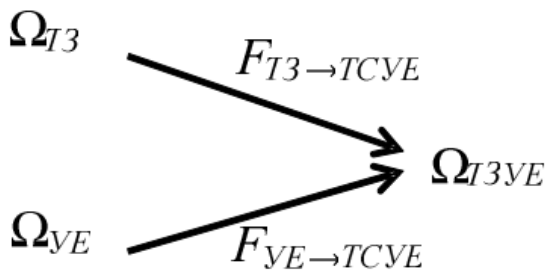


Рис.2 – Відображення процесу трансформації інформації параметрів технічного стану і процесів функціонування автотранспорту та умов їх експлуатації

де Ω_{T3} – множина моделей параметрів технічного стану і процесів функціонування транспортних засобів; Ω_{UE} – множина моделей параметрів умов експлуатації; $\Omega_{T3 UE}$ – множина моделей параметрів технічного стану ТЗ у відповідних умовах експлуатації; $F_{T3 \rightarrow TCUE}$ – функціональне відображення моделей параметрів технічного стану і процесів функціонування ТЗ; $F_{UE \rightarrow TCUE}$ – функціональне відображення моделей параметрів умов експлуатації ТЗ.

Основне призначення математичної моделі – це встановлення зв'язку між значеннями діагностичних параметрів та значеннями параметрів технічного стану (або безпосередньо із технічним станом). Таким чином, процес оцінювання та визначення можна відобразити наступним чином (рис.3).

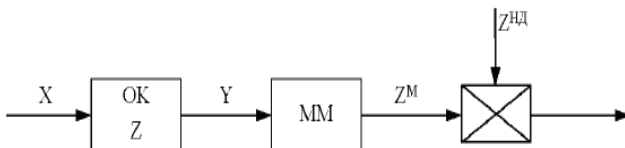


Рис.3 – Схематичне уявлення процесу визначення технічного стану

На об'єкт контролю (OK) з параметрами технічного стану Z діє зовнішнє збурення X. За допомогою засобів вимірювання визначають діагностичні параметри Y. Значення цих параметрів з використанням математичної моделі (MM) перетворюються в значення параметрів технічного стану Z^M . Далі ці значення порівнюються зі значеннями параметрів технічного стану (TC), які задані нормативно-технічною документацією Z^{HD} , та за результатами цього порівняння робиться висновок про технічний стан об'єкта контролю та технічного діагностування.

ВИСНОВКИ

При діагностуванні складних динамічних об'єктів, якими є автомобільні засоби, використання математичного моделювання ускладнене завдяки складності визначення аналітичних залежностей, які пов'язують зовнішні ознаки та відповідні їм несправності в деталях та вузлах об'єкта. Наведені та

розглянуті особливості моделювання процесів оцінювання поточного і прогнозованого технічного стану автотранспорту при діагностуванні.

REFERENCES ЛІТЕРАТУРА

- [1] Sivasubrahmaniyan, A., Jaishwal, A., Warriar, G., Tharaneetharan, S. et al., [Электронный ресурс] "A Systematic Approach of Improving Reliability Process through Development and Application of OnBoard Diagnostics System, for Commercial Vehicle," SAE Technical Paper 2015-26-0101, 2015, Published 01/14/2015, doi:10.4271/2015-26-0101. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [2] Denton T. Advanced Automotive Fault Diagnosis. 3rd edition. — New York: Routledge, 2012. XXII, 428 p. — ISBN: 978-0-08-096955-8, ISBN: 978-0-08-096956-5 (ebk).
- [3] Павлов, К. С. Модели выбора и замены оборудования в производственных системах машиностроительных предприятий [Электронный ресурс] / К. С. Павлов, Е. Н. Хоботов // Автоматика и телемеханика. – 2015. – № 2. – С. 125-140.
- [4] Гриньків, А. В. Використання методів прогнозування в керуванні технічним станом агрегатів та систем транспортних засобів /А.В.Гриньків // Техніка в сільськогосподарському виробництві, галузеве машинобудування, автоматизація. – 2016. – Вип. 29. – С. 25-32.
- [5] Сухарев, Э. А. Теория эксплуатационной надежности машин [Текст]: монография / Э. А. Сухарев. – Ровно: Изд-во УГАВХ, 1997. – 162 с.
- [6] Аулін, В. В. Забезпечення та підвищення експлуатаційної надійності транспортних засобів на основі використання методів теорії чутливості [Текст] / В. В. Аулін, А.В. Ериньків, Т.М. Замота // Вісник Інженерної академії України. – 2015. – №.3. – С. 66-73.
- [7] Kärpler W.D. Smart Vehicle Handling - Test und Evaluation in der Fahrzeugtechnik. pringer Vieweg, Kraftfahrzeugtechnik, 2015. — 249 p. — ISBN 978-3-662-46416-8, ISBN 978-3-662-46417-5 (eBook)
- [8] Карташевич А.Н. Диагностирование автомобилей. - Практикум : учеб. пособие / под ред. А.Н. Карташевича. — Минск : Новое знание ; М. : ИНФРА-М, 2011. — 208 с. : ил. — (Высшее образование). ISBN: 978-5-16-004864-2
- [9] Контроль роботи транспортного двигуна з використанням інформаційних технологій [Текст] / В. П. Матейчик, В. П. Волков, П. Б. Комов [та ін.] // Двигатели внутреннего сгорания: науч.-техн. журнал. – Харьков: НТУ «ХПИ», 2013. – №2. – С. 27-31.

Computer Modeling of Text Information Classification

Vadym Yakovenko

Department of Information System and
Technologies
University of Customs and Finance
Dnipro, Ukraine
yakovenko@ua.fm

Yuliya Uliyanovska

Department of Information System and
Technologies
University of Customs and Finance
Dnipro, Ukraine
yuliyauyv@gmail.com

Oleksii Kaliaka

Department of Information System and
Technologies
University of Customs and Finance
Dnipro, Ukraine
alexey.kalyaka97@gmail.com

Abstract - The paper considers the solution of the task of creating a neural network for the text document recognition. A research was conducted of existing models of neural networks, methods of vector to text representation. As a result of the experiment were identified the advantages and disadvantages of selected algorithms. Although an original algorithm for classifying texts using a neural network, using a perceptron with backpropagation method of training was developed.

Keywords: neural network, perceptron, classification

I. INTRODUCTION

The usage and improvement of computer technology opens up a range of new features for human, but with this there are a number of other problems associated with the growth of processed arrays of text documents. Working with full-text databases brings a problem of searching and classification of documents according to their content. Therefore, the task was to develop an algorithm capable of classifying text information.

Classification of documents - this is one of the tasks of information search, which consists in enrolling a document in one of several categories based on its content.

Classification tasks:

- Creating thematic directories
- Spam filtering
- Electronic systems of document flow
- Tagging
- Automatic text translation [1].

To solve this problem, were considered the most popular classification methods such as: classification using the decision tree, Bayesian classification method, reference vector method, k-neighbor method and artificial neural networks.

II. MAIN PART

The method of classification in the program is the neural network. Compared to other methods of classification, the neural network has the greatest accuracy, but requires high computing power.

An artificial neural network is an adaptive system that consists of a group of connected artificial neurons. [2]

Neural networks can be used to solve many tasks of information processing. As is known [3], an artificial (mathematical) neuron performs the transformation of the input vector of signals, as follows (1):

$$y = I(S); S = \sum w_i x_i, \quad (1)$$

w_i - Weighing vector of a neuron (weights of synaptic bonds);

S - Result of weighted addition;

I – neuron activation function.

Functionality of a neuron is simple, so neurons unite in networks for solving specific problems. Neural network learning is reduced to the selection of the weight coefficients of each neuron, provided that the chosen topology of the network and the selected activation function I [4].

The main types of classifying neural networks are: single-layer and multi-layer perceptron, a neural network Gaussian classifier, a Kohonen network, a network of embedded distribution, a cascade network [5].

To solve the problem were analyzed the different types of existing neural networks, their ability to perform the task and features. Methods of neural network learning and ways of converting text into a signal were also analyzed. As a result, the perceptron was chosen as a model of the neural network.

The developed model of neural network consist of three layers of neurons such as input, hidden and output.

The first two layers are dynamic, that's means that they can change the number of neurons depending on the size of the dictionary. The hidden layer is setting up depending on the input in the proportion of 2: 3. This gives the neural network some flexibility and adaptability.

The output layer has a fixed number of 3 neurons, but can be changed depending on user needs. The neuronal network with 3 output neurons can take up to 8 different results, but to facilitate learning in the developed model, each source neuron corresponds to a separate classification category.

The activation function in the model is sigmoidal function. This is an S-shaped, continuously differentiated, monotonous increasing function that can take values from 0 to 1 (2). Sigmoid function has the ability to amplify weak signals

better than large ones and prevents saturation from large signals.

$$f(x) = \frac{1}{1+e^{-x}} \quad (2)$$

Creation of input signal vector requires a dictionary. It must be created manually, or the program (you need to download the document with the data) can automatically create it. For better productivity it is desirable to use documents of similar subjects to the analyzed texts. The dictionary is created using N-grams, such as unigrams. The uploaded document is split into separate words, and then algorithm removes repetitions, symbols, and characters. Then they are copied to a separate document that can be edited optionally.

During the program, the documents subject to classification will be compared with this dictionary and as a result, a text document vector will be created.

The back propagation method is used in project as a learning method. This method is the most popular of learning multilayer perceptron. It can teach all layers of the neural network, but needs a continuous activation function to perform a gradient descent to set the weights of the neurons.

Was created an application that is capable to perform any number of any number of experiments with current neural network to test the system properly. The program is very flexible in setting up so you can set the number of neurons in the NN layers, the number of training periods, the number of training data and the number of experiments. You don't need text documents to train the program. Vectors are automatically generated and randomly assigned to a category. You can specify their number in the Train set field. Result sets field determines the number of vectors that will be generated for testing the trained network.

The result of the program is the average error of the neural network.

With this program were conducted several experiments with different data and the accuracy of the system was determined depending on the number of neurons on the layers of the neural network. For the accuracy of the results, the experiment was conducted several times (Fig. 2). As a result, it was found that with increasing number of Input neurons, reduced accuracy of the network. With 20 Input Neurons and 12 Outputs, the error was 1% and the training lasted 10 seconds. With 40 Input Neurons and 26 Outputs, the error was already 6%, training lasted 5 minutes 20 seconds. At 100 and 66 neurons, the error was 20% and the duration of training exceeded 8 hours (Tab. 1). According to the obtained data, we can conclude that the neural network will best recognize the small documents.

TABLE 1. EXPERIMENT RESULTS.

#	Experiment results		
	Neuron Input/Hidden	Time	Accuracy
1	20/12	10s	99%
2	30/20	1m	96%
3	40/26	5m20s	94%
4	50/33	9m30s	92%
5	60/40	45m	88%
6	70/46	2h	86%
7	100/66	8h	80%

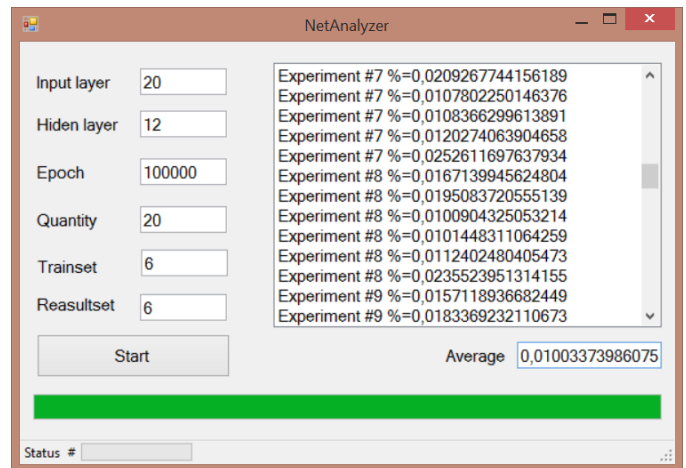


Fig 1. Program interface.

CONCLUSION

The result of the work is creation of a software product capable of classifying text documents by predefined categories. The system tests were performed, accuracy checked and all the advantages and disadvantages of the perceptron, the method of training the network and the vectorization of information were revealed. It have been proved experimentally that the system is better able to classify documents of a small volume.

REFERENCES ЛІТЕРАТУРА

- [1] A.S. Eprev. "Automatic classification of text documents". Mathematical structures and modeling. vol 21, 2010, pp. 65-81.
- [2] Volosyuk Yu.V. "Methods of text documents classification in the tasks of Text Mining" Scientific notes of the Ukrainian Research Institute of Communication. vol 6(34), 2014, pp. 76-81
- [3] V.V. Korneev, A.F. Gareev, S.V. Vasjutin, V.V. "Database. Intellectual information processing". Moscow, Knowkedge, 2003, 400p.
- [4] Ulianovska Yu. V. "Application of neural networks to determine the customs value of goods". Eastern-European Journal of Enterprise Technologies, vol 3/3 (57), 2012, P. 46-49
- [5] Reich. - M. : "Knowledge", 2000. – 352 c. Yang Y. A re-examination of text categorization methods / Y. Yang, X. Liu // Proc. of Int. ACM Conference on Research and Development in Information Retrieval (SIGIR-99), – 2007. – P. 42-49.

Застосування Згорткових Нейронних Мереж в Задачах Класифікації Акустичних Даних

Олексій Кудін
кафедра програмної інженерії
Запорізький національний університет
Запоріжжя, Україна
avk256@gmail.com

Анастасія Кривохата
кафедра програмної інженерії
Запорізький національний університет
Запоріжжя, Україна
krivohata@gmail.com

Application of Convolutional Neural Networks for the Acoustic Data Classification

Oleksii Kudin
software engineering dept.
Zaporizhzhya National University
Zaporizhzhya, Ukraine
avk256@gmail.com

Anastasia Kryvokhata
software engineering dept.
Zaporizhzhya National University
Zaporizhzhya, Ukraine
krivohata@gmail.com

Анотація— Останнім часом все більшого поширення набувають автоматичні системи розпізнавання звуку, зображень, відео. Серед методів, що застосовуються у таких задачах, слід виділити методи машинного навчання, як найбільш популярні та перспективні. У роботі пропонується підхід до застосування згорткових нейронних мереж в задачах класифікації звуку.

Ключові слова — акустичні дані, класифікація, машинний слух, машинне навчання

Abstract— Machine learning methods like K-means, SVM, KNN, decision trees etc. are traditionally used in audio classification. In the last two decades, the deep learning based methods have also gained popularity for audio tagging, for example, convolutional neural networks or recurrent neural networks. Deep neural networks can operate on both raw acoustical signal and features extracted from the signal. Further we build a proof-of-concept system for the classification of acoustic data on the basis of convolutional neural networks and construct a neural network ensemble. In addition, we outline a direction for further development of machine hearing systems based on our analysis and experimental model. The approach which follows this direction can use different types of ensemble learning methods with classifiers based on feature extraction and deep neural networks.

Keywords— acoustic data, classification, machine hearing, machine learning

I. ВСТУП

Аналіз мультимедійних даних, зокрема, акустичних записів є досить актуальною задачею. Це пов'язано, насамперед, з вибуховим ростом кількості звукової інформації в мережі Інтернет (музичні записи, записи лекцій, доповідей тощо). Для можливості пошуку серед таких даних, зазвичай, використовуються метадані, які описують в текстовому вигляді зміст відповідного аудіо файлу. Розробка автоматизованих систем анутовання таких акустичних даних, тобто формування метаданих

які б описували в текстовому виді зміст аудіофайлу є досить перспективним напрямом досліджень.

II. ОГЛЯД ПУБЛІКАЦІЯ

Процес аналізу акустичних даних можна поділити на декілька етапів. Спочатку, виконується попередня обробка неперервного акустичного сигналу з метою представлення його у дискретному цифровому вигляді. При цьому зазвичай використовується ряд стандартних підходів [1, 2]. Далі, виділяються ознаки акустичного сигналу, суттєві для розв'язання поставленої задачі аналізу. Серед найбільш широко вживаних ознак використовуються коефіцієнти перетворення Фур'є та автокореляції, мел-кепстральні коефіцієнти, хромограми. Після цього, отримані ознаки використовуються як вхідні параметри математичної моделі (наприклад, алгоритму класифікації, кластеризації або нейронної мережі). На заключному етапі, виконується верифікація отриманих результатів та впровадження розробленої системи акустичного аналізу.

На даний момент опубліковано досить багато спеціальних та оглядових робіт, присвячених тематиці розробки систем класифікації звуку. Так, в оглядових статтях [3, 4, 5] наводиться опис компонент системи автоматичної класифікації звуків, яка містить модулі попередньої обробки, екстракції ознак, алгоритм навчання та модуль обчислень.

В [3] детально розглянуто підходи до виділення ознак сигналу. Наводяться критерії, за якими можна класифікувати мову, музику та природні звуки. Виділяються методи, засновані на фізичних властивостях сигналів та особливостях людського сприйняття звуків. Частіше за все використовуються методи виділення ознак, які представляють акустичний сигнал у таких областях: часовій, частотній, кепстральній та вейвлет.

Огляди [4, 5] містять аналіз загальних підходів та публікацій з автоматичної класифікації музичних записів

за жанрами. Пропонується множина найбільш інформативних міток, які можуть використовуватись як класи при навчанні класифікаторів. Розглядаються найбільш вживані джерела розмічених акустичних даних, які можуть використовуватися в системах навчання з вчителем. Зазвичай, це відкриті музичні бази в мережі Інтернет, розмічені користувачами записи, наприклад, у соціальних мережах та дані, які згенеровано спеціально для розв'язання задач машинного слуху. В роботі [5] окремо розглядається питання оцінки ефективності систем класифікації музичних файлів за жанрами.

Для класифікації даних за певними ознаками, можуть використовуватися як статистичні методи (класифікатор Баєса [6, 7], дискримінантний аналіз [8], ЕМ алгоритм тощо), так і методи, які ґрунтуються на мірах схожості та відмінності (метод k-середніх, метод опорних векторів, метод k найближчих сусідів [9] тощо).

В останні роки все більше робіт присвячено використанню нейронних мереж як при вилученні ознак з даних, так і безпосередньо при класифікації [10, 11, 12].

Метою даної роботи є побудова системи класифікації акустичних даних різного походження на основі згорткових нейронних мереж та застосування підходу [13] для побудови ансамблю нейронних мереж. Перевагою такого підходу є його висока ефективність та адаптивність з точки зору вимогливості до обчислювальних ресурсів, оскільки, за необхідністю, можна коригувати кількість класифікаторів, які входять до ансамблю.

III. ОПИС ТЕСТОВОГО НАБОРУ АКУСТИЧНИХ ДАНИХ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ КЛАСИФІКАТОРА

Для тестування запропонованих підходів, в роботі використовуються дані з ресурсу www.kaggle.com, а саме – набір даних Urban Sound Classification. Набір даних містить 3449 звукових файли у форматі .wav для навчання та тестування системи. Вибірка для навчання містить звукові файли з 9 категорій.

При програмній реалізації використовувалися такі бібліотеки мови програмування Python: NumPy, Pandas, Librosa, Keras.

Повна реалізація наведена за посиланням <https://www.kaggle.com/avk256/fork-of-urban-sound-class-using-cnn-snapshot-ensem?scriptVersionId=5625604>.

ЛІТЕРАТУРА

- [1] Camastra F., Vinciarelli A. Machine learning for Audio, Image and Video analysis. London: Springer-Verlag. 2015. 561 p.
- [2] Mierswa I., Morik K. Learning feature extraction for learning from audio data. Technische Universität Dortmund. Technical Reports. 2004. No. 55.
- [3] Alias F., Socoró J.C., Sevillano X. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. *Applied Sciences*. 2016. № 6(5):143.
- [4] Bertin-Mahieux T., Eck D., Mandel M. Automatic tagging of audio: the state-of-the-art. Machine audition: principles, algorithms and systems. IGI Global. 2011. P. 334–352.
- [5] Salamon J., Jacoby C., Bello J.P. A dataset and taxonomy for urban sound research. <http://dx.doi.org/10.1145/2647868.2655045>, 2017. P. 1–4.

- [6] Dongge Li, Ishwar K. Sethi, Nevenka Dimitrova, Tom McGee. Classification of general audio data for content-based retrieval. *Pattern Recognition Letters*. 2001. № 22(5). P. 533–544.
- [7] Mierswa I., Morik K. Learning feature extraction for learning from audio data. Technische Universität Dortmund. Technical Reports. 2004. No. 55.
- [8] Burges C.J.S., Platt J.C., Jana S. Extracting noise-robust features from audio data. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, FL, USA, 13–17 May 2002. 2002. P. 1021–1024.
- [9] Camastra F., Vinciarelli A. Machine learning for Audio, Image and Video analysis. London: Springer-Verlag. 2015. 561 p.
- [10] Bach J.-H., Meyer A.-F., McElfresh D., Anemüller J. Automatic classification of audio data using nonlinear neural response models. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan. 2012. P. 357–360.
- [11] Burges C.J.S., Platt J.C., Jana S. Extracting noise-robust features from audio data. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, FL, USA, 13–17 May 2002. 2002. P. 1021–1024.
- [12] Kong Q., Xu Y., Wang W., Plumbley M.D. Convolutional gated recurrent neural network incorporating spatial features for audio tagging. The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, Alaska. 2017.

REFERENCES

- [1] Camastra F., Vinciarelli A. Machine learning for Audio, Image and Video analysis. London: Springer-Verlag. 2015. 561 p.
- [2] Mierswa I., Morik K. Learning feature extraction for learning from audio data. Technische Universität Dortmund. Technical Reports. 2004. No. 55.
- [3] Alias F., Socoró J.C., Sevillano X. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. *Applied Sciences*. 2016. № 6(5):143.
- [4] Bertin-Mahieux T., Eck D., Mandel M. Automatic tagging of audio: the state-of-the-art. Machine audition: principles, algorithms and systems. IGI Global. 2011. P. 334–352.
- [5] Salamon J., Jacoby C., Bello J.P. A dataset and taxonomy for urban sound research. <http://dx.doi.org/10.1145/2647868.2655045>, 2017. P. 1–4.
- [6] Dongge Li, Ishwar K. Sethi, Nevenka Dimitrova, Tom McGee. Classification of general audio data for content-based retrieval. *Pattern Recognition Letters*. 2001. № 22(5). P. 533–544.
- [7] Mierswa I., Morik K. Learning feature extraction for learning from audio data. Technische Universität Dortmund. Technical Reports. 2004. No. 55.
- [8] Burges C.J.S., Platt J.C., Jana S. Extracting noise-robust features from audio data. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, FL, USA, 13–17 May 2002. 2002. P. 1021–1024.
- [9] Camastra F., Vinciarelli A. Machine learning for Audio, Image and Video analysis. London: Springer-Verlag. 2015. 561 p.
- [10] Bach J.-H., Meyer A.-F., McElfresh D., Anemüller J. Automatic classification of audio data using nonlinear neural response models. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan. 2012. P. 357–360.
- [11] Burges C.J.S., Platt J.C., Jana S. Extracting noise-robust features from audio data. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Orlando, FL, USA, 13–17 May 2002. 2002. P. 1021–1024.
- [12] Kong Q., Xu Y., Wang W., Plumbley M.D. Convolutional gated recurrent neural network incorporating spatial features for audio tagging. The 2017 International Joint Conference on Neural Networks (IJCNN 2017), Anchorage, Alaska. 2017.
- [13] Howel J., Rooth M., Wagner M. Acoustic classification of focus: on the web and in the lab. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*. № 8(1):16. 2017. P. 1–41.

Застосування Алгоритму CART для Автоматичної Класифікації Кардіограм

Олег Федоришин

кафедра комп'ютерних наук та інформаційних систем
Прикарпатський національний університет

Івано-Франківськ, Україна

vout0rpvs@gmail.com

Application of CART Algorithm for the Automatic Classification of Cardiograms

Oleg Fedoryshyn

dept. of Computer Science and Information Systems
Precarpathian National University

Ivano-Frankivsk, Ukraine

vout0rpvs@gmail.com

Анотація—Розроблено метод автоматичного розпізнавання кардіограм. Оцінено ефективність застосування розробленого методу у якості компоненту автоматичних електрокардіографів.

Abstract—Method for performing automatic cardiograms classification was synthesized. Effectiveness of the described application was estimated as a part of automatic electrocardiograph.

Ключові слова—цифрова обробка сигналів; електрокардіограма; CART;

Keywords—digital signal processing; electrocardiogram; CART;

ВСТУП

Платформи машинного навчання є одними з найбільш конкурентоспроможних сфер діяльності корпоративних технологій, серед яких найбільші корпорації в галузях науки і медицини готуються перевести клієнтів на послуги автоматичних платформ, що охоплюють весь спектр функцій машинного навчання, включаючи збір і підготовку даних, діагностування захворювань, побудова моделі, її навчання та застосування в практичній медицині.

Тому така тематика досліджень є актуальною, оскільки в сучасних світових умовах домінуючою основою всієї медицини поступово стає машинне навчання та автоматизація роботи з пацієнтами, налаштування процесів для швидкого і якісного надання лікувальних послуг, для чого необхідне отримання точного діагнозу в найкоротші терміни.

II. ОСНОВНА ЧАСТИНА

Електрокардіограма (ЕКГ) являє собою графічне зображення пов'язаного з серцевим м'язом електричного сигналу, вимірюваного протягом деякого часу.

Застосування ЕКГ - від моніторингу серцевого ритму до діагностики конкретних серцевих захворювань. Основи вимірювання ЕКГ однакові для всіх випадків, але деталі та вимоги до електричних компонентів значно відрізняються.

Крім біологічних сигналів, більшість ЕКГ також виявляють два техногенних сигнали. Найважливіший з цих сигналів походить від імплантованих кардіостимуляторів і називається "темпом". Швидкий сигнал відносно короткий, десятки мікросекунд, а амплітуда коливається від декількох мілівольт до майже 1 вольт. Часто ЕКГ повинна виявити наявність сигналу швидкості, одночасно запобігаючи при цьому спотворенню інших сигналів від серця. Схема вимірювання ЕКГ та типове вхідне коло, наведені на рисунку 1.[1, 2]

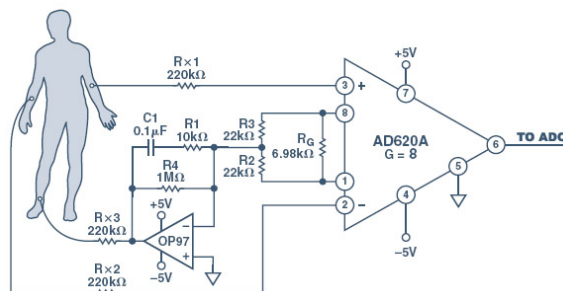


Рис. 1. Схема вимірювання ЕКГ та типове вхідне коло

При реалізації алгоритмів, які використовуються при проектуванні кардіографічних систем застосовуються спеціалізовані АЦП [5], цифрові хвильові фільтри [4,6], методи стиснення сигналів [3], та спеціалізовані системи кодування інформації [7]. Типовий вигляд фрагменту ЕКГ наведено на рисунку 2.

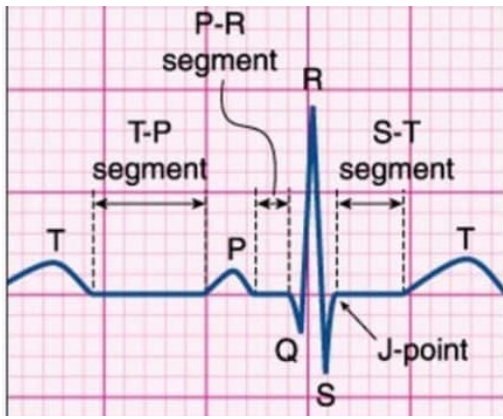


Рис. 2. Фрагмент ЕКГ

Для реалізації алгоритму розпізнавання вибрано алгоритм CART. Класифікаційні та регресійні дерева або CART для короткострокових вимірюваних даних - це термін, представлений Лео Брейманом для позначення алгоритмів дерева рішень, які можуть бути використані для задач прогнозування, моделювання, класифікації або регресії. Класично ці алгоритми отримали назву "дерева рішень", але на деяких платформах, таких як R, вони називаються більш сучасним терміном CART. Алгоритм CART є основою для важливих алгоритмів, таких як random forest та decision trees.

Представлення моделі CART - це двійкове дерево, що складається з алгоритмів і структур даних. Кожен кореневий вузол представляє собою єдину вхідну змінну (x) і точку розбиття для цієї змінної (припускаючи, що змінна числова). Листяні вузли дерева містять вихідну змінну (y), яка використовується для прогнозування. Даний набір даних з входами (x), взятий безпосередньо з реалізації алгоритму і є прикладом бінарного дерева рішень, зображеного на рисунку 3.

```
Is QRS >= 9.58083832335328?
--> True:
  Predict {'Ventricular Tachycardia': 16}
--> False:
  Is S >= 0.04197925382752761?
  --> True:
    Predict {'Sinus Bradycardia': 3}
  --> False:
    Is S >= -0.18836751558265838?
    --> True:
      Predict {'Myocardial Infarction': 7}
    --> False:
      Predict {'Sinus Bradycardia': 3}
```

Рис. 3. Бінарне дерево рішень

З представленням бінарного дерева для моделі CART, описаної вище, прогнозування є відносно простим. Враховуючи нові вхідні дані, дерево проходить через оцінку конкретного нового значення, розпочатою в кореневому вузлі дерева. Вивчення бінарного дерева насправді є розподілом вхідного простору. Можна придумати кожну вхідну змінну як розмірність у r -вимірному просторі. Деревом рішень проводиться

розбивка розмірності на прямокутники ($p = 2$ вхідних змінних) або на гіпер-прямокутник з більшою кількістю вхідних даних

Нові дані фільтруються через дерево і вводяться в одну з прямокутників, а вихідне значення для цього прямокутника є прогнозом, зробленим моделлю. Створення моделі CART передбачає вибір вхідних змінних та розподіл точок на ці змінні, доки не буде побудовано відповідне дерево. Вибір вхідної змінної, що використовується, та конкретної роздільної або роздільної точки вибирається за допомогою «жадібного» алгоритму, щоб мінімізувати функцію вартості.

Для регресійних задач прогнозування моделювання функція вартості, яка зводиться до вибору розподілених точок, є квадратичною помилкою суми у всіх навчальних зразках, які входять у прямокутник:

$$\sum (Y_{\text{actual}} - Y_p)^2 \quad (1)$$

де Y_{actual} - вихідний сигнал для зразка тренувань, а Y_p - прогнозований результат з прямокутника.

Для класифікації використовується функція індексу Джині або невизначеність Джині, яка дає вказівку на те, як "чисті" вузли листя (як змішані дані навчання, призначені для кожного вузла).

$$Gini = 1 - \sum_{i=1}^p (p_i)^2 \quad (2)$$

де Gini - індекс Джині над усіма класами, p - частка навчальних екземплярів класу i у прямокутнику. У вузлі, що має всі класи одного типу (ідеальна чистота класу), буде мати $G = 0$, в той час як G , який має 50-50 розбиття класів для задачі бінарної класифікації (найгіршої чистоти), буде мати $G = 0.5$.

Розрахунок індексу Джині для кожного вузла зважується на загальну кількість випадків у головному вузлі. Отже, коефіцієнт Джині для обраної роздільної точки в задачі двійкової класифікації розраховується таким чином:

$$G = \frac{\sum_{i=1}^n \sum_{j=1}^n |x_i - x_j|}{2n \sum_{i=1}^n x_i} \quad (3)$$

Для реалізації поставленої задачі було вирішено використати мову програмування Python версії 3.6.4, з використанням менеджера керування пакетами Anaconda. Блок схема реалізованого алгоритму наведена на рисунку 4.



Рис. 4.Робочий алгоритм для класифікації кардіограм

ВИСНОВКИ

Розв'язання проблеми автоматизації діагностування захворювань серцево-судинної системи по сигналу кардіограми на основі розробки використання парадигм машинного навчання дозволило виявити, що домінуючу роль у вирішенні поставленої задачі було визначення

оптимального алгоритму що визначає програмну реалізацію. Було обрано алгоритм CART, перевагами якого є простота у використанні та ефективність класифікації даних. Недоліком цього підходу виявилась складна підготовка даних для обробки алгоритмом, та ідентифікація необхідних сегментів у відповідних гілках.

Для покращення ідентифікації необхідних сегментів кожного серцевого циклу, було застосовано методи нормалізації з використанням b-сплайн функції і комплексного методу обробки сигналу для визначення необхідних даних. Окрім цього було застосовано сегментатор Гамільтона, для більш точного відмежування кожного серцевого циклу, що дало змогу краще виділити основні сегменти і зубці.

В результаті було реалізовано оптимальну підготовку даних та підібрано коректну множину значень для отримання достовірних діагнозів. Значення на виході виявились достовірними до актуальних діагнозів, згідно з якими було проведено тестування роботи програми, що доводить те, що реалізація алгоритму є ефективним засобом діагностування захворювань серцево-судинної системи

REFERENCES ЛІТЕРАТУРА

- [1] Франк А. Флакскampf: Практическая эхокардиография: руководство по эхокардиографической диагностике, М.: МЕДпресс-информ. 2013
- [2] С.Н. Коломиец: Азбука эхокардиографии, Одесса.: ОГМУ. 2010.
- [3] В.О. Горелов, Ю.Ю.Ляш, В.А.Ровінський, "Методи ущільнення даних та перетворення форми інформації у комп'ютерних іграх," Управління розвитком складних систем. 2018. № 35. С. 93-104.
- [4] В.А. Ровінський, О.Ю.Фуфалько, Ю.Й.Стрілецький, О.В.Євчук "Інформаційна технологія моделювання біполярних транзисторів хвиловими цифровими фільтрами," Вісник Вінницького політехнічного інституту. 2013. № 5(24). С. 81-84.
- [5] М. Л. Петришин, В. А. Ровінський, "АЦП порозрядного врівноваження для пришвидшеного перетворення технологічних параметрів," Інформаційні технології та комп'ютерна інженерія, №39, С. 42-52, 2017.
- [6] Ю. Й. Стрілецький, О.В.Євчук, В. А. Ровінський, "Реалізація аналогових активних елементів, керованих напругою, в системах цифрової обробки сигналів," Вісник Хмельницького національного університету, том.12 №4, С. 96-102, 2012.
- [7] М.Л. Петришин, В.А. Ровінський "Метод двовимірного кодування даних за допомогою рекурсивних кодових систем," Системи обробки інформації, №4, С. 56-58, 2016.
- [8] И.М. Андреев, " Описание алгоритма CART," Методы. Алгоритмы. Программы, №3-4(7-8), С. 48-53, 2004.
- [9] L.Breiman., J.H.Friedman, R.A. Olshen, C.T.Stone. Classification and Regression Trees.— Wadsworth, Belmont, California, 1984
- [10] J.R.Quinlan, C 4.5 Programs for Machine Learning.— Morgan Kaufmann, San Mateo, California,1993
- [11] Machine Learning, Neural and Statistical Classification / Editors: D. Michie, D. J. Spiegelhalter, C. C. Taylor.— London: Ellis Horwood, 1994
- [12] "Алгоритм CART," Лекция 9: Методы классификации и прогнозирования. Деревья решений, 2018. [Online]. Available: <https://www.intuit.ru/studies/courses/6/6/lecture/174?page=3/>. [Accessed 10 October 2018].

Використання автоматизованих інформаційних систем класу Learning Management System

М. І. Шевчук
Кафедра комп'ютерних наук та
інформаційних систем
ДВНЗ «Прикарпатський
національний університет імені
Василя Стефаника»
Івано-Франківськ, Україна
shevchukmisha@i.ua

М. В. Семаньків
Кафедра комп'ютерних наук та
інформаційних систем
ДВНЗ «Прикарпатський
національний університет імені
Василя Стефаника»
Івано-Франківськ, Україна
dlyamarii@gmail.com

Use of automated information systems of the class Learning Management System

М. Shevchuk
Department of Computer Science
Vasyl Stefanyk Precarpathian
National University
shevchukmisha@i.ua

M. Semankiv
Department of Computer Science
Vasyl Stefanyk Precarpathian
National University
Ivano-Frankivsk, Ukraine
dlyamarii@gmail.com

Анотація—У статті розглянуто систему дистанційного навчання Moodle та плагін для автоматизованої перевірки завдань CodeRunner на прикладі створення дистанційного курсу з мови програмування Python.

Abstract— The article considers the system of distance learning Moodle and a plugin for automated testing tasks of CodeRunner, on the example creation of a distance course in the Python programming language.

Ключові слова— дистанційне навчання, Moodle, Learning Management System, Python

Keywords—distance learning, Moodle, Learning Management System, Python

I. ВСТУП

Проблема вибору системи дистанційного навчання на початковому етапі вимагає обрати між безкоштовними та платними системами. На основі аналізу безкоштовних систем відзначено переваги системи дистанційного навчання Moodle – модульного об'єктно-орієнтованого динамічного навчального середовища, яке є у вільному доступі та відноситься до автоматизованих інформаційних систем класу LMS (Learning Management System) – систем управління навчанням [1].

Система дистанційного навчання Moodle орієнтована на організацію взаємодії між вчителем та учнем в процесі навчання, а також для організації традиційних дистанційних курсів, очного та заочного навчання.

II. МОЖЛИВОСТІ СИСТЕМИ ДИСТАНЦІЙНОГО НАВЧАННЯ MOODLE

Moodle навчальна платформа призначена для об'єднання вчителів, адміністраторів та учнів в одну систему для створення персоналізованого навчального середовища. Застосовується у більше, ніж 200 країнах світу. Роль вчителя полягає підвищенні мотивації й підтримці учнів через підготовку завдань для самостійного опрацювання, оцінювання результатів їх діяльності та коригування знань. І найбільші результати досягаються коли учні навчаються в співпраці, діляться своїм досвідом і думками.

Система дистанційного навчання відповідає основним критеріям, які вимагаються від систем дистанційного навчання, а саме:

- функціональність – наявність вибору функцій різного рівня, (форуми, чати, керування курсами);
- надійність – зручність адміністрування та керування навчанням, простота в оновленні контенту на базі шаблонів, захист курсів від несанкціонованого доступу;
- стабільність – високий рівень стійкості роботи системи у різних режимах роботи та активностях користувачів;
- вартість – використання системи безкоштовне, витрати на її впровадження, розробку курсів та супровід – мінімальні;
- відсутність обмежень на кількість учасників курсу;

- модульність – наявність в навчальних курсах блоків матеріалу, які можуть використовуватись в інших курсах;

- наявність вбудованих засобів розробки та редагування навчального контенту;

- наявність системи перевірки та оцінювання знань у режимі онлайн (тестування, завдання та ін.);

- зручність і простота використання та навігації;

Система дистанційного навчання Moodle надає можливість повноцінно організувати навчальний процес, включаючи засоби для навчання, систему контролю та оцінювання навчальної діяльності учнів, а також інші необхідні складові дистанційного навчання.[2, 3]

В системі Moodle пропонуються на вибір 3 типи форматів курсів: форум, структура (навчальні модулі без прив'язки до календаря), календар (навчальні модулі з прив'язкою до календаря).

III. РОЗРОБКА ДИСТАНЦІЙНОГО КУРСУ ДЛЯ ВИВЧЕННЯ МОВИ ПРОГРАМУВАННЯ PYTHON

Для вивчення мови програмування Python на основі вище вказаної системи дистанційного навчання розроблено практичний курс для учнів загальноосвітніх шкіл та студентів вищих навчальних закладів, який містить: вісім конспектів-лекцій на кожен підтему, в яких подано теоретичний матеріал; на основі теоретичного матеріалу розроблено практичні та тестові завдання, які оцінюються у 12-бальній шкалі; висвітлено додатковий матеріал, для тих, хто бажає знати більше, а саме відео-лекції, посилання на інші інтернет-ресурси, інші дистанційні курси, які не охоплюють шкільну тему.

У теоретичному матеріалі передбачено підготовку до вивчення мови програмування Python з початкового рівня: від встановлення IDLE до створення повноцінної програми. Кожне заняття подано у необхідному форматі із елементами самостійних завдань для зацікавлення учнів та студентів.

Для підтримки дистанційного навчання створено сайт на базі системи дистанційного навчання Moodle, який знаходиться за адресою <http://138.68.105.99/moodle/>.

В якості серверу для Moodle використано операційну систему Linux з використанням Apache, MySQL, PHP (Linux + Apache + MySQL + PHP = LAMP).

У якості сервера використовувано сервіс Digital Ocean, який надає гнучку можливість конфігурації серверу та сайтів, які на ньому знаходяться.

Під час проведення практичних занять використовуються модулі «Завдання» та «Тест». Завдання до виконання практичних робіт розміщуються в СДН Moodle. При підготовці до практичного заняття учні та студенти знайомляться зі змістом завдань, при необхідності повторюють теоретичний матеріал і

проходять пробне тестування. Кількість доступних спроб встановлюється налаштуваннями тесту.

Для додавання «Завдання» необхідно:

- увійти в режим редагування курсу, в меню «додати діяльність або ресурс обрати (завдання)» (рис. 1);

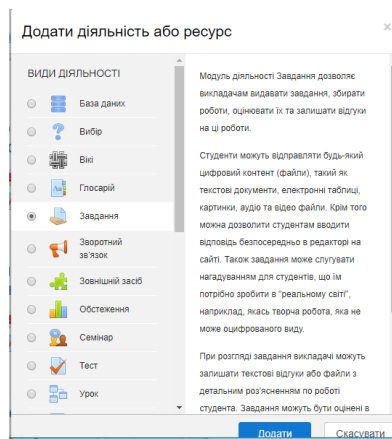


Рис. 1 – Вікно додавання «Завдання»

- у вікні, що відкрилось, ввести назву завдання, опис, відмітити «Показувати опис на сторінці курсу»;

- якщо необхідні терміни здачі виконаних робіт, обираєте необхідні дати в налаштуваннях, перед їх вибором необхідно активувати опцію «Включити», після закінчення кінцевого терміну подання робіт, учень більше не зможе надіслати свою роботу;

- у графі «Оцінка» вказуються максимальні бали за виконану роботу.

Для додавання питань до банку питань необхідно:

- в блоці Адміністрування вибрати Банк питань – Питання (рис.2) ;

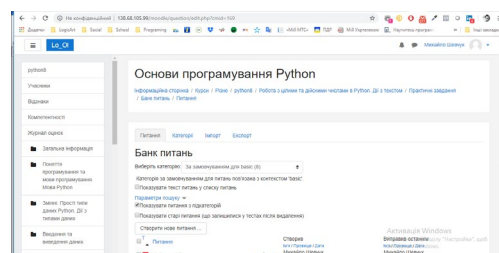


Рис. 2 – Банк питань

- в меню редагування банку питань необхідно обрати категорію, в яку буде додаватись питання та натиснути «Створити нове питання»;

- у вікні вибору типу питання найчастіше використовуються наступні типи питань: вбудовані відповіді, відповідність, коротка відповідь, множинний вибір, правильно/неправильно, CodeRunner;

Для автоматизованої перевірки завдань було проаналізовано наступні плагіни (надбудови), а саме: Code Runner (рис. 3) та Virtual Programming Lab (VLP). [4]

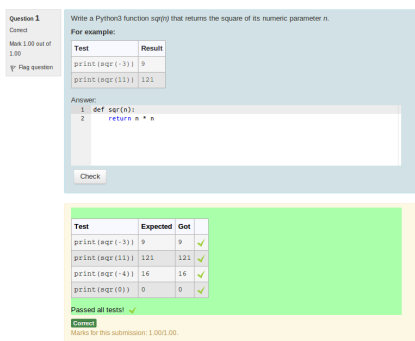


Рис. 3 – Приклад завдання з використанням плагіну CodeRunner

Аналіз показав, що плагін CodeRunner має більше вбудованих функцій, ніж VLP.

CodeRunner – типовий тип запитання Moodle, який дозволяє вчителям запускати програму, щоб оцінювати відповідь студента. На сьогоднішній день найбільш поширеним використанням CodeRunner є курси програмування, в яких студентам пропонується написати програмний код до деяких специфікацій, і тоді цей код класифікується, виконуючи його в серії тестів.

CodeRunner наразі підтримує Python2 (вважається застарілим), Python3, C, C++, Java, PHP, JavaScript (NodeJS), Octave та Matlab та інші. Архітектура дозволяє легко розширювати інші мови.

Тип запитання CodeRunner можна встановити на будь-якій сучасній системі Moodle (версії 2.9 або пізнішої, включаючи версію 3.2), на Linux, Windows і Mac.

Для створення завдань з автоматизованою перевіркою використовувався безкоштовний плагін CodeRunner (рис. 4), який підтримує компілювання багатьох мов програмування, зокрема Python 3.

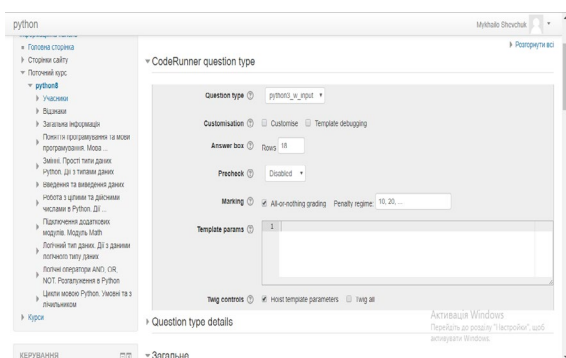


Рис. 4 – Створення питання у CodeRunner

У пункті Question type вибирається мова програмування на якій будуть відправлятися відповіді на сервер, у нашому випадку – python3_w_input. Модуль CodeRunner, також, передбачає можливість створення тест-кейсів (рис. 5) для перевірки правильності виконання програми та на основі результату оцінювання роботи у шкалі, яка задається при створенні питання.

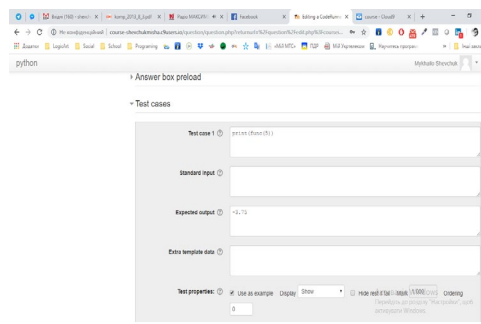


Рис. 5 – Створення «тест-кейсу»

У полі Test case 1 записується виклик функції, яку повинен користувач написати і параметри, які передаються. У expected output – результат виконання функції. Таких тест-кейсів можна створювати безліч.

ВИСНОВКИ

Підсумовуючи вище сказане, можна стверджувати, що процес навчання з використанням програмно-інструментальної платформи дистанційного навчання Moodle має ряд суттєвих переваг та зручні засоби управління контентом і різні форми організації занять. Дана система часто використовується для створення інформаційно-освітнього середовища навчального закладу, і орієнтована, насамперед, на забезпечення інтерактивної взаємодії між учасниками навчального процесу, застосовується для організації як заочного та дистанційного навчання, так і для підтримки очного навчання. На основі даної системи розроблено та апробовано на базі загальноосвітніх шкіл міста Івано-Франківськ навчально-методичні матеріали для вивчення мови програмування Python, в яких подається теоретичний матеріал, практичні завдання та завдання для перевірки знань учнів; всі вони наповнюють сайт для дистанційного навчання, що знаходиться за адресою <http://138.68.105.99/moodle>.

REFERENCES ЛІТЕРАТУРА

- [1] Джавала Л. Впровадження новітніх методів дистанційного навчання в системі освіти / Любомира Джавала // Інноваційні комп'ютерні технології у вищій школі : матеріали 4-ої Науково-практичної конференції, 20–22 листопада 2012 року, Львів / Національний університет "Львівська політехніка". – Львів : Видавництво Львівської політехніки, 2012. – С. 46–49. – Бібліографія : 3 назви.
- [2] Гасько Р. Створення електронних навчальних систем третього покоління: використанням робота віртуальної присутності та візуальної мови програмування / Роман Гасько // Інноваційні комп'ютерні технології у вищій школі : матеріали 4-ої Науково-практичної конференції, 20–22 листопада 2012 року, Львів / Національний університет "Львівська політехніка". – Львів : Видавництво Львівської політехніки, 2012. – С. 41–45. – Бібліографія: 7 назв.
- [3] Демида Б. Системи дистанційного навчання: огляд, аналіз, вибір / Б. Демида, С. Сагайдак, І. Копил // Вісник Національного університету "Львівська політехніка". – 2011. – № 694 : Комп'ютерні науки та інформаційні технології. – С. 98-107. – Бібліографія : 14 назв.
- [4] Фоменко А. Організація роботи з групами в середовищі Moodle 3 / Андрій Фоменко // Інноваційні комп'ютерні технології у вищій школі : матеріали 8-ої Науково-практичної конференції, 22–24 листопада 2016 року, Львів / Національний університет «Львівська політехніка». – Львів : Видавництво Наукового товариства ім. Шевченка, 2016. – С. 211–215.

Аналіз Методів Обробки Зображень в Системі Визначення Патологій Шкіри Людини

Наталія Герасименко
кафедра Прикладної математики
Харківський національний університет
радіоелектроніки
Харків, Україна
n.viktorovna18@gmail.com

Володимир Кобзев
доцент кафедри Прикладної математики
Харківський національний університет
радіоелектроніки
Харків, Україна
volodymyr.kobziev@nure.ua

Analysis of Image Processing Methods in Pathologies Determination System of Human Skin

Nataliia Herasymenko
Department of Applied Mathematics
Kharkiv national University of Radio Electronics
Kharkiv, Ukraine
n.viktorovna18@gmail.com

Volodymyr Kobziev
Department of Applied Mathematics
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
volodymyr.kobziev@nure.ua

Анотація—У роботі увага зосереджується на проблемі класифікації уражень шкіри, зокрема меланоми. Пропонується розбиття дермоскопічного зображення, що містить злоякісне чи доброякісне ушкодження шкіри. Запропоноване рішення побудовано на основі архітектури згорткової нейронної мережі VGGNet.

Ключові слова—ураження шкіри, меланома, медичні системи підтримки прийняття рішень, згорткові нейронні мережі.

Abstract—The work focuses on the problem of classification of skin lesions, in particular melanomas. It is suggested to split a dermoscopic image containing malignant or benign skin damage. The proposed solution is based on the architecture of the convolutional neural network VGGNet.

Keywords—skin lesions, melanoma, medical decision support systems, convolutional neural networks.

I. ВСТУП

Меланома є фатальною формою раку шкіри, яка часто не діагностується або неправильно діагностується як доброякісне ураження шкіри. За оприлюдненими оцінками [1], в Сполучених Штатах фіксується 76 380 нових випадків меланоми та 6750 смертей на рік. Раннє виявлення необхідно, тому що життя пацієнтів з меланомою залежить від точної і ранньої діагностики. Лікарі часто покладаються на особистий досвід та оцінюють рівень ураження в кожному конкретному випадку, беручи до уваги локальні ураження пацієнта в порівнянні з усім тілом [2].

Без комп'ютерної допомоги точність клінічного діагнозу для виявлення меланоми становить 65- 80% [3]. Використання дермоскопічних зображень покращує діагностичну точність уражень шкіри на 49% [4]. Візуальні відмінності між меланомою і доброякісними ураженнями шкіри можуть бути дуже незначними (рис. 1), що ускладнює встановлення їх відмінності навіть підготовленими медичними експертами.

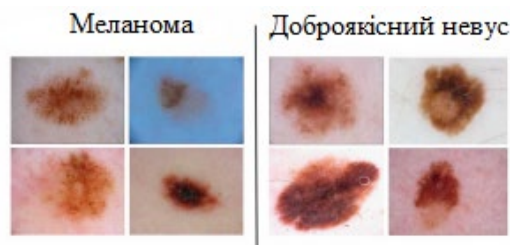


Рис. 1. Зразок зображень, створений з набору даних архіву ISIC [5]

У даній роботі пропонується метод класифікації пошкоджень шкіри за допомогою глибокого вивчення, який охоплює загальнодоступний дермоскопічний набір зображень шкірних покривів - набір даних архіву ISIC [5]. Він використовує існуючу архітектуру згорткової нейронної мережі VGGNet, яка є дуже глибокою згортковою мережею великомасштабного візуального розпізнавання.

II. ЗАПРОПОНОВАНЕ РІШЕННЯ

Для виконання вищевказаної задачі необхідно розглянути архітектуру згорткової мережі (ConvNet) та прояснити аспекти її реалізації.

A. Архітектура ConvNet

VGGNet є добре задокументованою та загальноприйнятою архітектурою для згорткових нейронних мереж [6]. Ця ConvNet стала популярна завдяки досягненню дивовижної продуктивності на наборі даних ImageNet [7]. Вона поставляється в декількох варіантах, з яких дві найефективніші (з 16 і 19 ваговими шарами) були оприлюднені. У цій роботі була обрана архітектура VGG16 адаптована з [6] (рис. 2), оскільки вона була добре впорядкована в інших наборах даних. Вхідний шар мережі чекає зображення у форматі RGB 224×224 пікселів. Вхідне зображення проходить через п'ять згорткових блоків. Використовуються невеликі згорткові фільтри з рецептивним полем 3×3.

Кожен збірний блок включає в себе операцію 2D-згортки (кількість фільтрів змінюється між блоками). Всі приховані шари обладнані ReLU (ректифікованим лінійним блоком) для виконання функцій рівня активації, включають просторовий пул за допомогою використання максимального об'єднання шару. Мережа пов'язується з блоком класифікаторів, що складається з трьох шарів Fully-Connected (FC).

Оригінальний VGG16 повинен бути змінений відповідно до наших потреб:

- Остаточний повністю підключений вихідний рівень повинен виконувати бінарну класифікацію (доброякісне проти злякисного).
- Функція активації в модифікованому шарі змінена на сигмоподібну.

B. Первинна обробка та збільшення даних

Вхідні зображення попередньо обробляють шляхом: (I) нормалізації значень пікселів до діапазону [0,1]; (II) обрізання зображення до квадратного співвідношення сторін (при необхідності); (III) зміни розміру зображення до очікуваного розміру 224×224 пікселів.

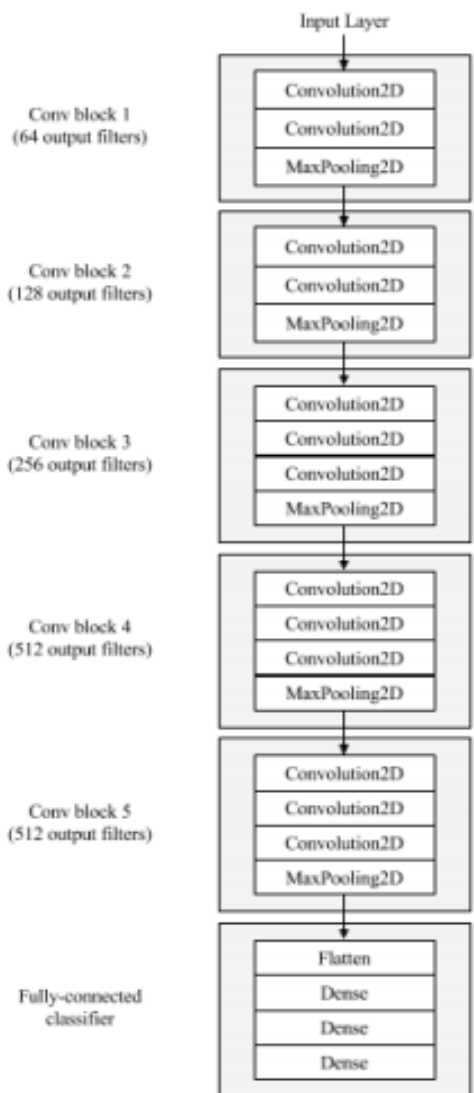


Рис. 2. Архітектура VGG16

Щоб максимально використати кілька навчальних прикладів та підвищити точність моделі, дані збільшені за допомогою ряду випадкових перетворень. Вибрані такі методи збільшення даних: повторне масштабування розміру, обертання 40°, горизонтальний зсув, масштабування зображення та горизонтальне перевертання. Крім того, очікується, що збільшення кількості даних також має допомогти уникнути перенавчання (загальна проблема з малими наборами даних, коли модель піддана занадто малій кількості прикладів та вивчає схеми, які не узагальнюють нові дані), і з цієї причини покращить здатність моделей до узагальнення.

C. Налаштування мережі

Одним із поширених методів передачі навчання полягає не лише в перепідготовці класифікатора у верхній частині мережі з новим набором даних, але і в застосуванні тонкої настройки мережі шляхом навчання лише частини верхнього рівня згорткових шарів та продовження зворотного розповсюдження. У цій роботі пропонується заморожувати рівні нижчого рівня мережі, оскільки вони містять більше загальних характеристик набору даних. Нас цікавить тренування лише верхніх шарів мережі через їх здатність виконувати вилучення більш специфічних функцій. У цьому підході перші чотири згорткові шари в остаточній архітектурі ініціалізуються з вагами з набору даних ImageNet. П'ятий згортковий блок ініціалізується випадковими вагами.

D. Аспекти реалізації

Для реалізації архітектури нейронної мережі пропонується використати Keras [20]. Він забезпечує шар абстракцій поверх Theano [21], який використовується як основна структура нейронної мережі.

Keras допускає:

- модульність - користувачі можуть створювати свою мережу послідовності, яка представляє собою лінійний стек шрифтів;
- мінімалізм - функції, включені в бібліотеку, дозволяють користувачеві легко створювати та змінювати мережеві рівні;
- розширюваність - щоденні оновлення забезпечують вирішення поточних проблем, з якими зустрічаються дослідники глибокого навчання.

Крім того, Keras працює в середовищі Python, що дає користувачам свободу використання додаткових залежностей Python, включаючи SciPy [22] і PIL [23]. В додаток до Keras, необхідні бібліотеки CUDA [24] для управління графічними процесорами NVIDIA GeForce GTX 1080 Ti, що використовуються для навчання та оцінки реалізації [25].

III. ЕКСПЕРИМЕНТИ ТА РЕЗУЛЬТАТИ

Для практичної перевірки позначеного підходу проведена низка експериментів з використанням запропонованих методів та вказаних аспектів реалізації.

А. Набір даних

Для експериментів був використаний набір даних ISBI 2016 року для аналізу пошкоджень шкіри на предмет виявлення меланоми. Набір даних містить репрезентативну суміш зображень, позначених як доброякісні або злоякісні, попередньо розбиті на групи з 900 навчальних зображень і 379 тестових зображень [5].

В. Параметри

Всі методи були реалізовані в Keras. Функцією оптимізації вибрана RMSProp [26], функція втрат описана в [27]. Для оптимізації випадання в повністю підключених шарах використовується значення 0,5. Розмір партії з 16 зображень обраний через невеликий розмір існуючого набору даних. Набір даних збалансований за допомогою дискретизації. В алфавітному порядку були відібрані перші 173 зображення з кожного класу в наборі навчальних матеріалів, і були обрані перші 75 зображень в кожному класі з набору тестових даних. Підсумковий набір даних складався з 346 навчальних зображень і 150 тестових зображень. Для збільшення даних використані такі варіанти: масштабування розмірів, повороти (кути), горизонтальний зсув, масштабування (коефіцієнт) і горизонтальне відображення.

С. Результати

Оцінка моделі виконується з використанням навчального та тестового наборів, які використовуються в наборі даних ISIC.

Використовуються наступні показники:

- функція втрати, визначена як кількісна оцінка узгодження між прогнозованими зображеннями та істинними мітками.
- чутливість, частка справжніх позитивних фактів, котрі правильно ідентифіковані.
- точність, частка завантажених екземплярів, які є релевантними.
- специфічність, частка справжніх негативів, які правильно ідентифікуються.
- точність, кількість правильних прогнозів, поділених на загальну кількість прогнозів.

Результати навчання та тестування представлені в таблиці 1.

ТАБЛИЦЯ 1. Оцінка моделі

	Похибка	Точність засобів вимірювання	Чутливість	Точність результату вимірювання
Навчальний набір	0.4891	76.88%	0.6903	0.8259
Тестовий набір	0.4337	81.33%	0.7866	0.7974

ВИСНОВКИ

Проаналізовані існуючі методи діагностики уражень шкіри. Розроблений двокласовий класифікатор, який отримує входні знімки з ушкодженнями шкіри, буде моделлю з використанням глибинних нейронних мереж зсуву та використовує цю модель для прогнозування того, чи є ураження шкіри доброякісним або злоякісним.

Запропонований підхід дає багатообіцяючі результати - його значення чутливості 78,66% та точності 79,74% значно вищі, ніж сучасний стан розпізнавання цього набору даних (50,7% та 63,7%, відповідно).

ЛІТЕРАТУРА REFERENCES

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2016," CA: a cancer journal for clinicians, vol. 66, no. 1, pp. 7–30, 2016.
- [2] J. Gachon, P. Beaulieu, J. F. Sei, J. Gouvernet, J. P. Claudel, M. Lemaitre, M. A. Richard, and J. J. Grob, "First prospective study of the recognition process of melanoma in dermatological practice," Archives of dermatology, vol. 141, no. 4, pp. 434–438, 2005.
- [3] G. Argenziano and H. P. Soyer, "Dermoscopy of pigmented skin lesions—a valuable tool for early diagnosis of melanoma," The Lancet Oncology, vol. 2, no. 7, pp. 443–449, 2001.
- [4] H. Kittler, H. Pehamberger, K. Wolff, and M. Binder, "Diagnostic accuracy of dermoscopy," The lancet oncology, vol. 3, no. 3, pp. 159–165, 2002.
- [5] "International Skin Imaging Collaboration: Melanoma Project Website," <https://isic-archive.com/>.
- [6] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
- [7] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein et al., "Imagenet large scale visual recognition challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, 2015.
- [8] E. H. Page, "Description of skin lesions," <https://goo.gl/m9ybFp>.
- [9] S. Dreiseitl, L. Ohno-Machado, H. Kittler, S. Vinterbo, H. Billhardt, and M. Binder, "A comparison of machine learning methods for the diagnosis of pigmented skin lesions," Journal of biomedical informatics, vol. 34, no. 1, pp. 28–36, 2001.
- [10] N. Codella, Q.-B. Nguyen, S. Pankanti, D. Gutman, B. Helba, A. Halpern, and J. R. Smith, "Deep learning ensembles for melanoma recognition in dermoscopy images," arXiv preprint arXiv:1610.04662, 2016.
- [11] C. Barata, M. Ruela, M. Francisco, T. Mendonca, "Two systems for the detection of melanomas in dermoscopy images using texture and color features," IEEE Systems Journal, vol. 8, pp. 965–979, 2014.
- [12] M. Walter, "Is this the end? machine learning and 2 other threats to radiologists future," goo.gl/M9X3SF, 2016.
- [13] S. Jha, "Will computers replace radiologists?" <http://www.medscape.com/viewarticle/863127>, 2016.
- [14] J. Kawahara, A. BenTaieb, and G. Hamarneh, "Deep features to classify skin lesions," IEEE International Symposium on Biomedical Imaging (IEEE ISBI), pp. 1397–1400.
- [15] H. Liao, "A deep learning approach to universal skin disease classification," [https://www.cs.rochester.edu/u/hliao6/projects/other/skin project report.pdf](https://www.cs.rochester.edu/u/hliao6/projects/other/skin%20project%20report.pdf).
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv preprint arXiv:1512.03385, 2015.
- [17] "Dermofit image library," <https://licensing.eri.ed.ac.uk/i/software/dermofit-image-library.html>.
- [18] "Dermnet – skin disease atlas," <http://www.dermnet.com/>.
- [19] "IEEE International Symposium on Biomedical Imaging," <http://biomedicalimaging.org/>.
- [20] "Keras documentation," <https://keras.io/>.
- [21] "Theano 0.8.2. documentation," <http://deeplearning.net/software/theano/>.
- [22] "Scipy Python Library," <https://www.scipy.org/>.
- [23] "Python Imaging Library (PIL)," <http://www.pythonware.com/products/pil/>.
- [24] "CUDA, Nvidia," http://www.nvidia.com/object/cuda_home_new.html.
- [25] "NVIDIA GeForce GTX 1080 Ti," <https://www.nvidia.com/en-us/geforce/products/10series/geforce-gtx-1080-ti/>.
- [26] "RMSProp Optimizer," <https://keras.io/optimizers/#rmsprop>.

Вибір Алгоритму Стемінгу Для Ефективного Повнотекстового Пошуку

Марина Кальметьєва
Кафедра Програмної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
mkalmetieva@gmail.com

Володимир Кобзєв
Кафедра Програмної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
volodymyr.kobziev@nure.ua

Selecting a Stemming Algorithm for Effective Full-Text Search

Marina Kalmetieva
Software Engineering department
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
mkalmetieva@gmail.com

Volodymyr Kobziev
Software Engineering department
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
volodymyr.kobziev@nure.ua

Анотація—У роботі визначена роль стемінгу для повнотекстового пошуку та загальні проблеми при використанні стемінгу. Розглянуті різні алгоритми стемінгу для англійської мови, зокрема, наявні у бібліотеці Apache Lucene, виявлені особливості алгоритмів, їх переваги та недоліки. Сформовані принципи для вибору найбільш ефективного алгоритму стемінгу в залежності від наявних умов та вимог до програмної системи.

Ключові слова—повнотекстовий пошук, стемінг.

Abstract—The paper identifies the role of stemming for full text search and common problems with using stemming. Different algorithms of stemming for English language are considered, in particular, available in the library of Apache Lucene, features of algorithms, their advantages and disadvantages are revealed. The established principles for choosing the most effective stemming algorithm depending on the existing conditions and requirements for the software system.

Keywords—full-text search, stemming.

I. ВСТУП

На даний момент повнотекстовий пошук є надзвичайно популярним. Він активно використовується у різноманітних інформаційних системах, і можна з впевненістю очікувати, що його розповсюдженість буде ще зростати. Адаже при роботі з великою кількістю даних набагато зручніше мати потужний, гнучкий та швидкий інструмент пошуку, схожий на Google, аніж обмежувати себе пошуком лише за точним входженням слова або використовувати регулярні вирази, виконання яких може бути повільним.

Для забезпечення високої швидкості обробки запитів у системах повнотекстового пошуку використовується така структура як інвертований індекс [1]. Інвертований індекс складається зі списку всіх слів, які містяться в будь-якому документі, і для кожного слова зберігається список документів, в яких слово присутнє. Документ у даному випадку – це одиниця даних, котру ми шукаємо, наприклад, наукова стаття чи кулінарний рецепт. Слова, що потрапляють у індекс, називаються термами.

Людська мова є дуже різнобічною, і часто те, що ми шукаємо, може бути присутнім у тексті в різних формах. Наприклад, коли ми шукаємо документи за словом «relax», нам буде все одно, чи написано це слово в документі з малої літери (в середині речення) чи з великої літери (на початку речення). Також, скоріш за все, нам підійдуть документи, де це слово присутнє в третій особі – «relaxes» – чи в іншому часі – «relaxed».

Таким чином, пошук по одному слову насправді може означати пошук по багатьох, хоча й схожих словах. У системах повнотекстового пошуку це вирішується шляхом використання попередньої обробки слів для перетворення їх на терми, що потім будуть записані в інвертований індекс. Попередня обробка може включати:

– Ігнорування стоп-слів, або шумових слів. Це слова, що не несуть смислового навантаження, тому їх користь та роль для пошуку не суттєва. Зазвичай це артиклі (a, an, the), сполучники (and, or), прийменники (in, on), займенники (I, he, she), форми дієслова to be (be, have, had). Ці слова просто не включаються до інвертованого індексу.

– Знаходження синонімів до слова у спеціальному словнику синонімів. Для слова «relax» ми можемо також віднайти слово «rest».

– Перевірка слова на одрук. Це допоможе виявити, що написавши «realx», людина на справді мала на увазі «relax».

– Стемінг, або зведення слова до його кореня. Стемінг перетворить слова «relax», «relaxes» та «relaxed» на один і той же терм «relax», що є коренем цього слова. Таким чином у індекс потрапляє лише обрізана форма «relax», а ось посилання від цього терму в індексі будуть присутні на всі документи, де є слово «relax» в будь-яких особі або часі.

Список методів попередньої обробки не обмежується вищезазначеними, однак ми розглянемо детальніше саме стемінг.

II. СУТНІСТЬ СТЕМІНГУ

Сутність стемінгу полягає у спробі зменшити кожне слово до його кореневої форми, або форми, близької до кореневої [2]. Таким чином ми зводимо різні, але однокореневі слова до єдиного загального терму. В англійській мові найбільш розповсюдженими є такі зміни форми слова:

- Різні частини мови: add, addition, additive.
- Особа: play, plays.
- Час: play, played, playing.
- Число: cat, cats.

Головна ідея можливості стемінгу полягає в тому, що саме корінь є основною значущою частиною слова, і відкинувши інше, у багатьох випадках ми зможемо зберегти смислову сутність слова.

Однак це не покриває всі випадки, і таким чином може виникнути проблема надлишкового стемінгу (overstemming). Надлишковий стемінг полягає в тому, що через обрізання двох слів з різним значенням ми зводимо їх до єдиного терму і таким чином змішуємо зміст слів в один. Це можна проілюструвати на прикладі слів «universe» та «university». Якщо наша система знає, що літера «e» часто присутня наприкінці дієслів, а «ity» – наприкінці іменників, та ці частки не мають істотного змісту, обидва наші слова будуть зменшені до терму «univers». Тепер при пошуку по одному із цих слів ми будемо знаходити і документи, де присутнє «university», і документи з «universe». Однак це є помилкою, адже початкові слова мають різне значення.

Для боротьби з проблемою надлишкового стемінгу, можна зменшити ступінь обрізання слів, тобто «пом'якшити» стемінг, однак при цьому може виникнути протилежна проблема – недостатній стемінг (understemming). Вона характеризується тим, що слова з однаковим значенням приводяться не до одного терму, а до різних. Наприклад, при дуже м'якому стемінгу наша система могла б не пов'язати між собою слова «relaxed» та «relaxing», вважаючи їх абсолютно відмінними.

Існують різні алгоритми стемінгу, а комп'ютерна програма, що реалізує певний такий алгоритм, називається стемером.

III. ОСНОВНІ ВИДИ СТЕМЕРІВ

Стемери можна поділити на дві великі групи: алгоритмічні та словникові.

Алгоритмічні стемери містять у собі ряд правил, які застосовують до слів, наприклад відкидання «s» з кінця слова, щоб прибрати закінчення третьої особи. Вони досить швидкі, не потребують багато додаткової пам'яті для своєї роботи, не містять ніякої інформації про окремі слова, лише загальні правила. Однак важливим недоліком є те, що ці стемери не можуть добре обробляти неправильні (irregular) слова, такі як дієслова «eat» – «ate», або іменники на кшталт «tooth» – «teeth».

Словникові стемери, на відміну від алгоритмічних, не містять майже жодних правил. Їх суть полягає у використанні словника, що попередньо заповнюється можливими формами слів, які ми очікуємо зустріти в корпусі документів. Кожне таке слово у словнику вказує на цільову обрізану версію слова. Це дозволяє стемеру

правильно обробляти неправильні слова («eat» – «ate») та розрізняти слова, схожі за формою, але різні за значенням («universe» – «university»). Головний недолік таких стемерів – досить важко записати усі форми слів, що очікуються, тобто якість роботи стемера дуже сильно залежить від якості заповненого словника. Суттєво, що підтримка такого словника та швидкого доступу до нього вимагає значних витрат по пам'яті.

У більшості випадків для використання у програмній системі буде обраний саме алгоритмічний стемер.

Розглянемо найбільш відомі алгоритмічні стемери, наявні для англійської мови.

A. Стемер Ловінса (Lovins stemmer) [3]

Це перший опублікований алгоритм стемінгу. Він був описаний Дж. Ловінсом у 1968 році для заклав основи для популярних зараз стемерів. На сьогодні стемер Ловінса вважається громіздким та не дуже надійним: для своєї роботи він використовує таблицю, у якій зібрано приблизно 300 існуючих суфіксів слів. Алгоритм намагається вирізати зі слова найдовший можливий суфікс.

Однак навіть таблиця такого розміру не є вичерпною, тож частина суфіксів в ній відсутня, що знижує якість стемінгу. Стемер Ловінса є досить жорстким, іноді може давати й зовсім неочікувані результати, дуже сильно обрізаючи слово.

B. Стемер Портера (Porter stemmer) [4]

Стемер Портера базується на алгоритмі стемера Ловінса, однак з деякими важливими змінами. Основною ідеєю є те, що існує багато суфіксів англійської мови, кожен з яких насправді розбивається на декілька маленьких суфіксів. Це дозволило спростити таблицю суфіксів, що підтримуються, та зробити правила по відсіченню суфіксів більш загальними. Цей алгоритм виконує помітно м'якший стемінг, ніж алгоритм Ловінса.

B. Стемер Портера 2 (Porter2 stemmer)

Нарівні зі стемером Портера була розроблена друга, трохи модифікована версія цього алгоритму – стемер Портера 2 (Porter2 stemmer). Porter2 відрізняється від Porter тим, що стемінг відбувається декілька швидше, але і є декілька жорсткішим.

На даний момент обидві версії алгоритму Портера є найбільш популярними. У таблиці 1 наведені результати роботи стемера Porter2 на різних формах слів.

Можна побачити, що в цілому стемер Porter2 добре справляється зі своєю задачею. Але в результатах наявні помилки як недостатнього стемінгу («adding» зводиться до «ad», а «adds» до «add»), так і надлишкового («universe» та «university»), тож треба буди готовим до цього.

Загалом проблема надлишкового стемінгу є досить неприємною. Адже навряд чи людина захоче знаходити дані про якийсь університет («university»), якщо вона робила пошук про космос («universe»), та навпаки. Це особливо суттєво, коли необхідний пошук за точними науковими термінами, а не просто за схожими словами.

ТАБЛИЦЯ 1. ВИКОРИСТАННЯ PORTER2 НА РІЗНИХ ФОРМАХ СЛІВ

Початкове слово	Porter2	Початкове слово	Porter2
marker	mark	added	ad
markers		add	add
market	market	user's	user
add	add	organization	organ
addition	addit	organ	organ
additive	addit	university	univers
adding	ad	universe	univers

Для більш точного пошуку бібліотека Apache Lucene надає спрощені реалізації алгоритмічних стемерів. Ці стемери містять меншу кількість правил та виконують м'якший стемінг. Розглянемо ці стемери в порядку зменшення їх жорсткості.

Г. Kstem

Цей стемер поєднує у собі алгоритмічний та словниковий підходи. Загалом він використовує лише три прості правила: зведення множини до однини («books» – «book»), перетворення минулого часу на теперішній («looked» – «look») та прибирання закінчення -ing («looking» – «look»). Також цей стемер містить у собі обмежений словник, що дозволяє знаходити відповідність між такими словами як «amplification» та «amplify», «european» та «europe».

Д. Мінімальний (Minimal, EnglishMinimalStemmer)

Цей стемер виконує мінімальну функціональність, що полягає в зведенні до однини: «books» – «book».

Е. Присвійний (Possessive, EnglishPossessiveFilter)

Стемер прибирає ознаки присвійної форми слова: «user's» – «user».

Результати роботи спрощених стемерів на різних формах слів приведені у таблиці 2. Для тестування були використані ті ж самі слова, що і для перевірки алгоритму Porter2 (таблиця 1), відсутні слова, що для усіх стемерів залишилися незмінними після стемінгу.

ТАБЛИЦЯ 2. ВИКОРИСТАННЯ СПРОЩЕНИХ СТЕМЕРІВ НА РІЗНИХ ФОРМАХ СЛІВ

Початкове слово	Kstem	Minimal	Possessive
markers	marker	marker	markers
adding	add	adding	adding
added	add	added	added
add	add	add	add
user's	user's	user'	user

Можна помітити, що Мінімальний стемер є дещо грубим, адже присвійна форма «user's» обрізана просто шляхом викидання останньої букви «s», і отриманий терм, до якого увійшов апостроф. Усі три стемери не мають помилки надлишкового стемінгу, адже слова

«university», «universe», «organization», «organ» не підпадають під жодне їх правило. Також для стемеру Kstem відсутня проблема недостатнього стемінгу, і всі форми дієслова «add» приводяться до єдиного терму.

IV. ВИБІР ОПТИМАЛЬНОГО СТЕМЕРУ

Вибір певного алгоритму стемінгу залежить в першу чергу від того, який ступінь жорсткості обрізання слів є найбільш сприятливим для користувачів програмної системи. При значному розмірі корпусу документів на вибір алгоритму також можуть впливати вимоги до швидкості та наявні технічні ресурси, зокрема пам'ять.

Тут вже постає питання віднайдення певної рівноваги. Жорсткий стемер може призвести до знаходження нерелевантних результатів через змішування змісту різних слів. М'який стемер може знаходити менше результатів, ніж хотілося б, бо буде вважати схожі слова зовсім різними та індексувати окремо, до того ж розмір такого індексу може бути дуже великим.

Використання жорсткого стемеру (Lovins або Porter2) має сенс для максимального угруповання схожих слів у задачі пошуку ключових слів [5]. Більш агресивний стемінг також корисний у випадках, коли результати пошуку направляються на подальшу машинну обробку, наприклад, споживаються алгоритмом кластеризації.

У випадку коли результати пошуку призначені для споживання людиною, м'який стемер зазвичай буде призводити до кращих результатів [6]. Найбільш м'яким серед популярних стемерів є Porter.

Якщо ж програмна система вимагає дуже «тонкого» пошуку, тобто надзвичайно важливо, щоб було знайдене саме те, що шукається, а не схожі слова, тоді можна використовувати стемери зі спрощеним набором правил, що, зокрема, надає Apache Lucene. Прикладом може бути система з пошуком по точних наукових термінах у медичних текстах.

ВИСНОВКИ

Розглянута роль стемінгу для повнотекстового пошуку, проаналізовані популярні стемери, а також спрощені алгоритми, що надають дуже м'який стемінг. Встановлено, що не існує єдиного оптимального алгоритму стемінгу, що підійшов би кожній програмній системі. На вибір алгоритму може впливати специфіка тексту документів, наявні обмеження щодо швидкості пошуку та доступної пам'яті, а також те, як саме використовуються результати пошуку.

ЛІТЕРАТУРА REFERENCES

- [1] Grainger T., Potter T., "Solr in Action", Manning, 2014, с. 11-15.
- [2] Hatcher E., Gospodnetic O., "Lucene in Action", Manning, 2004, с. 102-148.
- [3] Lovins J. B., "Development of a stemming algorithm," Mechanical Translation and Computer Linguistic., vol.11, no.1/2, с. 22-31, 1968.
- [4] Porter M.F., "An algorithm for suffix stripping". Program, 1980; 14, с. 130-137.
- [5] Вуль В.А., "Электронные издания». БХВ-Петербург, 2003, с. 411-412.
- Gormley C., Tong Z., "Elasticsearch: The Definitive Guide", O'Reilly Media, Inc., 2015, с. 369-371

CQRS-підхід в Розробці Програмного Забезпечення

Ілля Шутєєв
Кафедра Програмної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
iliaking97@gmail.com

Володимир Кобзєв
Кафедра Програмної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
volodymyr.kobziev@nure.ua

CQRS-Approach in Software Development

Illia Shutieiev
Software Engineering department
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
iliaking97@gmail.com

Volodymyr Kobziev
Software Engineering department
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
volodymyr.kobziev@nure.ua

Анотація—В основі CQRS лежить просте поняття використання різних моделей для поновлення і читання інформації. Однак це просте поняття веде до серйозних наслідків в проектуванні інформаційних систем. Ці наслідки розглянуті на прикладі додатку, що взаємодіє з CRM-системами.

Ключові слова – CQRS, CRM, архітектура, сховище даних.

Abstract—At the heart of CQRS is a simple notion of using different models for updating and reading information. However, this simple concept leads to serious consequences in the design of information systems. These implications are discussed on an example application that interacts with CRM systems.

Keywords - CQRS, CRM, Architecture, Data Warehouse.

I. ВСТУП

Command-query separation (CQS) або command-query responsibility segregation (CQRS) – це принцип імперативного програмування, винайдений Бертраном Мейером [1] під час роботи над мовою програмування Eiffel. Принцип вказує, що метод повинен бути або командою, яка виконує якусь дію, або запитом, що повертає дані, але не одночасно і тим, і іншим. Інакше кажучи, задавання питання не повинно змінювати відповідь. Більш формально, повертати значення можна тільки чистими методами, тобто такими, які не мають побічних ефектів. У даній роботі цей підхід порівнюється з класичним підходом до зчитування та запису даних.

Навіть за межами контрактного програмування, застосування CQRS розглядається його прихильниками як спосіб спрощення програми, роблячи доступ до її стану (через запити) і зміну її стану (через команди) більш зрозумілими, аналогічно тому як уникнення використання команди [goto](#) спрощує розуміння того, як програма виконується.

II. ОСНОВНА ТЕОРІЯ CQRS

A. CQRS на рівні методів класу

Розглянемо принцип Command-query separation (CQS). Основна ідея CQS, згідно [1, 2], полягає в тому, що в об'єкті методи можуть бути двох типів:

- **Queries:** Методи повертають результат, не змінюючи стан об'єкта. Іншими словами, у Query не існує будь яких side effects.
- **Commands:** Методи змінюють стан об'єкта, не повертаючи значення. Насправді коректніше називати ці методи modifiers або mutators, але так історично склалося, що вони називаються командами.

Для прикладу розглянемо клас User з одним методом IsValidEmail:

```
public class User
{
    public string Email { get; private set; }

    public bool IsValidEmail(string email) // Query
    {
        return Regex.IsMatch("email pattern", email);
    }

    public void ChangeEmail(string email) // Command
    {
        if (IsValidEmail(email) == false)
            throw new ArgumentOutOfRangeException(email);
        Email = email;
    }
}
```

Ми запитуємо у користувача (робимо Query) чи є дійсним email. Якщо так, то чекаємо у відповідь true, інакше false. Крім повернення значення, програміст, який писав метод IsValidEmail, вирішив в разі

валідного email відразу привласнювати його значення (робимо Command) полю Email.

Скористаємося принципом CQS і розділимо методи на Command і Query:

```
public class User
{
    public string Email { get; private set; }

    public bool IsEmailValid(string email) // Query
    {
        return Regex.IsMatch("email pattern", email);
    }

    public void ChangeEmail(string email) // Command
    {
        if (IsEmailValid(email) == false)
            throw new ArgumentOutOfRangeException(email);
        Email = email;
    }
}
```

B. CQRS на рівні дизайну додатка

Точно така ж ідея лежить в основі принципу Command Query Responsibility Segregation (CQRS). Але в цьому випадку принцип піднімається на рівень об'єктів, а не методів в об'єкті. Для зміни стану системи робиться клас Command, а для вибірки даних клас Query. Таким чином, ми отримуємо набір об'єктів, які змінюють стан системи, і набір об'єктів, які повертають дані.

Якщо клас описується як Command, то він:

- Змінює стан системи.
- Нічого не повертає.
- Добре описує предметну область, як дії користувачів над системою.
- Контекст команди зберігає потрібні для її виконання дані.

Приклад класу, який є командою:

```
public class DeleteUserCommand : ICommand<DeleteUserContext>
{
    private readonly ISession session;

    public DeleteUserCommand(ISession session)
    {
        this.session = session;
    }

    public void Execute(DeleteUserContext context)
    {
        session.Delete<User>(context.UserId);
    }
}
```

Зауваимо, що у класі типу ICommand є тільки один метод Execute, який повертає void [3]. Так виглядатимуть всі команди в вашому проекті. Клас DeleteUserContext – це той самий контекст, який несе в собі дані, необхідні для команди.

Якщо клас описується як Query, то він:

- Не змінює стан системи. Ніяких side effects.

- Контекст запиту зберігає потрібні для її виконання дані (пейджинг, фільтри і т.п.).
- Повертає результат.

Приклад класу, який є запитом:

```
public class FindUserIdQuery : IQuery<FindByIdContext, User>
{
    private readonly ISession session;

    public FindUserIdQuery(ISession session)
    {
        this.session = session;
    }

    public User Ask(FindByIdContext context)
    {
        return session.Query<User>()
            .SingleOrDefault(x => x.Id == context.Id);
    }
}
```

Зауважимо, що у класі типу IQuery є тільки один метод Ask, який повертає дані. Так виглядатимуть всі запити в вашому проекті. Клас FindByIdContext - це контекст, який несе в собі дані, необхідні для запиту.

III. ПРОБЛЕМНИЙ ШАБЛОН REPOSITORY

Реалізація шаблону Repository [4] фактично перетворюється в статичний клас (навіть, якщо у нього немає слова static) з великою кількістю методів. Крім цього, виникає низка проблем під час реалізації шаблону Repository, на які немає простих відповідей. Наведемо декілька з них:

- Незрозуміло, який саме репозиторій використовувати, якщо метод вибірки працює з декількома сутностями.
- Незрозуміло, що робити, коли різні репозиторії повинні використовувати закриті методи один одного.
- Тяжко зрозуміти чи необхідний взагалі шар сервісів.

Найпростішим робочим рішенням є поділ репозиторіїв на невеликі об'єкт-запити. У проекті ця реалізація тісно переплітається з концепцією CQRS. Суть в тому, що у CQRS є набір об'єктів, кожен з яких робить одну операцію читання. Тільки читання, без запису. Одна логічна операція читання дорівнює одному об'єкту запиту.

IV. ПРИКЛАД ЗАСТОСУВАННЯ CQRS

Підхід CQRS було вирішено використовувати під час розробки проекту, що взаємодіє з CRM-системою. Основною задачею веб-сайту було зчитування даних з CRM, їх коректне відображення та запис усіх змін даних.

Відомо [5], що CRM - управління відносинами з клієнтами (Customer relationship management) – це поняття, що охоплює концепції, які використовуються компаніями для управління взаємовідносинами зі споживачами, включаючи збір, зберігання й аналіз інформації про споживачів, постачальників, партнерів та інформації про взаємовідносини з ними.

Архітектура проекту наведена на рис. 1.

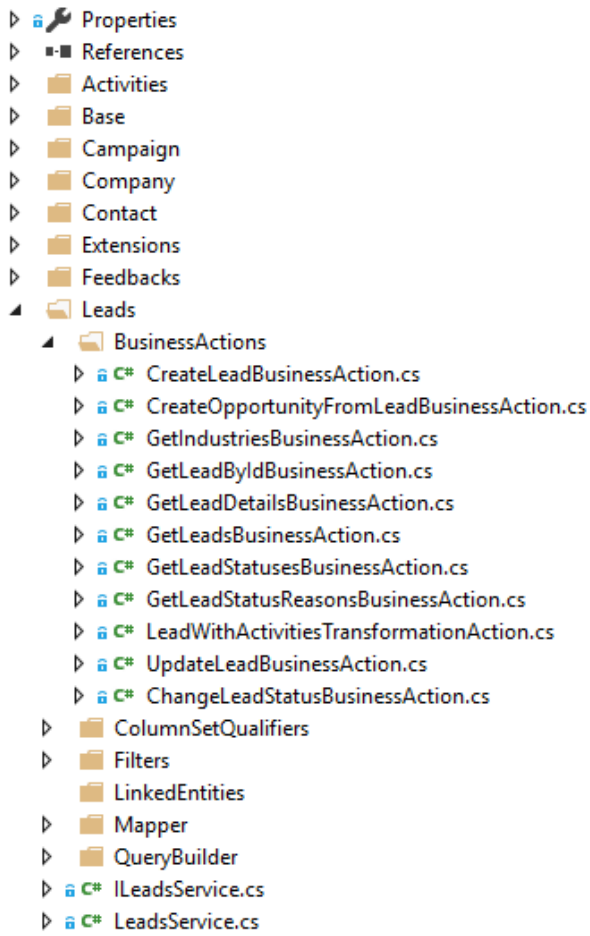


Рис. 1 – Архітектура проекту

Як бачимо, для кожної дії є свій Business Action, що виконує лише одну дію. Це дозволяє комбінувати такі елементарні дії в шарі сервісів (LeadService на рис. 1). Розглянемо приклад такого комбінуння:

```
public BaseLeadModel CreateLead(CreateLeadRequest request)
{
    // create lead
    var leadId = _createLeadBusinessAction.Execute(request);

    var company =
_getCompanyByIdBusinessAction.Execute(new
GetCompanyByIdRequest(request.Data.CompanyId));

    // set the same company industry if it is not set
    if (company.InheritPropertiesFromLead())
```

```
{
    _setInheritedPropertiesFromLeadBusinessAction.Execute(
        new UpdateInheritedFieldsRequest(
            company.Id,
            request.Data.Industry?.Id,
            request.Data.Potential?.Id));
}

// create activity notifying about status
_addChangePropertyActivityBusinessAction.Execute(
    new AddChangePropertyActivityRequest
    {
        NewValue = ((int)LeadStatus.Assigned).ToString(),
        Property = PropertyChange.Status,

        RegardingObjectType = Constants.LeadLogicalName,
        Subject = $"Status changed to
{LeadStatus.Assigned.ToString()}"
    });

// return new lead
return _getLeadByIdBusinessAction.Execute(new
GetLeadByIdRequest(leadId));
}
```

Як бачимо, в методі створення сутності використовуються дії з отримання, створення та оновлення сутностей. Такий підхід робить код легко зрозумілим, легко маштабуємим та знижує поріг ознайомлення з проектом.

ВИСНОВКИ

CQRS-підхід добре підходить для методології об'єктно-орієнтованого програмування, але може бути застосований і поза ООП, адже поділ всіх операцій на команди і запити не вимагає ООП.

CQRS-підхід може бути застосований в будь-якій парадигмі програмування, де потрібно турбуватись про побічні ефекти.

ЛІТЕРАТУРА REFERENCES

- [1] Meyer, Bertrand (1988). Object-oriented Software Construction. Prentice Hall. ISBN 0-13-629049-3.
- [2] Jeffrey Richter CLR via C #; Петроград. Друкарня Суворіна. Новий час - Москва, 2012. - 896 с.
- [3] Троелсен, Ендрю Мова програмування C # 5.0 і платформа .NET 4.5 / Ендрю Троелсен. - М.: Вільямс, 2015. - 486 с.
- [4] Steven John Metsker Design Patterns in C #; Фортуна Ел - Москва, 2011. - 480 с.
- [5] С. Шовкопляс. Як підвищити конкурентоспроможність за допомогою CRM // Office - 2005. - № 3-4, с. 12-18.

Реалізація Технології Data Mining у Аналізі Ризиків Страхової Компанії

Михайло Лазар
Кафедра Програмної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
mykhailo.lazar@nure.ua

Володимир Кобзєв
Кафедра Програмної інженерії
Харківський національний університет радіоелектроніки
Харків, Україна
volodymyr.kobziev@nure.ua

Implementation of Data Mining Technology in the Insurance Risk Analysis

Mykhailo Lazar
Software Engineering
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
mykhailo.lazar@nure.ua

Volodymyr Kobziev
Software Engineering Department
Kharkiv National University of Radioelectronics
Kharkiv, Ukraine
volodymyr.kobziev@nure.ua

Анотація—Робота присвячена питанням використання технологій Data Mining у страховій галузі, зокрема, для аналізу страхових ризиків.

Ключові слова—страхові ризики, Big Data, технології Data Mining.

Abstract— The work is devoted to the use of Data Mining technologies in the insurance industry, in particular, for the analysis of insurance risks.

Keywords—insurance risks, Big Data, Data Mining technologies.

I. ВСТУП

Ризик став об'єктивною передумовою виникнення страхової справи. Страхіві компанії, надаючи клієнтам різні послуги страхування, намагаються знизити свої потенційні витрати, що пов'язані з виплатою сум, передбачених зобов'язаннями, при настанні страхових випадків. Для цього існує набір моделей ризиків при здійсненні певного виду страхування. Такі моделі враховують ймовірність настання страхових випадків в залежності від ряду характеристик клієнтів, які можуть вплинути на час настання страхового випадку та розмір суми відшкодування збитків. Побудова таких моделей базується на попередньому досвіді страховиків та технології роботи, що затверджена для практичної діяльності у окремих страховій компанії.

Цілеспрямовані дії з обмеження або мінімізації ризику в системі економічних відносин називають управлінням ризиком або ризик-менеджментом. Ризик-менеджмент на практиці являє собою сукупність аналітичних, організаційних, фінансових заходів, що мають комплексний послідовний характер, спрямований на зменшення або запобігання негативних наслідків настання ризиків [1]. Завдяки технологіям Big Data можна отримувати та обробляти величезні обсяги неструктурованих, постійно оновлюваних даних, що відкриває нові можливості для інтелектуального аналізу даних - Data Mining [2]. Величезні обсяги накопичених даних у галузі страхування потребують використання як окремих методів Data Mining, так і їх комбінацій [3].

II. АНАЛІЗ РИЗИКІВ СТРАХОВОЇ КОМПАНІЇ

Розглянемо на практиці процес аналізу страхових ризиків з урахуванням найпоширеніших характеристик клієнтів. Для прикладу завантажимо "Набір даних для прогнозування звернень щодо страхових виплат", який містить наступні дані про клієнтів:

1. age: вік;
2. sex: стать (жінка = 0, чоловік = 1);
3. bmi: індекс маси тіла;
4. children : кількість дітей
5. smoker (не палить = 0; палить = 1);
6. region: код житлового району;
7. charges: сума страховки;
8. insuranceclaim: страхові претензії (так = 1, ні = 0).

Практичну реалізацію програмного коду виконаємо за допомогою хмарної платформи IBM Bluemix та сервісу Apache Spark.

Завантажуємо файл:

```
df <- read.df(
  cos$url("test-donotdelete-pr-cjtek95pwoqsc", "insurance2.csv"),
  source = "org.apache.spark.sql.execution.datasources.csv.CSVFileFormat",
  header = "true")
head(df)
nrow(df)
Obtaining Spark session....
Spark session obtained.
age sex bmi children smoker region charges insuranceclaim
19 0 27.9 0 1 3 16884.924 1
18 1 33.77 1 0 2 1725.5523 1
28 1 33 3 0 2 4449.462 0
1338
```

Для початку аналізу даних використаємо метод кластеризації k-середніх Мак-Кіна [2], в якому сам користувач повинен задати шукане число кінцевих кластерів, що позначається як "k". Головною перевагою цього методу є можливість обробляти дуже великі масиви даних, оскільки немає необхідності зберігати в пам'яті комп'ютера всю матрицю відстаней цілком.

Нормалізуємо дані.

```
df.stand <- as.data.frame(scale(df))
health<-select(df.stand, "age", "sex", "bmi",
"children", "smoker", "region", "charges")
```

Для визначення оптимальної кількості кластерів використаємо метод "ліктя" (elbow method), який розглядає характер зміни розкиду W_{total} зі збільшенням числа груп k . Об'єднавши всі n спостережень в одну групу, ми маємо найбільшу внутрішньокластерну дисперсію [2], яка при

$k \rightarrow n$) буде знижуватися до 0.

```
k.max <- 15 # максимальна кількість кластерів
wss <- sapply(1:k.max, function(k)
+ {kmeans(df.stand, k, nstart=10)$tot.withinss})
```

Будуємо графік (рис. 1).

```
plot(1:k.max, wss, type="b", pch = 19, frame = FALSE,
+ xlab="кількість кластерів K",
+ ylab="Загальна внутрішньогрупова сума квадратів")
```

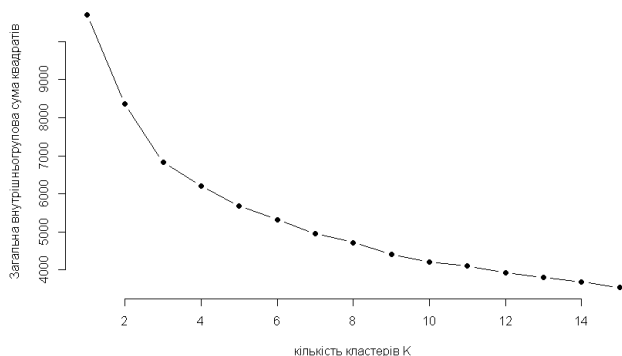


Рис. 1

У нашому випадку обгрунтовано 3 кластери

```
km <- kmeans(health, 3)
km $size
[1] 562 502 274
```

Знаходимо центри кластерів:

```
km$centers
age sex bmi children smoker region charges
1 0.80187302 -0.11007238 0.257996573 0.0969114 -
0.5072734 0.111409764 -0.1437772
2 -0.87081483 0.04131626 -0.292865189 -0.1167444 -
0.5072734 -0.122381053 -0.6854731
3 -0.04928319 0.15007267 0.007387777 0.0151149
1.9698501 -0.004295617 1.5507673
```

Оскільки наші дані стандартизовані, позитивні значення того чи іншого показника показує, що відповідна величина вище середнього значення, негативні значення вказують на те, що вона нижче середнього, а нульове значення показує, що вона збігається із середнім. Додавши номер кластеру до наших даних, ми можемо аналізувати показники в тому чи іншому кластері.

```
DF<-data.frame(df, km$cluster)
DF[1:5,c("km.cluster", "sex", "age", "smoker", "charges",
"insuranceclaim")]
km.cluster sex age smoker charges insuranceclaim
1 2 0 19 1 16884.924 1
2 3 1 18 0 1725.552 1
3 3 1 28 0 4449.462 0
4 2 1 33 0 21984.471 0
5 3 1 32 0 3866.855 1
```

Таким чином, можна поділити страховий портфель за групами ризику для розрахунку потенційних збитків, чи для розробки страхових продуктів націлених на певну групу клієнтів. Отримані таким чином дані є також джерелом інформації для ризик-менеджменту компанії.

Для передбачення кластеру, до якого потрапить потенційний клієнт, використаємо бібліотеку машинного навчання MLlib, вбудовану в Apache Spark. Ділимо дані на дві вибірки: тренувальну (train) та тестову (test):

```
df list <- randomSplit(df, c(8,2), 2)
train <- df list[[1]]
test <- df list[[2]]
kmModel <- spark.kmeans(train, insuranceclaim ~ age +sex+
bmi+children+smoker+region+charges,k = 3)
kmPrediction<-kmPredict(kmModel, newData=test)
```

Аналізуючи дані, отримані при кластеризації тестової вибірки, знаходимо, у першому кластері кількість звернень за відшкодуванням складала 50%, у другому - 97%, у третьому - 72%.

Таким чином, клієнтів в кластерах умовно можна розділити на групи ризику з подальшим використанням цієї інформації в роботі.

Для передбачення ймовірності звернення за страховим відшкодуванням спочатку застосуємо логістичну регресію, що широко використовуються для дослідження страхового ринку [3].

```
logModel <- spark.logit(train, insuranceclaim ~ age +sex+
bmi+children+smoker+region+charges, regParam = 0.5)
LogPrediction<-logPredict(logModel, newData=test)
```

Відсоток вірного передбачення звернення - 77,37%.

Для покращення результатів передбачення спробуємо використати алгоритм машинного навчання random forest (за даними [4] він є ефективним інструментом прогнозування та має стійкість до шуму).

```
rfModel<-
spark.randomForest(train,insuranceclaim~age+sex+
bmi+children+smoker+region+charges,"classification",
numTrees = 10)
rfPrediction<-rfPredict(rfModel, newData=test)
```

Відсоток вірного передбачення звернення - 90,88%.

ВИСНОВКИ

Отримані результати підтверджують можливість використання вище наведених моделей для аналізу ризиків страхової компанії. Отримані таким чином дані можна використовувати як аналітичну інформацію для цілей ризик-менеджменту, або комплексно як вхідні дані для автоматизованих систем страхування.

ЛІТЕРАТУРА REFERENCES

- [1] Сокиринська І.Г., Журавльова Т.О., Аберніхіна І.Г., "Страховий менеджмент"- Дніпропетровськ, "Пороги", 2016.
- [2] Шитиков В.К., Мастицкий С.Э. "Классификация, регрессия и другие алгоритмы Data Mining с использованием R"- Тольятти, Лондон – 2017.
- [3] "Using logit on big data – from iterative methods to analytical solutions", Birgit Stoltenberg, WORKING PAPER /// NO. 3 / 2016.
- [4] "RANDOM FORESTS", Leo Breiman Statistics Department University of California Berkeley, CA 94720 January 2001.

Особливості Технології Захисту Інформації в Сучасних Месенджерах

Збаражський Кирило Андрійович
Харківський національний
університет радіоелектроніки
факультет Комп'ютерних Наук,
Харків, Україна
kyrylo.zbarazhskiy@nure.ua

Давидова Вероніка Павлівна
Харківський національний
університет радіоелектроніки
факультет Комп'ютерних Наук,
Харків, Україна
veronika.davydova@nure.ua

Мороз Дмитро Дмитрович
Харківський національний
університет радіоелектроніки
факультету Комп'ютерних Наук,
Харків, Україна
dmytro.moroz@nure.ua

Features of Technology for Protecting Information in Modern Instant Messengers

Zbarazhskiy Kirill Andreevich
Kharkiv National University of Radio
Electronics
Faculty of Computer Science
Kharkiv, Ukraine
kyrylo.zbarazhskiy@nure.ua

Davydova Veronika Pavlovna
Kharkiv National University of Radio
Electronics
Faculty of Computer Sciences
Kharkiv, Ukraine
veronika.davydova@nure.ua

Moroz Dmytro Dmytrovich
Kharkiv National University of Radio
Electronics
Faculty of Computer Sciences
Kharkiv, Ukraine
dmytro.moroz@nure.ua

Abstract – Today, many people are using file messengers. We convey information without even thinking about its security. Compare two messengers - Telegram and Viber. We find out their security and reliability of encryption.

Keywords – Telegram, Viber, MTProto.

I. ВСТУП

Сьогодні доступ до будь-якої інформації доступний кожному. Все можна дізнатися в інтернеті. Але є інформація, до якої доступ повинен бути обмежений. Інформаційна безпека як раз і захищає нас, всіх користувачів соціальної павутини.

Інформаційна безпека - практика запобігання несанкціонованого доступу, використання, розкриття, спотворення, зміни, дослідження, запису або знищення інформації. Ніхто не хоче, щоб їх персональні дані потрапили до рук чужих людей або зловмисників. Тому кожен намагається забезпечити себе як може. Основне завдання інформаційної безпеки - збалансований захист конфіденційності, цілісності і доступності даних, з урахуванням доцільності застосування і без будь-якої шкоди продуктивності організації.

II. ПОСТАНОВКА ПРОБЛЕМИ

Порівнюємо надійність Telegram і Viber. Telegram - безкоштовний багатофункціональний месенджер, що дозволяє обмінюватися повідомленнями і медіафайлами різних форматів. Засновником є Павло Дуров. Незважаючи на те, що його месенджер вийшов на ринок пізніше своїх головних конкурентів WhatsApp і Viber, він досить швидко набув репутації одного з найбільш безпечних сервісів. Щодня сервіс поповнюють 350 тисяч нових користувачів, а сервери доставляють близько 15 мільярдів повідомлень на добу. Зробивши порівняльний аналіз Telegram та інших месенджерів ми виділили такі його переваги:

- трафік нікому не вдалося розшифрувати;
- відкритий код;
- відсутня реклама;
- сервери знаходяться в 5 різних точках світу;
- до 200 учасників в одній бесіді.

Telegram використовує, спеціально створений для додатка, протокол зв'язку MTProto. У месенджері є хмарні і секретні чати. В останніх використовується шифрування end-to-end, тобто повідомлення шифрується на телефоні відправника, а розшифровується тільки у одержувача. Telegram не бачить зміст повідомлення, тільки метадані, тобто хто, кому і коли написав, а також приблизний обсяг тексту, вони шифруються і не залишають ніяких слідів на серверах, не зберігаються в хмарі і можуть бути автоматично знищені після прочитання. Решта чатів, у тому числі приватні повідомлення, канали і групи — хмарні, вони шифруються на телефоні, передаються за допомогою протоколу MTProto на сервери месенджера, там розшифровуються і зберігаються. Тобто ключі від них є в розпорядженні Telegram. На своєму сайті розробник стверджує, що єдина реальна небезпека, пов'язана з витоком даних, - це людина, яка стоїть у вас за спиною і читає ваші повідомлення. Чи так це?

Proto — криптографічний протокол, який використовується в системі обміну повідомленнями Telegram для шифрування листування користувачів. Протокол MTProto використовує два шари шифрування - сервер-сервер і клієнт-сервер. Він працює на основі наступних алгоритмів:

- AES – симетричний 256-бітний алгоритм, прийнятий урядом США як стандарт;

- RSA – криптографічний алгоритм, в основі якого лежить обчислювальна складність завдання факторизації великих цілих чисел;
- Метод Діффі-Хеллмана - дозволяє отримати двом і більш співрозмовникам секретний ключ по незахищеному від прослуховування, однак захищеному від підміни каналу зв'язку;
- SHA-1, MD5 – хеш-алгоритми, використовуються в багатьох криптографічних протоколах і додатках для безпечного хешування.

На відміну від протоколу Double Ratchet, який застосовується WhatsApp і вже встиг отримати схвалення відомих фахівців в області захисту інформації, розробники MTPROTO не поспішають надавати свій продукт для незалежного аудиту. З одного боку, це робить алгоритм потенційно вразливим для атак, з іншого - на сьогоднішній день не зафіксовано жодної успішної дії, що призвела до розшифровки повідомлень.

Чи реальний взлом?

Дані в дата-центрах зберігаються на дисках в зашифрованому вигляді, кожен кластер зашифрований окремим ключем, який зберігається в іншому кластері під інший юрисдикцією. Тобто навіть якщо хтось руками добереться до цих дисків, їм ще доведеться мозок зламати, щоб ваше повідомлення перечитати. А повідомлення, що самоліквідуються не зберігаються ніде.

Творці месенджера заявляють про гарантії безпеки у відношенні передачі зашифрованих даних. Для підтвердження своїх слів Павло Дуров періодично організовує конкурси, в яких учасникам пропонується розшифрувати листування двох співрозмовників. Призовий фонд становить 200 тис. доларів, однак до теперішнього часу жоден хакер не зміг прочитати зашифровані повідомлення.

Навіть якщо прийняти як аксіому, що MTPROTO дійсно має кращі параметри захисту серед сучасних месенджерів, зловмисники все ж мають можливість зламати акаунт користувача. При цьому сам протокол тут абсолютно ні до чого. Уразливість можлива при авторизації користувача, або при реальній крадіжці телефону.

Розглянемо Viber. Viber — це мобільний додаток, що дозволяє дзвонити і відправляти текстові повідомлення іншим користувачам цього месенджера безкоштовно. Користуватися ним можна через Wi-Fi або мобільний інтернет. Його заснували ізраїльтяни Тальмон Марко і Ігор Магазіннік. Viber налічує близько 450 млн активних користувачів. Переваги Viber:

- можна здійснювати платні дзвінки на міські і мобільні телефони;
- величезна кількість користувачів;
- доступність на всіх платформах.

Окрім переваг Viber має також і недоліки:

- відсутність бекапа даних;

- поганий зв'язок при голосових і відео дзвінках;
- багато реклами.

Тепер розповімо про приховані можливості Viber, що стосуються безпеки. У Viber теж піклуються про безпеку своїх користувачів. Так, тут все текстові, голосові та відеочати захищені наскрізним шифруванням, тобто всі дані (всі типи повідомлень, фото, відео, голосові і відеодзвінки) шифруються в момент відправки з вашого пристрою і розшифровуються в момент прийому на пристрій одержувача. Отже, ніяка третя сторона не зможе розшифрувати і прочитати дані під час їх передачі, а тільки учасники чату. Таке шифрування здійснюється як в особистих, так і в групових чатах. Крім того, система безпеки Viber забезпечує шифрування даних, що передаються між пристроями, тобто ключі для шифрування / дешифрування повідомлень є тільки на пристроях відправника і одержувача. Тому ніхто, навіть співробітники Viber, не зможуть зламати шифрування і підслухати вашу розмову.

На відміну від Telegram, який використовує в якості протоколу end-to-end шифрування, реалізацію свого алгоритму автори Viber написали з нуля. При цьому, шифруванню піддаються:

- текстові повідомлення;
- дзвінки;
- файли.

Для організації end-to-end шифрування кожен з клієнтів використовує пару ключів: відкритий і закритий. Ця пара ключів алгоритму шифрування 256-bit Curve-25519 (ID ключа) для клієнта Viber генерується при установці месенджера на основний пристрій. Публічна частина ключа відправляється на сервер Viber, а приватна зберігається на пристрої клієнта для розшифровки повідомлень, що приходять. Інші пристрої, які працюють з цим акаунтом Viber, також отримують копію закритого ключа від основного пристрою з використанням спеціального криптографічного алгоритму. Viber встановлює захищене підключення з усіма пристроями клієнтів, коли вони захочуть обмінятися повідомленнями. Це означає, що, якщо у однієї сторони месенджер встановлений на основному пристрої, на ПК і планшеті, а в іншого також використовується на декількох пристроях, захищене підключення буде встановлено між кожним з пристроїв протилежних сторін.

Але з конфіденційністю в Viber можуть бути певні проблеми. Адже стовідсоткової гарантії про захист ваших даних ніхто не дасть. Варто відзначити, що один з допоміжних сервісів Viber Support був зламаний в липні 2013 року групою, яка назвалася Сирійською електронною армією. Вони зламали сайт Viber і отримали доступ до бази даних. На веб-ресурсі компанії вони залишили повідомлення: «На жаль, не змогли зламати всі системи Viber». За словами зловмисників, вони успішно завантажили кілька резервних копій з даними про телефонні номери та адреси електронної пошти користувачів. Висновок

Viber. Як ви розумієте, плюси є, але мінусів більше. Хтось не звертає увагу на ці недоліки, а для когось вони мають велике значення, що змушує повністю відмовлятися від використання цього месенджера, на користь інших. Саме з цієї причини Viber не є найпопулярнішим сервісом для спілкування.

Ступінь захищеності даних в Telegram, є одним з кращих. Тому даний месенджер - це найкраще рішення для захисту ваших даних і листування.

ЛІТЕРАТУРА (REFERENCES)

- [1] Вікіпедія. Telegram. – Режим доступу: <https://ru.wikipedia.org/wiki/Telegram> (дата звернення: 11.11.2018).
- [2] Рябухіна Ю. Telegram. Что это такое и почему это круто. – Режим доступу: <https://keddr.com/2014/11/telegram-что-это-такое-i-pochemu-eto-kruto/> (дата звернення: 11.11.2018).
- [3] Вікіпедія. Viber. – Режим доступу: <https://ru.wikipedia.org/wiki/Viber> (дата звернення: 11.11.2018).
- [4] Habr. Viber получил шифрование end-to-end по умолчанию. – Режим доступу: <https://habr.com/company/eset/blog/283010/> (дата звернення: 11.11.2018).

Розробка Програмного Забезпечення Вибору Транспортних Засобів для Автотранспортних Підприємств за Допомогою Веб-Технологій

Вячеслав Асаян
Автоматизації та комп'ютерно-інтегрованих технологій
Харківський національний автомобільно-дорожній університет
Харків, Україна
vulgoverbatim@gmail.com

Леонід Нефьодов
Автоматизації та комп'ютерно-інтегрованих технологій
Харківський національний автомобільно-дорожній університет
Харків, Україна
nefedovli@i.ua

Дмитро Маркозов
Автоматизації та комп'ютерно-інтегрованих технологій
Харківський національний автомобільно-дорожній університет
Харків, Україна
markozov.d@gmail.com

Development of Software for Selection of Vehicles for Road Transport Enterprises Using Web Technologies

Viacheslav Asaian
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkov, Ukraine
vulgoverbatim@gmail.com

Leonid Nefodov
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkov, Ukraine
nefedovli@i.ua

Dmytro Markozov
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkov, Ukraine
markozov.d@gmail.com

Анотація – Проаналізовано вплив характеристик автомобіля на його ефективність експлуатації. Розроблено web-додаток для автоматичного розрахунку експлуатаційних властивостей автомобіля. (

Abstrac – The influence of car parameters (characteristics) on its efficiency of operation is analyzed. A web application has been developed for the automatic calculation of vehicle operation properties.

Ключові слова– експлуатаційні властивості, дорожньо-транспортні пригоди, графік прискорень, web-додаток

Key words– operation properties, traffic accidents, acceleration schedule, web-application.

I. ВСТУП

Сьогодні автотранспортні засоби мають великий вплив на побут кожного з нас. Для вирішення щоденних завдань його використовує сільське господарство, перевезення пасажирів і інші області діяльності людини. Автомобільна індустрія стрімко розвивається, але не можна забувати про безпеку. Автотранспортне підприємство повинно правильно визначити склад транспортних засобів, які б відповідали конкретним вимогам і забезпечували ефективність експлуатації.

Дослідження параметрів руху автомобіля з урахуванням всіх зв'язків між його окремими елементами являє собою задачу великої складності. Тому при дослідженні будь-яких експлуатаційних властивостей, в тому числі і його керованості і стійкості, автомобіль замінюється розрахунковою

моделлю, яка в більшій чи меншій мірі відображає реальний автомобіль. Складність розрахункової моделі і ступінь її наближення до реального автомобілю диктується рядом розмірковувань. При виборі розрахункової моделі перш за все слід враховувати цілі завдання, яке вирішується.

II. АНАЛІЗ І ПОСТАНОВКА ПРОБЛЕМИ

Під час експлуатації автомобіля потрібно постійно відслідковувати багато його властивостей, оскільки це складний механізм. Теорія автомобіля розглядає найважливіший аспект – визначення готовності транспортного засобу до експлуатації [1]. Експлуатаційні характеристики ділять на такі властивості:

- тягово-швидкісні;
- економічність витрати палива;
- плавність пересування;
- розгін-гальмові та ін.

Е. А. Чудаков в 1935-му році вперше сформулював основні поняття науки під назвою «теорія автомобіля» [2], видавши однойменну книгу. Теорія автомобіля поступово розширювалася, розвивалися закладені в неї ідеї. В подальшому виходили праці, присвячені методам: розрахунку плавності ходу (Р. В. Ротенберг [3]), керованості і стійкості (А.С. Литвинов [4]), оцінки прохідності (Я. Е. Фаробін) і розрахунку тягово-швидкісних властивостей автомобіля.

Щоб самостійно розрахувати експлуатаційні властивості, потрібно витратити багато часу і звертатися до додаткових літературних джерел. Метою дослідження

є розробка програмного забезпечення вибору транспортних засобів для автотранспортних підприємств за допомогою веб-технологій.

III. РОЗРОБКА WEB-ДОДАТКА

Форма, з якою взаємодіє користувач (рис. 1), має кнопку "Зробити розрахунки", при натисканні на неї відбувається автоматичне обчислення по заданим параметрам і побудова графіків. Кнопка "File" відкриває стандартний браузер файлів локального пристрою, що дозволяє завантажити збережені раніше дані для повторного використання.

Нижче розташований список моделей автомобіля, що випадає з переліком заздалегідь занесених розробником в БД автомобілів. Після вибору потрібної моделі майже всі поля, які відносяться до характеристик автотранспорту, заповнюються автоматично, потім залишається вибрати зовнішні фактори опору руху.

Перерахування полів даних:

- вікно «Дані автомобіля», що розгортається при натисканні на неї (рис. 2). У цьому вікні присутні такі поля введення: «Повна маса автомобіля, кг», «Максимальна потужність, кВт», «Частота обертання валу двигуна при макс. потуж., об/хв.», «Передавальне число Головної передачі», «Передавальне число I-ої передачі», «Передавальне число II-ої передачі», «Передавальне число III-ої передачі», «Передавальне число IV-ої передачі», «Передавальне число V-ої передачі», «Висота, м», «Ширина, м» і перемикач «Тип двигуна»;
- діалоговий список «ККД трансмісії автомобіля»;
- діалоговий список «Тип кузова автомобіля»;
- діалоговий список «Тип і стан дорожнього покриття»;
- діалоговий список «Категорія ділянки»;
- бігунок «Ухил дороги у %»;
- діалоговий список «Розмір шин»;
- блок з перерахуванням деяких підказок для користувача.

Всі значення в наданих полях введення можна змінювати вручну у рамках вказаного формату.

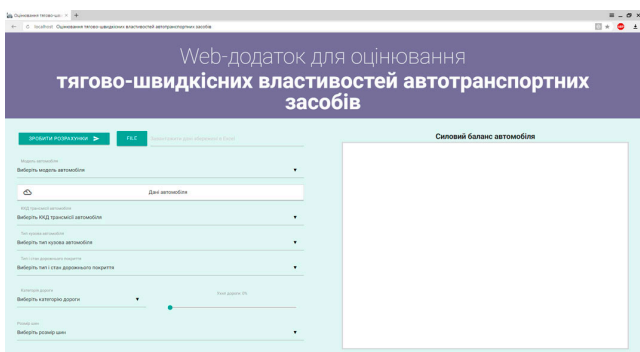


Рис. 1 – Веб-сторінка в нормальному вигляді, що завантажена через браузер користувачем

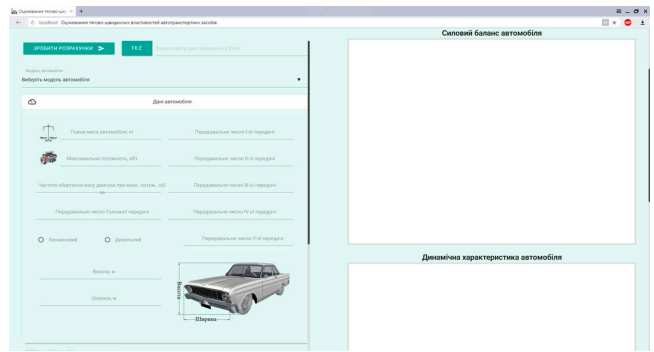


Рис. 2 – Вікно «Дані автомобіля»

Додаток справно відображається на ноутбучі (1366x768 пікселів), планшеті і смартфоні (рис. 3).

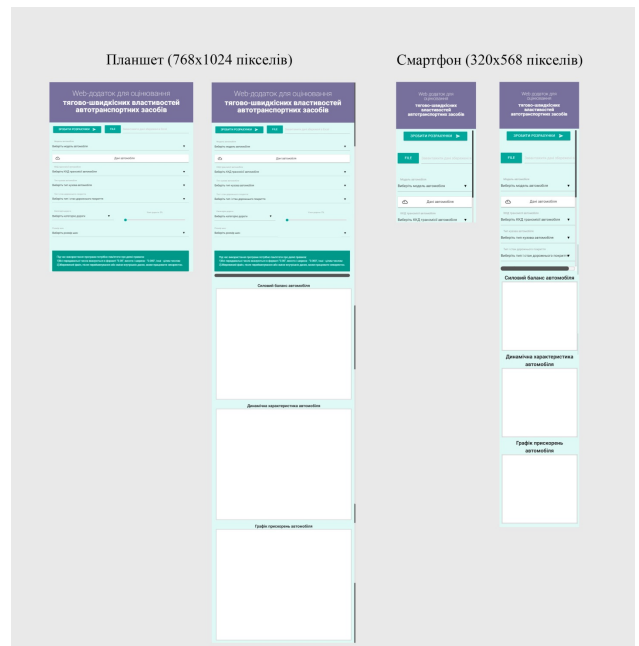


Рис. 3 – Веб-додаток, що відкрито на екрані планшета і смартфона

Після правильного заповнення форми і натискання на кнопку "Зробити розрахунки", з'являється діалогове вікно (рис. 4), яке дозволяє зберегти початкові і отримані в кінці обчислень дані у вигляді звіту (рис. 5) формату «.xls».

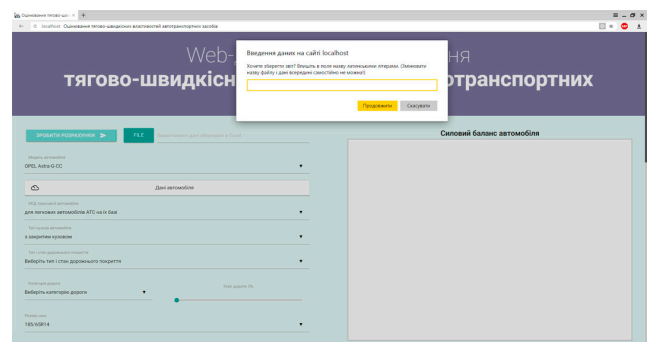


Рис. 4 – Діалогове вікно збереження файлу

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1															
2															
3															
4															
5															
6															
7															
8															
9															
10															
11															
12															
13															
14															
15															
16															
17															
18															
19															
20															
21															
22															
23															
24															
25															
26															
27															
28															
29															
30															

Рис. 5 – Внутрішня структура Excel-звіту

У результаті справа екрану по ним будується три графіка з підзаголовками: "Силовий баланс автомобіля", "Динамічна характеристика автомобіля" і "Графік прискорень автомобіля" (рис. 6).

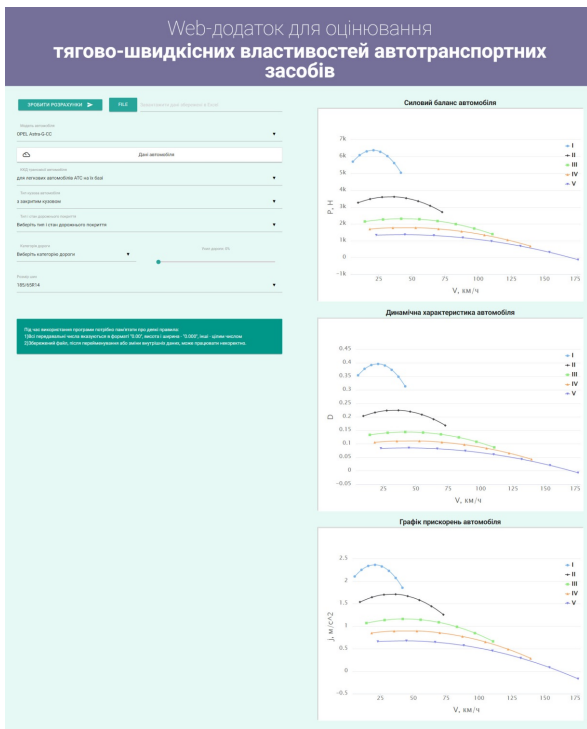


Рис. 6 – Кінцевий результат з побудованими графіками

Висновки

Розроблено програмне забезпечення вибору транспортних засобів для автотранспортних підприємств за допомогою веб-технологій, що відрізняється від існуючих аналогів готовою базою автомобілів і доступністю для користувача.

Користувач враховує розраховані експлуатаційні властивості конкретного транспортного засобу при проектуванні, щоб домогтися кращої продуктивності.

REFERENCES ЛІТЕРАТУРА

- [1] Подригало М.А. Стойкость колесных машин как сложное эксплуатационное свойство / М.А. Подригало, Д.М. Клец, Н.П. Артемов // Автомобильный транспорт. – 2011 г. – 21 июля. – Вып. 29. – С. 179-183.
- [2] Чудаков Е.А. Теория автомобиля / Е.А. Чудаков. – Государственное научно-техническое издательство машиностроительной литературы, 1950. – 344 с.
- [3] Ротенберг Р. В. Подвеска автомобиля. Колебания и плавность хода / Р. В. Ротенберг. – изд. 3-е перераб. и доп. – Москва: Машиностроение, 1972. – 392 с.
- [4] Литвинов А.С. Устойчивость и управляемость автомобиля / А.С. Литвинов. – М.: Машиностроение, 1971. – 416 с.

СЕКЦІЯ 3

Системи управління

Management Systems

Параметрування Швидкісних Лічильників PLC Simatic S7-1200 для Систем Управління з Імпульсними Перетворювачами

Николайчук М.Я.

Кафедра інформаційно-телекомунікаційних технологій та систем

Івано-Франківський національний технічний університет нафти і газу

Івано-Франківськ, Україна

nmj_2010@ukr.net

Гедзик Д.Я.

Кафедра інформаційно-телекомунікаційних технологій та систем

Івано-Франківський національний технічний університет нафти і газу

Івано-Франківськ, Україна

dmytro.hedzyk16@gmail.com

Parameterization of high speed counters PLC Simatic S7-1200 for control systems with pulsed transducers

Nykolajchuk M.Y.

Department of Information and Telecommunication Technologies and Systems

Ivano-Frankivsk National Technical University of Oil and Gas

Ivano-Frankivsk, Ukraine

nmj_2010@ukr.net

Hedzyk D.Y.

Department of Information and Telecommunication Technologies and Systems

Ivano-Frankivsk National Technical University of Oil and Gas

Ivano-Frankivsk, Ukraine

dmytro.hedzyk16@gmail.com

Анотація—розроблено апаратно-програмні засоби і методику параметрування швидкісних лічильників програмованих логічних контролерів (ПЛК) для систем управління з імпульсними витратомірами.

Abstract—hardware and software tools and method of parameterization of high speed counters of programmable logic controllers (PLC) for control systems with pulsed flowmeters have been developed.

Ключові слова—системи управління, програмований логічний контролер (ПЛК), параметрування швидкісних лічильників, імпульсні витратоміри.

Keywords—control systems, programmable logic controller (PLC), parameterization of high speed counters, pulsed flowmeters.

I. ВСТУП

Об'єктом дослідження є інформаційні процеси в системах управління, пов'язані з методами збору, обробки і передачі даних в системах управління з витратомірними імпульсними перетворювачами.

Актуальність роботи полягає в необхідності побудови розподілених систем управління з витратомірними перетворювачами на основі новітніх апаратно-програмних засобів і технологій з розширеними функціональними можливостями.

Метою роботи є створення апаратно-програмного комплексу для дослідження інформаційних процесів при параметруванні швидкісних лічильників PLC та розширення функціональних можливостей

розподілених систем управління за рахунок організації локального WEB-сервера на PLC і відкритої "Open User Communication" комунікації на основі протоколів TCP/IP [1].

II. ОБГРУНТУВАННЯ ЗАСТОСУВАННЯ ПЛК ДЛЯ СИСТЕМ УПРАВЛІННЯ

Побудова сучасних систем управління передбачає їх багаторівневу структуру (рівень первинних перетворювачів і виконавчих механізмів, рівень PLC, комунікаційний рівень, верхній рівень на базі серверних робочих станцій і SCADA-систем).

Для організації і апаратно-програмного забезпечення вказаної структури, на основі світового досвіду і сучасних тенденцій розвитку інформаційних технологій, переважним і оптимальним є застосування PLC на відповідному структурному рівні.

Особливостями застосування PLC в системах управління, порівняно з мікроконтролерними є:

- модульна конструкція (забезпечує можливість побудови різномірних конфігурацій для роботи з різними типами сигналів, функцій і комунікаційних задач);

- спрощення проектних процедур за рахунок підтримки уніфікованого вертикально-інтегрованого програмного забезпечення (Simatic TIA Portal) [2];

- стандартизовані мови програмування IEC 61131-3;

- підтримка промислових комунікаційних шин (Industrial Ethernet, ProfiBus-DP, AS-interface, ModBus RTU, ModBus TCP та інш.);

- вбудовані WEB-сервери з підтримкою HTML-5 і JavaScript [3].

III. ПАРАМЕТРУВАННЯ ШВИДКІСНИХ ЛІЧИЛЬНИКІВ SIMATIC S7-1200

PLC Simatic S7-1200 підтримують шість швидкісних лічильників з вхідною частотою до 100 кГц (CPU 1217 – до 1 МГц). Особливістю швидкісних лічильників Simatic S7-1200 є їх апаратна реалізація в складі PLC з програмним керуванням, що забезпечує можливість їх функціонування поза програмним циклом PLC, що складає одиниці мілісекунд [4].

Параметрування швидкісних лічильників PLC S7 1200 має розширену функціональність і виконується за трьома типами:

- прямиий і реверсивний рахунок в діапазоні (від -2147483648 до +2147483647);
- вимірювання частоти (для періоду 0.01, 0.1 і 1.0 с);
- вимірювання періоду (від 0 до 4294967280 нс).

На рис. 1 наведено конфігурацію апаратних засобів на базі PLC S7-1200, що включає два PLC (Master CPU 1214C і PLC_2 CPU 1212C), серверну робочу станцію SIMATIC PC Station зі SCADA-системою та промислову мережу PN/IE_1.

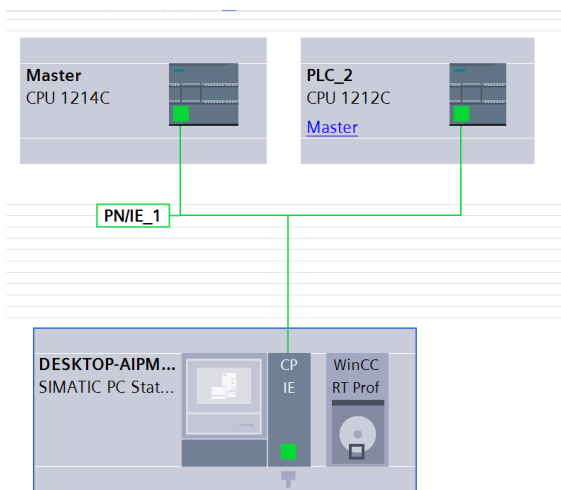


Рис. 1. Конфігурація апаратних засобів на базі PLC S7-1200

Для керування швидкісними лічильниками PLC виконано параметрування програмного блоку на мові FBD (Functional Block Diagram) (рис. 2). Параметри включають апаратний ідентифікатор «Local~HSC_4», ідентифікатор екземплярного (instance) блоку даних «Data_block_4_HSC_4_EXT», та діагностичні виходи (DONE, BUSY, ERROR і STATUS) [5].

Крім того, для параметрування швидкісних лічильників для систем управління з імпульсними перетворювачами і розширення їх функціональних можливостей, реалізовано функції переривання основного програмного циклу PLC за досягненням заданого значення, старт-стопні операції, функції внутрішньої та зовнішньої синхронізації.

Таким чином, реалізовано та апробовано методи організації та параметрування компонентів систем управління з імпульсними перетворювачами на основі швидкісних лічильників PLC, що працюють поза

програмним циклом PLC і мають розширену функціональність порівняно зі стандартними цифровими входами.

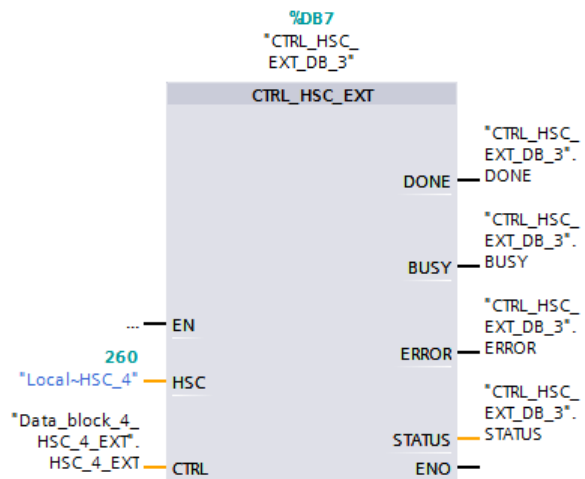


Рис. 2 Програмний блок для параметрування швидкісного лічильника на мові FBD (Functional Block Diagram)

ВИСНОВКИ

В роботі виконано огляд технологій збору, обробки і передачі даних з імпульсних витратомірних перетворювачів.

Обґрунтовано застосування PLC для вирішення задач побудови систем управління з імпульсними витратомірними перетворювачами.

Розроблено апаратно-програмні засоби для реалізації алгоритмів збору даних з витратомірних перетворювачів, організації локального WEB-сервера та Open User Communication на базі PLC S7-1200.

Результати роботи апробовані в лабораторних та промислових умовах і можуть бути корисними для вирішення аналогічних задач і досліджень інформаційних процесів в системах управління.

- [1] Open User Communication with TSEND_C and TRCV_C. Simatic S7-1200 CPU (V1.0, Item ID: 67196808). SIEMENS: – FAQ Jenuary – 2013 – 22 p.
- [2] Siemens. STEP 7 and WinCC Engineering V15.1. System Manual 10/2018. - P. 79-189.
- [3] Siemens. Creating User-defined Web Pages on S7-1200 / S7-1500. (Entry ID: 68011496, V2.2), 04/2017. – 34 p.
- [4] SIMATIC S7. S7-1200 Programmable Controller. System Manual: Siemens (A5E02486680-AK). – 09/2016. – P. 555-588.
- [5] Siemens. Application examples for High-Speed Counters (HSC) Application Example IAPortal, S7-1200 V4.2 Entry ID: 109742346, V1.0, 11/2016. – 24 p.

Система Антипомпажного Захисту Газоперекачувального Агрегату на Базі Апаратно-Програмних Засобів Siemens

Татарин Т.М., Іванюк Н.І.

Кафедра інформаційно-телекомунікаційних технологій і систем
Івано-Франківський національний технічний університет
нафти і газу

Заміховський Л.М.

Кафедра інформаційно-телекомунікаційних технологій і систем
Івано-Франківський національний технічний університет
нафти і газу

Anti-Surge Protection System for Gas-Pumping Unit on the Basis of Siemens Hardware and Software

Tataryn T.M., Ivanuyk N.I.

Department of Information and Telecommunication
Technologies and Systems

Ivano-Frankivsk National Technical University of Oil and Gas
Ivano-Frankivsk, Ukraine
taras_tataryn@ukr.net

Zamikhovskiy L.M.

Department of Information and Telecommunication
Technologies and Systems

Ivano-Frankivsk National Technical University of Oil and Gas
Ivano-Frankivsk, Ukraine
leozam@ukr.net

Abstract - Запропонована система антипомпажного захисту газоперекачувального агрегату (ГПА), визначені контролюємі параметри, обґрунтований алгоритм прогнозування помпажу та його програмна реалізація. The system of anti-surge protection of gas pumping unit (GPA) is given. The controlled parameters are defined. Surge prediction algorithm and its software are substantiated.

Keywords-помпаж, технологічні параметри, газоперекачувальний агрегат, система, прогнозування, алгоритм, програмне забезпечення, антипомпажний захист. *Surge, technological parameters, gas-pumping unit, system, prediction, algorithm, software, anti-surge protection.*

I. ВСТУП

Об'єктом дослідження є процес зриву потоку газу з лопаток колеса нагнітача ГПА з можливим проривом його з нагнітального патрубку у всмоктуючий, що супроводжується значною вібрацією, шумом і різкою зміною робочих параметрів ГПА та може призвести до виникнення аварійних ситуацій. Кожен помпажний удар зменшує моторесурс ГПА мінімум на 50 мотогодин [1].

В газотранспортній системі України працює значна кількість ГПА різного типу, найбільш чисельними з яких є ГПА з газотурбінним приводом. Режим роботи нагнітача ГПА постійно змінюється під час експлуатації, що пов'язано із зміною споживання газу, його складу, умов навколишнього середовища, а також зміною технічного стану ГПА в процесі експлуатації. Таким чином, забезпечення надійності роботи ГПА, зокрема його захисту від помпажу, є актуальною проблемою.

Метою роботи є розробка системи антипомпажного захисту ГПА на базі сучасних апаратно-програмних засобів фірми Siemens з використанням, як одного із

інформативних параметрів, вібраційних процесів, що супроводжують роботу ГПА, алгоритму прогнозування помпажу та його програмної реалізації.

II. СТАН ПРОБЛЕМИ АНТИПОМПАЖНОГО ЗАХИСТУ ГПА

Аналіз літературних джерел показує, що методи ідентифікації помпажу розвиваються в двох напрямках [2,3,4 і ін]:

- шляхом аналізу поточного стану нагнітача та порівняння його стану з наперед відомим, для чого використовують розподіл області допустимих режимів роботи ГПА на області стійкої та нестійкої роботи нагнітача. Перевагою методу є можливість визначення робочої точки за параметрами нагнітача, що підлягають вимірюванню, а недоліком - висока чутливість методу до похибки вимірювань та зміні характеристик нагнітача від часу його напрацювання;

- шляхом аналізу параметрів коливань (амплітуди, частоти) технологічних параметрів, специфічних для автоколивань газодинамічної системи, який дозволить зробити висновок про виникнення помпажу. Так, пульсації витрати та тиску призводять до виникнення пульсації потужності, обертового моменту, осьових зусиль у підшипниках, температури за турбіною низького тиску та інших параметрів нагнітача і приводу, що зв'язані функціональними залежностями з витратою. Рішення про те, який з параметрів вибрати як інформаційний, залежить від вимог його надійного контролю. Перевага методу полягає в тому, що для ідентифікації помпажу не використовуються витратно-напірні характеристики нагнітача, отже, на результати аналізу не впливає дрейф фактичних характеристик ГПА.

III. ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕНЬ

Виходячи з викладеного, в роботі необхідно обґрунтувати вибір інформаційних технологічних параметрів ГПА, розробити алгоритм прогнозування параметрів їх коливань (амплітуди, частоти) для попередження виникнення помпажу, а також його програмну і технічну реалізацію.

IV ОТРИМАНІ РЕЗУЛЬТАТИ

З метою вибору інформативних технологічних параметрів роботи ГПА, стосовно ГПА-Ц-16С, був проведений аналіз зміни його параметрів перед виникненням помпажних коливань у нагнітачі. На рис. 1 зображене вікно реєстрування параметрів процесу регулювання ГПА-Ц-16С ст. №3 при помпажній ситуації на КС-3 Долинського ЛВУМГ УМГ «Прикарпаттрансгаз».

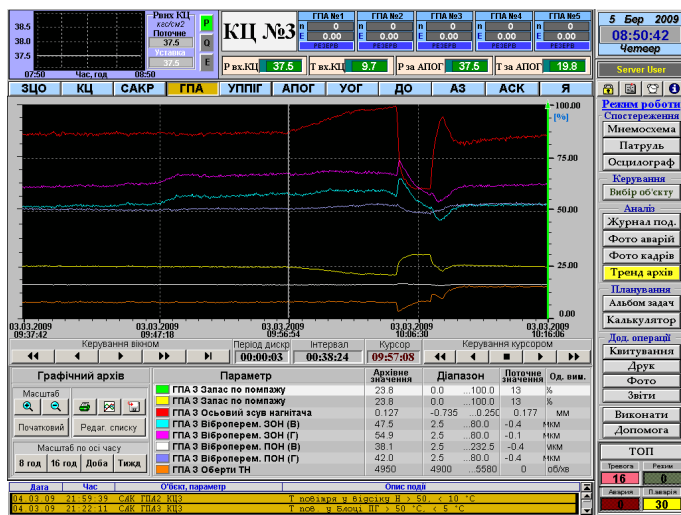


Рисунок 1 – Тренд роботи ГПА-Ц-16С при помпажній ситуації на КС-3

Аналізі параметрів роботи ГПА показав, що перед появою помпажних коливань в нагнітачі збільшується рівень параметра вібрації (вібропереміщення) на валу його задньої опори та осьовий зсув нагнітача ГПА-Ц-16С, а інші показники залишаються в нормі. При цьому також варто відзначити, що незмінними залишаються показники вібрації передньої опори нагнітача.

Розглянемо детальніше перелік технологічних параметрів ГПА-Ц-16С, які можна використати як інформативні для прогнозування виникнення помпажу, а також їх діапазон вимірювання:

- перепад тиску на конфузорі – діапазон вимірювання від 0 до 1 кгс/см², робоча точка «оптимальна» - 0.2 кгс/см², істотна зміна в межах 10-15%;

- оберти турбіни нагнітача – діапазон вимірювання від 0 до 5200 об/хв, робоча точка «оптимальна» - 5000 об/хв, істотна зміна в межах 10 - 15%;

- осьовий зсув нагнітача – діапазон вимірювання від -500 мкм до +500 мкм, загалом 1 мм, робоча точка «ідеальна» - 0 мм, істотна зміна в межах 10%;

- вібропереміщення задньої опори нагнітача (по горизонталі та вертикалі) – діапазон вимірювання від 0 до 80 мкм, робоча точка «ідеальна» - 40 мкм, істотна зміна в межах 10%.

В даному випадку поняття «ідеальної» та «оптимальної» робочої точки означають наступне: «ідеальна» - параметри які можна отримати при роботі агрегату за відсутності дефектів, пошкоджень та ін.; «оптимальна» - параметри при яких отримується максимальна продуктивність роботи агрегату.

Для прогнозування зміни розглянутих вище параметрів був використаний інтерполяційний многочлен Лагранжа першого порядку, оскільки поліноми Лагранжа зручно застосовувати для інтерполяції, а також чисельного інтегрування [5]. Екстраполяційна модель, що використовує поліноми Лагранжа, має вигляд:

$$F(t) = \sum_{i=0}^n \xi_i \varphi_i(t) = \xi_0 \varphi_0(t) + \xi_1 \varphi_1(t) + \dots + \xi_n \varphi_n(t),$$

$$\varphi_i(t) = \frac{(t-t_0)(t-t_1) \dots (t-t_{i-1})(t-t_{i+1}) \dots (t-t_n)}{(t_i-t_0)(t_i-t_1) \dots (t_i-t_{i-1})(t_i-t_{i+1}) \dots (t_i-t_n)}$$

$$\equiv \prod_{j=0, j \neq i}^n \frac{(t-t_j)}{(t_i-t_j)}$$

$i = 0, n, \xi_i$ – значення $\zeta(t)$, в момент часу $t_i \in T_1$.

В розгорнутому вигляді екстраполяційну модель записують наступним чином:

$$F(t) = \frac{(t-t_1)(t-t_2) \dots (t-t_n)}{(t_0-t_1)(t_0-t_2) \dots (t_0-t_n)} \xi(t_0) +$$

$$+ \frac{(t-t_0)(t-t_2) \dots (t-t_n)}{(t_1-t_0)(t_1-t_2) \dots (t_1-t_n)} \xi(t_1) +$$

$$+ \frac{(t-t_0)(t-t_1) \dots (t-t_{n-1})}{(t_n-t_0)(t_n-t_1) \dots (t_n-t_{n-1})} \xi(t_n).$$

Спосіб обчислення многочленів, що був запропонований Лагранжом, є зручним для реалізації за допомогою мов програмування високого рівня, зокрема для мови програмування SCL (Structured Control Language) – високорівнева, Pascal-орієнтована мова програмування, яка поряд із LAD та FBD є однією із мов, що призначена для створення програм для ПЛК фірми Siemens. В роботі розглядається процедура прогнозування зміни вибраних параметрів роботи ГПА з використанням даної моделі.

Була розроблена система антипомпажного захисту ГПА, яка реалізує розглянуту модель. Апаратне забезпечення системи складається з контролера Siemens S7-1200 з CPU 1212C AC/DC/Rly, аналогового модуля вводу для підключення датчиків технологічних параметрів ГПА з уніфікованим вихідним сигналом 4-20 мА, а також інноваційного вібраційного модуля SM 1281 від фірми Siemens (рис. 2).

Для розробки програмного забезпечення системи управління на основі технологічного модуля SM 1281

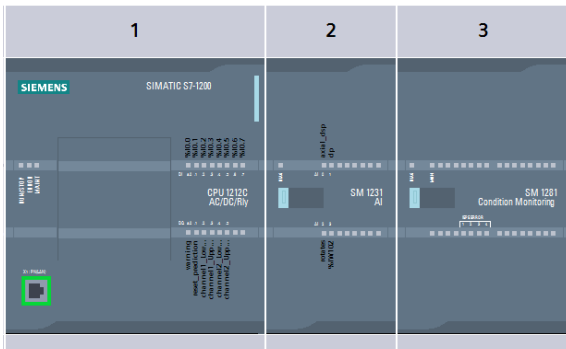


Рисунок 2 – Апаратне забезпечення системи антипомпажного захисту ГПА

використовується програмний пакет TIA Portal - інтегроване середовище розробки програмного забезпечення систем автоматизації технологічних процесів. Структура проекту в середовищі TIA Portal V14 виглядає наступним чином (рис. 3).

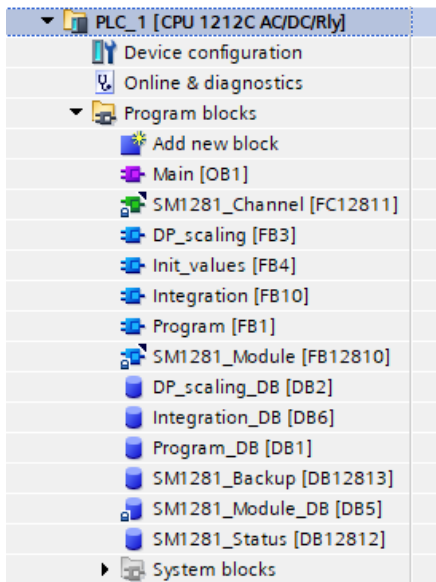


Рисунок 3 – Структура проекту

Для програмування контролера Siemens S7-1200 використовується програмний пакет Step 7, а у випадку реалізації верхнього рівня системи - для візуалізації технологічного процесу та забезпечення роботи системи в режимі реального часу, а також архівування інформації про стан ГПА може бути використана SCADA-система Simatic WinCC.

Першим етапом прогнозування є накопичення інформація про зміну значення параметрів ГПА. В розробленій програмі це досягається за допомогою запису вимірних значень параметрів у відповідний масив даних, в якому зберігаються початкові значення для можливості їх подальшого прогнозування. Причому, для більш коректного формування екстраполяційної моделі здійснюється вимірювання кількох значень для одного технологічного параметра ГПА в дискретні моменти часу, та знаходиться їх середнє значення.

Після того, як були обраховані початкові значення технологічних параметрів для прогнозування, за допомогою інтерполяційного поліному Лагранжа можна побудувати екстраполяційну модель для визначення значень вибраних технологічних параметрів роботи ГПА на майбутній момент часу. В результаті прогнозування отримуємо прогнозовані значення параметрів на певний, наперед заданий, проміжок часу i , після досягнення цього моменту часу, прогнозоване значення порівнюється із фактичним. Якщо отримане значення абсолютної похибки знаходиться в допустимих межах (зазвичай 10-15%), робиться висновок щодо коректності побудованої екстраполяційної моделі і здійснюється наступний прогноз на подальший проміжок часу. За результатами наступних ітерацій робиться висновок щодо можливості виникнення явища помпажу ГПА. Таким чином, використовуючи результати прогнозування зміни інформаційних технологічних параметрів роботи відцентрового нагнітача в часі, можна попередити виникнення явища помпажу.

ВИСНОВКИ

Проведений аналіз технологічних параметрів роботи ГПА дозволив обґрунтувати вибір додаткових інформаційних параметрів - осьового зсуву нагнітача та параметрів вібрації його задньої опори в вертикальному і горизонтальному напрямках до осі нагнітача для визначення моменту виникнення явища помпажу.

Розроблена процедура прогнозування виникнення явища помпажу з використанням екстраполяційної моделі на базі поліномів Лагранжа.

Створена система антипомпажного захисту ГПА на базі сучасних апаратно-програмних засобів фірми Siemens та розроблено її програмне забезпечення з використанням програмного пакету TIA Portal V14.

REFERENCES ЛІТЕРАТУРА

- [1] Ніщета В.В., Свістельник О.В., Мамонов А.І., Таргонський В.О. Помпажні явища та їх вплив на моторесурс ГПА // Інформаційний огляд ДК «Укртрансгаз». – 2006. – № 3(39). – С. 6–7. – № 4(40). – С. 11–13.
- [2] Цабенко М. В. Системи антипомпажного керування відцентровим компресором / М. В. Цабенко // Електротехніка та електроенергетика. - 2016. - № 1. - С. 35-41.
- [3] Антипомпажное регулирование компрессоров в Дельта V // Emerson Process Management. – апрель 2002.–9с.; на сайте:www.EmersonProcess.ru.
- [4] Противопомпажная (антипомпажная) защита ТВЛ, ТК, ТН и ЭК.–18.07.07. –2с.; на сайте <http://teraurs.kharkov.ua/application-pompag.htm>.
- [5] Заміховський Л. М. Основи теорії надійності і технічної діагностики систем: Навчальний посібник / Л. М. Заміховський, В. П. Калявін, - Івано-Франківськ: Вид-во «Полум'я», 2004. – 360 с.

Розроблення Автономного Пристрою Моніторингу Температури на Базі Мк Esp8266 Та Середовища Lua У Межах Концепції "Інтернету Речей" (Iot)

Юрій Паньків

*Івано-Франківський національний технічний
університет нафти і газу*

*Кафедра інформаційно-телекомунікаційних технологій
та систем*

Івано-Франківськ, Україна
yuriy.pankiv@gmail.com

Олександр Танюк

*Івано-Франківський національний технічний
університет нафти і газу*

*Кафедра інформаційно-телекомунікаційних технологій
та систем*

Івано-Франківськ, Україна
taniuk.oleks@ukr.net

I ВСТУП

Анотація—Розроблено автономний пристрій на базі мікроконтролера ESP-8266 для моніторингу температури. Передбачено можливість дистанційного керування температурним режимом шляхом передавання керуючих команд через інфрачервоний протокол по мережі Internet.

Ключові слова — інтернет речей, бездротові технології передавання даних, вимірювальні пристрої.

Сьогодні значної популярності набуває концепція Інтернету речей (IoT). А із появою дешевих, достатньо продуктивних пристроїв, з малими розмірами і з вбудованим безпроводним мережевим інтерфейсом, з'являється можливість впровадження концепції IoT в усі пристрої які дотепер не мали такої можливості.

Такі пристрої, зокрема, можуть збирати інформацію про навколишнє середовище, та ефективно управляти системою обігріву та кондиціонування, і стануть добрим стимулом до ефективного впровадження енергоефективних технологій.

Впровадження енергетичного моніторингу є інструментом скорочення споживання енергії та підвищення ефективності її використання, дає змогу отримати детальну картину споживання ресурсів.

II ПОСТАНОВКА ЗАВДАННЯ

На сьогоднішній момент найбільші виробники побутової техніки, систем опалення та кондиціонування мають свої рішення для концепції Інтернету речей але інтерфейс взаємодії цих пристроїв вимагає уніфікації, що тягне за собою вибору одного виробника, для всіх пристроїв. Як наслідок весь Інтернет речей орієнтований на просте управління цими пристроями користувачем через WEB, та й ціна цих приладів надто висока.

Для системи опалення існує проблема моніторингу температури в місцях кінцевого тепло

обміну, яку дуже важко вирішити без допомоги IoT так як кінцевий теплообмін може розміщуватися на досить віддаленій відстані від цих систем, а також кількість кінцевих теплообмінників може бути досить великою.

Згідно з вище сказаним планувалось розробити пристрій який автономно моніторить температуру в місцях кінцевого теплообміну, отримувати вказівки від користувача за допомогою WEB інтерфейсу, передавати дані безпроводним способом через Інтернет до пристрою управління, та реалізувати інтерфейс управління цим пристроєм системою обігріву та кондиціонування.

III ПРОПОНОВАНИЙ ПІДХІД

Для вирішення задачі було розглянуто різні пристрої, та з необхідними параметрами і з доступною ціною виявився пристрій ESP12E від компанії Ai-thinker [1]. Пристрій має WiFi модуль, та продуктивність його достатня для запуску скриптового інтерпретатора, та запуску WEB сервера, для отримання налаштувань від користувача.

Можливості стандартної прошивки вибраного WiFi модуля не давали змоги використовувати модуль як повноцінний автономний пристрій, тому було вирішено замінити її на прошивку від NodeMcu [2] з інтерпретатором мови програмування LUA, що дозволило полегшити процес програмування, а також перепрограмувати пристрій не прошиваючи його.

Було використано давачі температури DS18B20 в силу своєї доступності, а також підтримкою цифрового інтерфейсу 1-Wire, що забезпечило зменшення кількості окремих компонентів та виводів головного модуля [4].

Проаналізувавши інтерфейси за допомогою яких можна здійснювати управління системою обігріву та кондиціонування було вибрано інфрачервоний (ІЧ) протокол [5], використано сторонню бібліотеку

IRremoteESP8266 [6]. Пристрої об'єднуються в мережу, як показано на рис. 1.

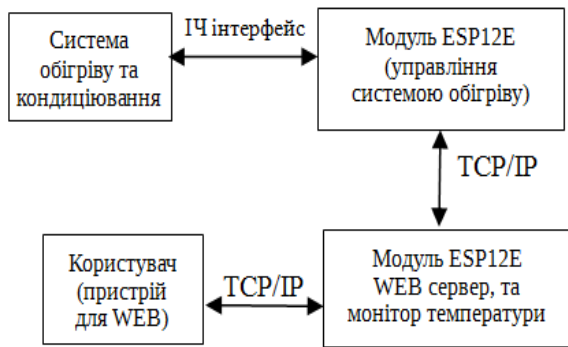


Рис. 1 - Схема взаємодії пристроїв

IV ОГЛЯД МАКЕТНОГО ПРИСТРОЮ

Схему електричну структурну наведено на рис. 2.

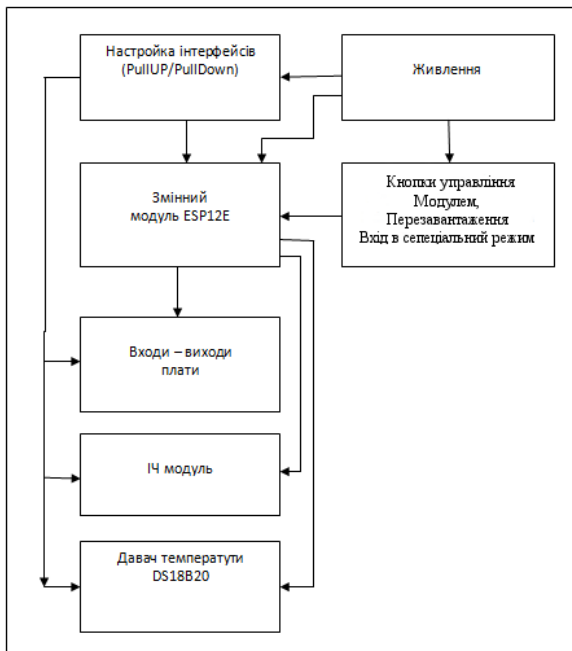


Рис. 2 - Структурна схема проектованого пристрою

Принципова схема знімного модулю зображена на рис 3.

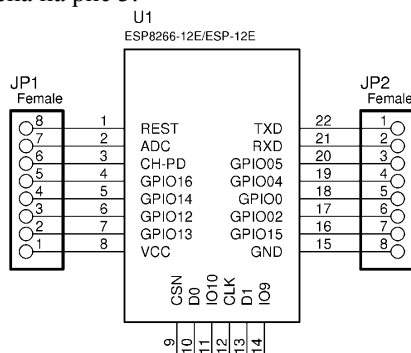


Рис. 3 - Знімний модуль, схема електрична принципова

За допомогою такої реалізації схеми було зменшено площу основної плати шляхом

багатшаровості. Принципова схема основної плати зображена на рис. 4.

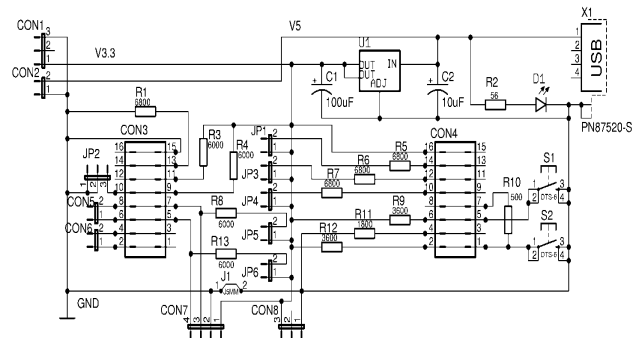


Рис. 4 - Основний модуль, схема електрична принципова

Вигляд реалізованих змінного та головного модулів показано на рис. 5

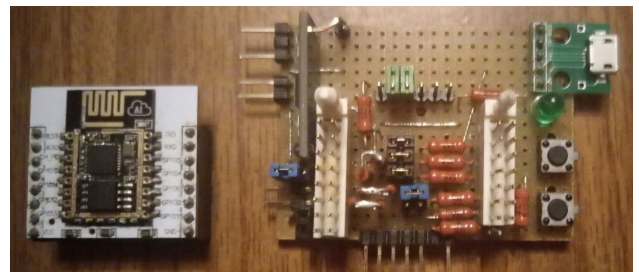


Рис. 5 - Вигляд реалізованих змінного та головного модулів

ВИСНОВОК

Розроблений пристрій для автономного моніторингу температурних показників та управління системою обігріву та кондиціонування через послідовний I2C інтерфейс. Подальші дослідження будуть пов'язані із удосконаленням методів статистичного аналізу та удосконаленням програмного пакету для вирішення поставлених задач.

ЛІТЕРАТУРА REFERENCES

- [1]. <http://esp8266.ru> – Ресурс МК ESP-8266 .
- [2]. <https://github.com/nodemcu/nodemcu-firmware> - проект прошивки NodeMcu.
- [3]. <https://www.lua.org> - The Programming Language Lua.
- [4]. <https://datasheets.maximintegrated.com/en/ds/DS18S20.pdf> - Dallas Semiconductor DS18B20 .
- [5]. <http://users.telenet.be/davshomepage> - I2C протоколи: RC5, Sony SIRC, Panasonic, JVC, Daewoo (en).
- [6]. <https://github.com/markszabo/IRremoteESP8266> - проект, бібліотека для віддаленого управління по I2C інтерфейсу (англ.).

Розробка Концепції Інформаційної Системи Зберігання і Обробки Інформації про Реабілітацію Осіб з Особливими Потребами

Панфьорова Ірина
Кафедра Інформаційних управляючих систем
Харківський Національний Університет Радіоелектроніки
Харків, Україна
iryna.panforova@nure.ua

Кислинський Єгор
Кафедра Інформаційних управляючих систем
Харківський Національний Університет Радіоелектроніки
Харків, Україна
yehor.kyslynskyi@nure.ua

Development of the concept of information system for the storage and processing of information on rehabilitation of persons with special needs

Panforova Iryna
Department of Information System
Kharkiv National University of Radio Electronic
Kharkiv, Ukraine
iryna.panforova@nure.ua

Kislinskiy Egor
Department of Information Systems
Kharkiv National University of Radio Electronic
Kharkiv, Ukraine
yehor.kyslynskyi@nure.ua

Анотація. У цій статті наводиться короткий опис автоматизованої системи керування реабілітаційним терапевтом та переваги даної системи.

Ключові слова: автоматизована система управління, інваліди, ефективність реабілітації, переваги.

Abstract— This article is a brief description of an automated control system for a rehabilitation therapist and the advantages of this system.

Keywords— automated control system, disabled people, rehabilitation effectiveness, advantages.

I. ВСТУП

Збереження здоров'я значною мірою залежить від середовища проживання людини. Виділяють три групи факторів, що впливають на здоров'я людини:

- фізичні (забруднення повітря, води, ґрунту, харчових продуктів, шум, електромагнітні поля, радіація та ін.);
- психологічні (трудові, сімейні, особистісні, культурні відносини, психоемоційні впливи та ін.);
- соціальні (соціальні та політичні зміни, безробіття, дефіцит часу, енергії та ін.);

Різні обставини можуть привести до значного погіршення здоров'я людини.

На сьогоднішній день в Україні ситуація навколо людей з обмеженими можливостями м'яко кажучи не дуже сприйнятлива. Таким людям треба отримувати не тільки психологічну підтримку, але й фізичну. Саме тому зараз затребуваною стає професія тренера-реабілітолога. Це спеціаліст, який займається адаптивною фізичною культурою з людьми з обмеженими можливостями та особами, що мають відхилення в стані здоров'я, всіх вікових і нозологічних груп з метою створення реальних можливостей для їх інтеграції в суспільство.

Одразу виникає потреба у розробці необхідного програмного забезпечення для тренера-реабілітолога, а саме інформаційно-управляючої системи обліку (ІУС) та контролю процесу реабілітації клієнтів (пацієнтів).

II. ПЕРЕВАГИ СИСТЕМИ

ІУС повинна мати архітектуру, що задовольняє вимогам відкритості, стандартності, модульності, розподіленості та іншим вимогам, що пред'являються до сучасних системам автоматизації.

Застосовувані при розробці АСУ інструментальні засоби та інформаційні технології повинні забезпечувати реалізацію основних вимог. Розглянемо основні переваги такої системи як для пацієнта, так і для тренера:

Зручність, швидкість і повнота. Система повинна містити картотеку пацієнтів, що включає в себе не тільки стандартні дані про пацієнта (ім'я, місце проживання, номер телефону, тощо), але й дані про його первинний стан здоров'я, а також можливість відслідковувати стан протягом проходження пацієнтом реабілітації. Записи у такій картотеці можна швидко відсортувати і знаходити необхідні дані по будь-яким з полів.

Аналіз даних про реабілітацію. Саме аналіз даних про реабілітацію пацієнта, про який було сказано у попередньому пункті є важливим плюсом такої системи. Записані дані у процесі аналізу скорочують частку рутинної роботи тренера, зменшують ймовірність помилок, їх не можна втратити, а інформація може використовуватися навіть в роботі одночасно декількома тренерами або надана лікарю. Також можливість відслідковування прогресу лікування звісно допоможе побачити як проходить реабілітація. За цими даними тренер має змогу переконатися, що реабілітація проходить правильно чи зафіксувати недостатні позитивні зміни, або навіть погіршення стану клієнта і тоді, зробивши певні висновки, додати нові фізичні навантаження або змінити вже використані раніше.

Сфера застосування. Система може використовуватись тренерами-реабілітологами не тільки у реабілітаційних центрах і спортивних комплексах, а й у соціальних центрах адаптації людей з обмеженими можливостями, спеціалізованих школах і дитячих садках, санаторно-курортних комплексах.

Надійний захист і легкість модифікації. Гнучка і високонадійна система аутентифікації необхідна для підтримки різних застосовуваних в реабілітаційному процесі моделей авторизації та процедур розмежування доступу. Також система повинна відповідати таким параметрам, як масштабованість, модульність і легкість модифікації системи. А її інтерфейс бути інтуїтивно-зрозумілим і легким в освоєнні і роботі[1, 2].

III. ПРИНЦИПИ РОБОТИ СИСТЕМИ

Тренер повинен не тільки зібрати необхідну інформацію про пацієнта та занести її у систему, очевидно, що в такій системі немає необхідності, адже інформацію про зміни стану здоров'я треба обробляти та використовувати для покращання процесу реабілітації.

Процес фізичної реабілітації обов'язково повинен супроводжуватися фіксацією основних показників здоров'я. Такими показниками можуть бути:

- маса тіла;
- артеріальний тиск систолічний (верхній);
- життєва ємність легенів;
- сила кисті;
- частота серцевих скорочень (при виконанні різних вправ, а також у спокійному стані);
- тривалість сну;
- функціональний стан органів і систем: норма спокою, норма реакції, резервні можливості та ін.

Також треба мати на увазі, що для людей з обмеженими можливостями більшість показників буде відрізнятися від стандартів здорової людини, а тому алгоритм роботи системи базується саме на показниках для людей з обмеженими можливостями[3].

Показники стану здоров'я дають змогу системі, на основі введених тренером даних відслідковувати ефективність реабілітації. Ефективність – складний феномен, він визначається перш за все поставленими цілями і досягненими результатами, тобто ефективність можна розглядати як ступінь реалізації мети або ступінь досягнення очікуваних результатів.

Оцінка ефективності реабілітації неможлива без використання критеріїв і показників, сутність яких полягає в тому, що вони показують, на скільки кінцеві результати збігаються з поставленою метою. Саме користувач системи (тренер) встановлює мету – необхідні кінцеві, або проміжні результати реабілітації.

Математична реалізація знаходження ефективності реабілітації на перший погляд дуже проста, але й у той самий час раціональна і ефективна. Кожен показник стану здоров'я оцінюється за такими правилами:

- якщо показники критерію оцінки повністю задовольняють поставленим цілям, то такому критерію відповідає 1 бал оцінки;
- якщо частково задовольняє, то 0,5 бала;
- якщо зовсім не задовольняє, то 0 балів;

Результати оцінки по окремим критеріям (в балах) сумуються і діляться на кількість оцінюваних критеріїв:

$$E = (N_1 + N_2 + \dots + N_n) / n,$$

де E - ефективність реабілітації;

$N_1 + N_2 + \dots + N_n$ - сума кількісних показників критеріїв;

n – кількість оцінюваних критеріїв.

Отриманий кінцевий результат оцінюється таким чином:

- висока ефективність реабілітації - підсумкове значення в межах 0,9 - 1 балів;
- помірна ефективність реабілітації - підсумкове значення в межах 0,7 - 0,89 балів;
- недостатня ефективність реабілітації - підсумкове значення в межах 0,5 - 0,69 балів;
- низька ефективність реабілітації - підсумкове значення в межах 0,3 - 0,49 балів;
- неефективна реабілітація - підсумкове значення нижче 0,3 бала.

Таким чином необхідні критерії та методи оцінки ефективності реабілітації мають високу інформативність і практичну змістовність, що дозволяє тренеру своєчасно вносити корективи в процес реабілітації[4].

ВИСНОВКИ

Запропонована ІУС виконує аналіз введених тренером даних, щодо процесу реабілітації, що робить можливим отримати показники ефективності реабілітації та зробити на основі цих даних висновки, щодо успішного проходження реабілітації.

Метод розрахування ефективності реабілітації є ефективним рішенням оцінки якості самого процесу реабілітації. На його основі тренеру набагато зручніше планувати протікання процесу і корегувати його для досягнення поставлених цілей.

ЛІТЕРАТУРА REFERENCES

- [1] Zharin V. A. Voronov Automated control system for medical center / V. A. Zharin, E. N. Voronov. – 2004
- [2] Rakhimov, T.N. Basics of building ACS / T.N. Rakhimov, O. A. Zaikin, B. Ya. Sotov. - Tashkent: Ukituvchi, 1984. - 376 p.
- [3] Marushchuk, LA, Complex Rehabilitation of Disabled Persons by Means of Adaptive Physical Culture: Dis. Cand. ped. Sciences: 13.00.04 / Marushuk L. A., 2009. - 176 p.
- [4] T. N. Fedorova. Complex rehabilitation of patients and disabled people / T. N. Fedorova, N. A. Nalobina, N. N. Lazareva., 2012. - 190 p. - (SibGUFK). - (5981)

Implementation of Marketing Information System (MIS) at The Enterprise for Decision-Making Tasks

Vladislav Shevchenko
Department of Software Engineering and Management
Information Technologies
NTU «KhPI»
Kharkov, Ukraine
ffshevchenko@gmail.com

Mariia Kozulia
Department of Software Engineering and Management
Information Technologies
NTU «KhPI»
Kharkov, Ukraine
mariya.kozulya7@gmail.com

Впровадження Маркетингових Інформаційних Систем (МІС) на Підприємстві для Задач Прийняття Рішень

Владислав Шевченко
Програмна інженерія та інформаційні технології
управління НТУ «ХПІ»
м. Харків, Україна
ffshevchenko@gmail.com

Марія Козуля
Програмна інженерія та інформаційні технології
управління НТУ «ХПІ»
м. Харків, Україна
mariya.kozulya7@gmail.com

Abstract— Relevance of marketing researches for any enterprise activity is an important tool in the complex solution of such tasks as: drawing up market characteristics, its potential opportunities, research on the distribution of market shares among different firms, sales variant, careful study of business activity trends, the discovery of competitors' goods. Research results on target possibility assessment of marketing information systems introduction in enterprises operating in uncertainty conditions of competition in the market is given. It is proposed to use the findings on companies state analysis of Acer and Lenovo based on the method of assessing the degree and level of risk when disclosing the uncertainty of counterparties.

Keywords—marketing information systems, decision making tasks, marketing decisions.

Анотація— Актуальність маркетингових досліджень для діяльності будь-якого підприємства є важливим інструментом при комплексному вирішенні таких завдань як: складання характеристики ринку, його потенційних можливостей, дослідження з розподілу часток ринку між різними фірмами, варіант збуту, ретельне вивчення тенденцій ділової активності, виявлення товарів конкурентів. представлені результати досліджень за оцінкою цільової можливості впровадження маркетингових інформаційних систем на підприємствах, що працюють в умовах невизначеності конкуренції на ринку. При цьому запропоновано використовувати висновки щодо аналізу стану компаній Acer і Lenovo на основі методу оцінки ступеня та рівня ризику під час розкриття невизначеності дій протидіючих сторін.

Ключові слова — маркетингові інформаційні системи, задачі прийняття рішень, маркетингові рішення,

I. INTRODUCTION

Relevance of marketing researches for any enterprise activity is an important tool in the complex solution of such tasks as: drawing up market characteristics, its potential opportunities, research on the distribution of market shares among different firms, sales variant, careful study of business activity trends, the discovery of competitors' goods. Marketing is becoming more diverse and complex form of activity, which requires a continuously integrated information process. It is imperative that the company develop and use a system of continuous environmental monitoring and data storage so that they can be analyzed in the future. A marketing information system can be defined as a set of procedures and methods designed to create, analyze and disseminate information for proactive marketing decisions on a regular, regular basis [1].

II. AIM

The aim of this work is to analyze the situation of certain quality goods sales, oriented to the needs of the market in the context of limited production opportunities, based on the inclusion of information components in the decision-making system (DM) in the market research opportunities for supplying products of the required quality.

III. PROBLEM FORMULATION

In this paper, it is proposed to solve the following problems in achieving the objectives of marketing information systems (MIS):

1) give a description of the MIS capabilities in solving the tasks of marketing research at enterprises in solving the

product promotion problems in the conditions of competitor existence in the market;

2) determine the decision making strategy by the manufacturer in uncertainty conditions of action against the manufacturer.

IV. RESULTS OF THE RESEARCH AND THEIR ANALYSIS

The traditional MIS model consists of four subsystems, each of which performs its task (Fig. 1) [2]: internal reporting; external marketing information (marketing surveillance or intelligence); marketing research; marketing information analysis.

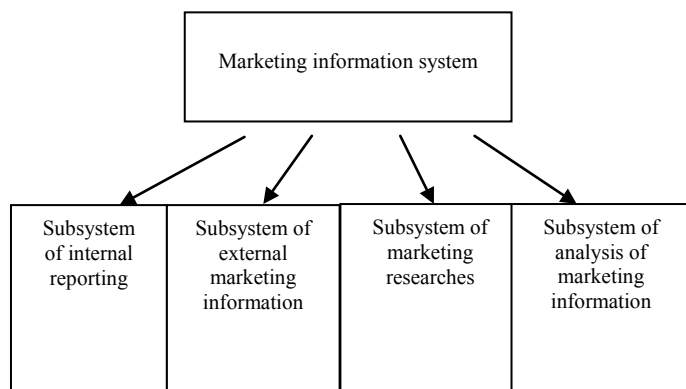


Fig. 1. Marketing information system

The basis of marketing information analysis subsystem is the bank of methods and models.

Bank of methods – a set of modern methods of statistical information processing, allowing to produce arithmetic operations, graphical and statistical processing of data and establish the degree of their statistical reliability. These techniques are used to solve tasks such as [3]:

- 1) variables definition that influence the sales and determine the significance of each of them;
- 2) dependence establishment of sales volume on the price and cost of advertising and promotion;
- 3) optimal variables determination for market segmentation, etc.

The practice of creating and using the MIS by foreign companies highlighted a number of problems associated with its operation. These include, in the first place, the rigid information system structure, which does not allow to change data format, as well as methods and models of analysis. Another disadvantage is the complexity of system's development, and it is virtually impossible to create a system data format that is optimal for all users [4].

When creating the MIS, there are transparent factors, methods of analysis and models that guide managers of different levels when making decisions on specific issues. Not everyone is ready to disclose their experiences and secrets, and without it it's almost impossible to develop satisfactory reports forms. It should be noted high material and time costs necessary to develop, install and study the use of application software packages. All these problems determine the need to look for other variants of information

systems for implementation in the system of decision making in enterprises operating in market uncertainty conditions with competitors. To do this, the paper proposes to assess the level and degree of risk when disclosing the counteraction uncertainty of the firm, for example, firms Acer and Lenovo [5].

When disclosing the counterparty uncertainty, there is a task of system risk level and degree assessment. The approach does not take into account the probability of choosing an opposing firm for a particular strategy, that is, the choice probability, $x_2 = x_2^*$, $x_2 = x_2^{**}$ or any other $x_2 = x_2^{***}$ actions in competitive disinformation conditions at market.

In the real firms counteraction conditions situation on information space of economic activity of enterprises or firms associated with intentional information "leakage" about the technical products characteristics, in particular functional features of the laptop, which does not correspond to reality, and so on. The risk degree in the case of counteracting the parties depends both on the probability of the opponent's a particular strategy choice, and on the decision recognition likelihood. Risk level determines the size of each firm loss. To solve the problem of minimizing the losses risk and economic losses of firms Acer and Lenovo in a competitive market in which misinformation is present, it is proposed to employ system analysis tools in the part of assessing method the level and degree of risk in disclosure counterparties uncertainty.

For example, the second firm made a decision $x_2 = x_2^{**}$, and the first has built a strategy on condition $x_2 = x_2^*$, then the loss level U_1 of the first firm is defined as

$$U_1 = \Delta f_1 = \max_{x_1} f_1(x_1, x_2^{**}) - \max_{x_1} f_1(x_1, x_2^*) \quad (1)$$

The first firm is likely to make a decision $x_1 = x_1^{**}$, and the second builds the strategy provided that $x_1 = x_1^*$. In this case, the level of loss U_2 of the second company is

$$U_2 = \Delta f_2 = \max_{x_2} f_2(x_1^{**}, x_2) - \max_{x_2} f_2(x_1^*, x_2) \quad (2)$$

Accepted estimates (1) and (2) show how much the result exceeds the guaranteed, that is, they are estimates below.

In the this method framework of market situation analysis, another approach is proposed to determine the level of risk – a results compare with the potentially achievable, that is, how much the result is approaching the maximum possible.

In this case, it is assumed that the maximum possible value of the target function of the first firm will be the highest priority

$$f_1^* = \max_{x_1} \max_{x_2} f_1(x_1, x_2),$$

Then company loss level will determine the extent to which the result is different from the potentially achievable.

Level of loss U_1^* for the first firm subject to the choice of actions x_2 the second company in the form $x_2 = x_2^*$ is determined by the ratio

$$U_1 = f_1^* - \max_{x_1} f_1(x_1, x_2^{**}),$$

For the second firm, the loss rate is

$$U_2 = f_2^* - \max_{x_2} f_2(x_1^{**}, x_2),$$

where the maximum possible value of the target function is defined as

$$f_2^* = \max_{x_2} \max_{x_1} f_2(x_1, x_2),$$

For critical situations analysis, when one firm causes irreparable harm to another, with this description, the relationship between sales companies is determined as the situation approaches the critical, which is characterized by an absolute minimum. The absolute minimum for the first firm is calculated from the condition

$$f_1^- = \min_{x_1} \min_{x_2} f_1(x_1, x_2),$$

Then value U_1^- will determine difference level between the target functions of the situation and the critical situation

$$U_1^- = f_1(x_1', x_2') - f_1^-,$$

where x_1', x_2' – numerical values x_1, x_2 , that correspond to the situation that occurred for the first firm.

The second firm state is determined on the condition of such losses existence

$$U_2^- = f_2(x_1'', x_2'') - f_2^-,$$

where the absolute minimum equals

$$f_2^- = \min_{x_1} \min_{x_2} f_2(x_1, x_2),$$

where x_1'', x_2'' – numerical values x_1, x_2 in the situation that occurred for the second firm.

Thus, the applied situation analysis on the sales market of HP and Lenovo firms by the level and degree risk assessment method in disclosure of uncertainty actions counter parties can give a basis for making decisions on the certain quality products preparation for its effective marketing on the market in a competitive environment.

CONCLUSIONS

In this paper, research results on target possibility assessment of marketing information systems introduction in enterprises operating in uncertainty conditions of competition in the market is given. It is proposed to use the findings on companies state analysis of Acer and Lenovo based on the method of assessing the degree and level of risk when disclosing the uncertainty of counterparties.

REFERENCES ЛІТЕРАТУРА

- [1] V. P. Melnikov, A. G. Shirladze "Research of control systems. Textbook for Academic Bachelor ". M.: Urait, 2018, pp. 250-296.
- [2] G. A. Polinskaya "Information Systems of Marketing. Textbook and Workshop for Academic Bachelor's Degree ". M.: Urait, 2014, pp. 203-205.
- [3] Varenko V.M. "Information and Analytical Activity" Textbook K.: University "Ukraine", 2013, 416 c.
- [4] Soloviev B.A. "Marketing": Textbook M.: INFRA-M, 2007, 383 p.M. Z. Zgurovsky, N. D. Pankratova, "Fundamentals of System Analysis". K.: BHV, 2007, pp. 169-170.

Система Тестування Психофізіологічних Властивостей Водіїв

Поворознюк А.І.
НТУ «ХПІ»
Харків, Україна
ai.povoroznjuk@gmail.com

Черних О.П.
НТУ «ХПІ»
Харків, Україна
lenachernikh@gmail.com

Стрюков А.О.
НТУ «ХПІ»
Харків, Україна
andrey.strukov.96@gmail.com

Двараковська О.В.
НТУ «ХПІ»
Харків, Україна
sasha_dvarakovsk@mail.ru

Testing system of Psychophysiological Properties of Drivers

Povoroznyuk A.
NTU «KhPI»
Kharkiv, Ukraine
ai.povoroznjuk@gmail.com

Chernykh O.
NTU «KhPI»
Kharkiv, Ukraine

Stryukov A.
NTU «KhPI»
Kharkiv, Ukraine
andrey.strukov.96@gmail.com

Dvarakovskaya A.
NTU «KhPI»
Kharkiv, Ukraine
sasha_dvarakovsk@mail.ru

Анотація — Досліджені основні методики для оцінки психофізіологічного стану людини. На основі результатів розроблено систему тестування психофізіологічних властивостей водіїв. Для тестування водія були обрані напрямки оцінювання: інтелектуальний розвиток, психомоторні характеристики, особливості особистості.

Ключові слова—Психофізіологічні Властивості, Водій, Тестування, Увага, Зір, Реакція, Дальтонізм, Знання Правил Дорожнього Руху.

I. ВСТУП

Проблема безпеки дорожнього руху з року в рік стає більш гострою. У функціональній системі «водій – автомобіль – середовище руху» однією з основних ланок забезпечення безпеки руху є водій. За даними Управління безпеки дорожнього руху України у ДТП з постраждалими приблизно 48,5% випадків скоєні з вини водіїв.

II. ПОСТАНОВА ЗАДАЧІ

Мета роботи – розробка системи тестування психофізіологічних властивостей водіїв.

Для досягнення зазначеної мети були поставлені основні задачі:

- якісна та швидка оцінка психофізіологічного стану водія;
- інформування водіїв про психофізіологічні особливості їх організму;
- підвищення рівня безпеки дорожнього руху.

В ході розробки реалізовано наступні функціональні можливості системи:

- тестування водія за обраним видом тесту в обраній категорії;
- забезпечення фіксування інформації під час тестування та її швидка обробка;
- оперативна видача отриманих результатів у вигляді протоколу та можливість його збереження у форматі PDF;
- підтримка української, російської та англійської локалізацій.

III. СИСТЕМА ТЕСТУВАННЯ ВОДІІВ

Оцінка психофізіологічних особливостей [1] та опис обстежуваного у розроблюваній системі повинні виконуватися з точки зору:

- психомоторних характеристик;
- особливостей особистості, у тому числі таких, як здатність діяти в складних ситуаціях і соціальна зрілість психомоторних характеристик;
- інтелектуального розвитку та процесів пізнання.

Для тестування водіїв обрано наступні групи і види тестів [2-5], результати яких оцінюються за відповідними формулами:

1) «Увага» – таблиці Шульте:

$$K_l = \tau/30, \quad (1)$$

де τ – час проходження тесту, с.

Суть тесту: якнайшвидше обрати числа (від 1 до 25) у порядку зростання / спадання. Оптимальний час для виконання завдання – 30 секунд. Результат тесту – коефіцієнт, що обчислюється як відношення реального часу виконання тесту водієм до оптимального часу.

2) «Зір» – кільця Ландольта:

$$K_2 = \sum_{k=1}^3 \left(\frac{m}{n} * k \right) \quad (2)$$

де m – кількість правильних відповідей;

n – загальна кількість відповідей;

k – номер етапу тестування.

Суть завдання: обрати кільця потрібного вигляду. Тестування проходить в три етапи. На кожному етапі на екран виводиться комбінація з n -ої кількості кілець. Результат тесту – коефіцієнт, що обчислюється як сума проміжних результатів на кожному етапі.

3) «Реакція» – тест червоних і чорних квадратів:

$$K_3 = \left(\frac{m_1}{n_1} + \frac{m_2}{n_2} \right) / 2, \quad (3)$$

де m – кількість правильних відповідей;

n – загальна кількість відповідей.

Тест проходить в два етапи. На екрані розміщено квадрати червоного та чорного кольору. Суть завдання: на першому етапі потрібно обрати квадрати тільки червоного кольору, на другому – тільки чорного. Результат тесту – коефіцієнт, що обчислюється як середнє арифметичне від проміжних результатів на кожному етапі.

4) «Дальтонізм» – таблиці Рабкіна:

$$K_4 = \frac{m_1 + m_2 + \dots + m_n}{p}, \quad (4)$$

де n – номер питання;

p – загальна кількість відповідей;

$$m_n = \begin{cases} 0, & \text{якщо відповідь неправильна} \\ 1, & \text{якщо відповідь правильна.} \end{cases}$$

5) «Знання правил дорожнього руху» – тестування у вигляді питань з кількома варіантами відповідей:

$$K_5 = m/n * 100\%, \quad (5)$$

де m – кількість правильних відповідей;

n – загальна кількість відповідей.

Тест проводиться в n -ну кількість етапів. На кожному етапі на екран виводиться картинка, в якій зашифрована певна цифра. Людина з порушеним кольоровим сприйняттям не зможе побачити правильну відповідь. Результат тесту – коефіцієнт, що обчислюється як середнє арифметичне від проміжних результатів на кожному етапі.

IV. ПРОГРАМНА РЕАЛІЗАЦІЯ

Для розробки системи психофізіологічного контролю водіїв [6] було обрано мову програмування JavaScript та спеціальний фреймворк Electron.js, який дозволяє запускати програмний застосунок на найбільш популярних сучасних операційних системах. Для роботи з базами даних для збереження результатів проходження тестів було обрано систему MongoDB, яка забезпечує швидку роботу та оновлення бази даних.

По закінченню тестування водія видається протокол тестування, що складається з особистих даних тестованого і таблиці з результатом пройдених ним тестів. Система має можливість його збереження у форматі PDF. Таблиця містить наступні колонки: назва тесту, результат тестованого і норма тестування. Тести, які мають незадовільний результат, будуть виділені і в кінці протоколу будуть залишені рекомендації щодо поліпшення поточних показників. Розроблена система має підтримку української, російської та англійської локалізацій.

V. ОБГОВОРЕННЯ

Керування автомобілем зобов'язує мати гострий зір, швидку реакцію, зосереджену увагу, знати правила дорожнього руху. Звичайно, не всі вищеописані характеристики мають бути ідеальними. Але, знаючи про свої недоліки, водій може компенсувати їх більшою обережністю або стилем водіння. Дослідження вчених довели, що коли водії проінформовані про психофізіологічні особливості їх організму, аварійність зменшується на дві третини.

ВИСНОВКИ

Для якісної та швидкої оцінки стану водія необхідна система тестування їх психофізіологічного стану, яка включає в себе тести на увагу, на зір і реакцію, на дальтонізм і знання правил дорожнього руху.

Застосування такої системи приведе до зниження помилкових дій водіїв, від яких безпосередньо залежить безпека дорожнього руху.

REFERENCES ЛІТЕРАТУРА

- [1] Гюлев Н. У. Особливості ергономіки та психофізіології в діяльності водія: навч. посібник / Н. У. Гюлев; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2012. – 185 с.
- [2] Методика «Таблицы Шульце» / Альманах психологических тестов. М., 1995. – С.112-116.
- [3] Лещенко И. А. О системах и правилах определения остроты зрения // Вестник оптометрии. 2009. № 3. С. 54–58.
- [4] Рабкин Е. Б. Полихроматические таблицы для исследования цветоощущения. – Минск, 1998. – 144 с.
- [5] Фоменко О.Я., Раціборинський Б.Л., Гусар В.С. Правила дорожнього руху України 2017. – Київ, 2017. – 112 с.
- [6] Э. Фримен, Э. Фримен. Изучаем HTML, XHTML и CSS. – П.: «Питер», 2010. – 656 с.
- [7] Поворознюк А.І., Черних О.П., Двараківська О.В., Стрюков А.О. Розробка системи психофізіологічного контролю водіїв. Інформатика, управління та штучний інтелект. Матеріали третьої міжнародної науково-технічної конференції студентів, магістрів та аспірантів. – Х.: НТУ "ХПІ", 2016. – С. 68.

Модель Адаптивного Контролера на Основі Нечіткої Логіки для Управління Системою "Розумний Дім"

Ємельянов Г.С

Факультет математики, фізики
та інформаційних технологій

Одеський національний університет імені І.І. Мечникова

Одеса, Україна
emelyanovjora95@gmail.com

Крапівний Ю.М

Факультет математики, фізики
та інформаційних технологій

Одеський національний університет імені І.І. Мечникова

y.krapivny@gmail.com

Model of Adaptive Controller on the Basis of Fuzzy Logic for the Management of the System "Smart House"

Emelyanov Georgy

Faculty of Mathematics, Physics and Information Technologies

Odessa National University Mechnikov

Odessa, Ukraine

emelyanovjora95@gmail.com

Krapivniy Yu.M

Faculty of Mathematics, Physics and Information Technologies

Odessa National University Mechnikov

Odessa, Ukraine

y.krapivny@gmail.com

I. ВСТУП

Анотація — Робота присвячена створенню програмно-апаратного комплексу "Розумний дім", а також розробці програмних модулів ЧПУ для управління цим комплексом. Ключовою метою даного проекту, стане впровадження методів штучного інтелекту в практику автоматичного управління складними технічними об'єктом «Розумний дім, а також генерація керуючих команд для апаратного комплексу, що забезпечують його коректну роботу і перегляд інформації про стан приміщення на основі датчиків.

Ключові слова—Розумний дім, Нечітка логіка, логічний висновок Мамданді

Abstract—The work is devoted to the creation of software and hardware complex "Smart Home, as well as the development of CNC software modules to control this complex. The key objective of this project will be the introduction of artificial intelligence methods into the practice of automatic control of complex technical objects "Smart House", as well as the generation of control commands for the hardware complex, ensuring its correct operation and viewing information about the state of the room based on sensors.

Keywords—Smart House, Fuzzy Logic, Mamdani fuzzy logic conclusion

При вирішенні завдань управління та прийняття рішень в умовах нечіткості (ненадійності) вихідних даних важко, а часто і неможливо використовувати точні методи. У таких випадках хороші результати можна

отримати, використовуючи методи інтелектуальних систем, наприклад - методи, засновані на нечіткій логіці.

II. РЕАЛІЗАЦІЯ ЗАДАЧІ

Робота присвячена розробці програмно-апаратної моделі управління робототехнічної системою, яка може бути покладена в основу управління "Розумним будинком" [1]. Основою програмної частини системи є інтерпретатор завдань нечіткого управління, що використовує нечітку базу знань і нечіткий логічний висновок Мамдані [2]. Програмна система призначена, для використання на одноплатному комп'ютері Arduino, яка забезпечує моделювання багатофункціонального адаптивного нечіткого контролера (рис. 1).



Рис. 1. Архітектура комплексу «Розумний будинок»

На вході інтерпретатор приймає код на мові опису моделі нечіткої системи управління. Кожна модель здатна генерувати і підтримувати роботу окремого програмного контролера на основі нечіткої логіки, що взаємодіє з апаратною частиною системи. Створити код з описом нечіткої системи управління можна двома способами:

- Вручну, задавши спеціальні команди і провівши налаштування через мобільний додаток.
- В пакеті Fuzzy Logic Toolbox через спеціалізоване програмне забезпечення - MatLab [3].

Для роботи системи потрібен чітко заданий набір вхідних параметрів, які можуть бути зокрема добуті з різних датчиків (датчик руху, відстані, температури і т.п.). На виході буде отриманий спеціальний сигнал як результат нечіткого виведення для формування керуючого сигналу [4]. Цей керуючий сигнал потрапить на плату управління, а потім піде на об'єкт управління.

Для зручної роботи користувача з системою, а також попередньої настройки і перегляду статистики буде створено мобільний додаток. Мобільний додаток в даному проєкті грає ключову роль, так як є прошарком між апаратною частиною проєкту, як постачальник даних і базою даних, як обробник і сховище даних [5]. Завдяки наявності у клієнта інтуїтивно зрозумілого інтерфейсу, користувач може без будь-яких проблем знайти потрібну йому інформацію. Мобільний додаток буде реалізований на операційній системі Android.

III. ПРАКТИЧНЕ ЗАСТОСУВАННЯ

Інтерпретатор мови нечіткої логіки, може бути використаний для розробки і налагодження систем нечіткого управління в самих різних предметних областях.

Різноманітним може бути застосування системи в оборонній сфері. Даним комплексом можна обладнати різні типи приміщень: бараки, склади та інші

приміщення, в яких потрібно автоматизація процесу збереження стан середовища, спрощення управління побутовою технікою, функції охорони на великих просторах. Просте управління і можливість налаштування через мобільний пристрій сильно спростить роботу оператора даної системи. Використання нечіткої бази знань, допоможе налаштуватися під різні умови середовища і умови задані оператором. Також, ця системи може бути використана для околівених цілей. Наприклад, спростити життя учасникам бойових дій, які постраждали під час несення бойової служби [6].

При удосконаленні системи, в майбутньому, буде доступно підключення більшої кількості датчиків, яке істотно розширить можливості системи.

ЛІТЕРАТУРА REFERENCES

1. Емельянов Г.С., Крапивный Ю.Н., «Числовое программное управление «Умным домом». Тезисы доклада XIII всеукраинской конференции студентов и молодых учёных "Информатика, информационные системы и технологии". - Одесса: - 2018г., с.152-154.
2. Штовба С. Д. - Проектирование нечетких систем средствами MATLAB. М.: Телеком, 2007.-288 с.
3. Fuzzy Logic Toolbox [Електронний ресурс] - Режим доступу: <https://www.mathworks.com/help/fuzzy/>.
4. Штовба С.Д. - Введения в теорию нечітких множин і нечітку логіку.
5. Мобільні додатки для «Розумного будинку» [Електронний ресурс] - Режим доступу: <https://ichip.ru/6-mobilnykh-prilozhenij-kotorye-sdelayut-vash-dom-umnym.html>.
6. Smart Homes for Vets [Електронний ресурс] - Режим доступу: <https://www.military.com/military-report/smart-homes-for-vets.html>

Математична Модель Задачі Планування Групових Розміщень в Системах Управління Готельним Бізнесом

Оксана Мазурова
кафедра програмної інженерії
Харьковский национальный университет
радиоэлектроники
Харьков, Украина
oksana.mazurova@nure.ua

Юлія Грушко
кафедра програмної інженерії
Харьковский национальный университет
радиоэлектроники
Харьков, Украина
yuliia.hrushko@nure.ua

Mathematical Model of the Problem of Planning of Group Placements in Hotel Business Management Systems

Oksana Mazurova
Department of Software Engineering
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
oksana.mazurova@nure.ua

Yuliia Hrushko
Department of Software Engineering
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
yuliia.hrushko@nure.ua

Анотація—Робота присвячена математичному моделюванню задачі планування та оптимального розміщення груп постояльців в системах керування готельним бізнесом. Для підтримки групових розміщень людей за готельними номерами в роботі розглянута задача календарного планування з елементами багатокритеріальної оптимізації на базі адитивної згортки.

Abstract—The work is devoted to mathematical modeling of the problem of planning and optimal placement of groups of guests in hotel business management systems. To support group allocations of people by hotel numbers, the task of scheduling with elements of multicriteria optimization based on additive convolution is considered.

Ключові слова—математична модель, задача планування, багатокритеріальна оптимізація, критерій, план, бронювання, готельний номер.

Keywords—mathematical model, planning problem, multicriteria optimization, criteria, plan, reservation, hotel room.

I. ВСТУП

Готельний бізнес за популярністю займає перше місце в світі, а в багатьох країнах є основою економіки. Головна мета успішної організації готельного бізнесу - це прагнення до рівномірного заповнення всіх номерів, незважаючи на сезонний характер діяльності в цьому сегменті. Однією з найважливіших функцій управління

готелем є бронювання та управління готельними номерами.

На сьогодні успішні в своїй галузі готелі використовують веб-додатки, що надають можливість клієнтам швидко та зручно обирати номери, бронювати їх [2]. Управління номерним фондом готелю передбачає обслуговування заявок, в тому числі, і на групове заселення клієнтів. Але ефективно групове заселення з урахуванням усіх потреб клієнтів підтримується далеко не в усіх програмних додатках. Наприклад, система www.booking.com спрямована на пошук номерів у зазначених містах та їх порівняння, а система khariv-palace.phng.com – це типовий веб-додаток, що підтримує базові функції готелю. Зазвичай, функція бронювання в таких системах є або вторинною, або не дозволяє робити більш складні запити на бронювання, що не в повній мірі відображає усі можливості окремих готелів.

II. НАУКОВА ПРОБЛЕМАТИКА

Науково-обґрунтоване планування розселення з урахуванням усіх потреб клієнтів та оптимальне навантаження номерів є важливим чинником успіху та покращення діяльності готельного бізнесу. Існують різні види планування, які розрізняються за метою, рівнем, предметом, періодом, призначенням та вмістом.

Календарне планування дозволяє моделювати процеси визначення черговості використання матеріальних і

виробничих ресурсів на певний часовий період з метою забезпечення заданих термінів виконання замовлень [1]. Отже, календарне планування є одним перспективних напрямків, який дозволить ефективно промодельовати задачу групового заселення клієнтів та врахувати розподілені за часом заселення та броні готельних номерів.

Крім того, ефективне обслуговування групових заселень передбачає необхідність врахувати переваги цілої низки клієнтів: тих, що заселяються групою, та тих, яких ці заселення можуть потурбувати. Отже, існує необхідність в багатокритеріальній оптимізації можливих планів заселення. Теорія корисності та її багатокритеріальні згортки дозволяють ефективно вирішувати оптимізаційні задачі, що характеризуються помірною множиною кількісних критеріїв [4].

Для системи підтримки планування розселення було обрано календарне планування, враховуючи те, що зайнятість номерів упорядкована в часі. Для знаходження оптимального плану розселення було обрано метод багатокритеріального прийняття рішень з використанням лінійної адитивної згортки [3].

III. ПОСТАНОВКА ЗАДАЧИ

В роботі була поставлена задача побудувати математичну модель задачі планування групових розміщень клієнтів та підтримати вибір оптимального плану з урахуванням переваг клієнтів. Розроблена модель буде використана під час розробки програмної системи управління готельним бізнесом. Така система має стати зручним інструментом для менеджерів готелю та інших працівників, за допомогою якого вони змогли би побудувати оптимальний план заселень номерів на обраний період, у яких треба враховувати побажання клієнтів, надавати можливість контролю і управління номерами.

Для досягнення поставленої мети необхідно:

- розробити математичну модель задачі планування розселення;
- розробити оптимізаційну модель вибору найкращого плану розселення;
- програмно реалізувати систему на основі розроблених моделей.

IV. МОДЕЛЮВАННЯ

Задача планування Z розселення в системі готельного бізнесу може бути розглянута як:

$$Z = \langle H_b, H_e, P^{opt} \rangle,$$

де H_b – початкова ситуація бронювання в готелі,
– кінцева ситуація бронювання в готелі,

P^{opt} – оптимальний (найкращий) план з досягнення результуючої ситуації - $P^{opt} : H_b \rightarrow H_e$.

Ситуація H в готелі складається з таких важливих частин, як бронювання, наявні заявки на заселення та характеристики номерів на даний момент часу. Тобто ситуацію в готелі $H = H(t)$ в момент часу t можна описати як:

$$H = H(t) = H(R(t), N, A(t_b, t_e)),$$

де $R(t)$ – сукупність бронювань в момент часу t , N – наявні номери та їх характеристики, $A(t)$ – заява на заселення на проміжок часу від t_b до t_e .

Важливою складовою ситуації в готелі є бронювання номерів на певний проміжок часу, тобто час, який певний номер зайнятий нині або у майбутньому. Саме планування розкладу бронювань є основною метою даної роботи. Але в нашій ситуації нам важливо враховувати не тільки час бронювання, але й момент часу t , у який розглядається заявка, щоб мати правильне уявлення про ті бронювання, які можна змінювати, а які ні. Сукупність бронювань у готелі можна виразити як:

$$R(t) = \{R_i(t_b, t_e)\}_{i=1}^n$$

де R_i – конкретне i -те бронювання номеру,
 t_b – теперішній час або час розглядання заявки,
 t_e – дата виїзду,
 n – загальна кількість бронювань.

Певне бронювання у готелі можна визначити наступним чином:

$$R_i(t_b, t_e) = \langle N(t_b, t_e), B, C, K \rangle,$$

де t_b – теперішній час або час розглядання заявки,
 t_e – дата заїзду,
 t_e – дата виїзду,
 N – власне номер,
 B – мітка бронювання,
 C – клієнт, на якого оформлено бронювання,
 K – кількість клієнтів.

Заява на бронювання A також є важливою частиною ситуації, адже саме через виконання заяв змінюється ситуація в готелі та з'являються/змінюються бронювання в готелі. Заяву на заселення можна описати наступним чином:

$$A(t_b, t_e) = \langle C, D, Q, F \rangle,$$

де t_b – дата заселення,
 t_e – дата виїзду,
 C – клієнт, на якого оформлено заяву,
 D – сукупність вимог,
 Q – загальна кількість людей,
 F – бажаний поверх.

На рисунку 1 показано, як можуть накладатися бронювання та нові заявки один на одне.

Сукупність вимог, що необхідно врахувати під час обслуговування заявки виглядає як:

$$D = \{D_j\}_{j=1}^m$$

де m – загальна кількість вимог.

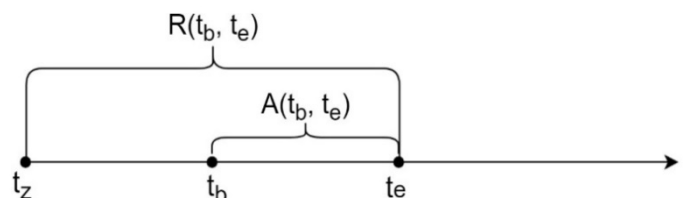


Рис. 1. – Графік накладання бронювання та заявки у часі для одного номеру

В свою чергу одну конкретну вимогу можна описати як:

$$D_j = \langle O, O_N, L_w, T_w, U_w, \{N^i\} \rangle,$$

де O – загальна кількість людей, для яких треба виконати вимогу,

O_N – бажана кількість людей у номеру,

L_w – бажаний поверх,

T_w – бажаний рівень комфорту,

U_w – бажані зручності,

$\{N^i\}$ – множина запропонованих номерів.

У ході рішення задачі ми маємо отримати множину планів P , із яких потім буде вибрано оптимальний. Множину планів P опишемо як:

$$P = \{P_i\}_{i=1}^k,$$

де P_i – конкретний i -тий план,

k – загальна кількість планів.

План із множини може бути описаний таким чином:

$$P_i = \langle P_s^i, P_r^i \rangle,$$

де P_s^i – множина заселень i -того плану,

P_r^i – множина переселень i -того плану.

Множина заселень представимо у вигляді:

$$P_s^i = \{P_{sj}^i\}_{j=1}^d,$$

де d – загальна кількість необхідних заселень.

Аналогічно представимо множину переселень:

$$P_r^i = \{P_{rj}^i\}_{j=1}^q,$$

де q – загальна кількість необхідних переселень.

У свою чергу кожне заселення представляє собою:

$$P_{sj}^i = \langle C, N \rangle,$$

де C – клієнт, на якого оформлено заяву,

N – номер, у який виконується заселення.

Переселення можна зобразити так:

$$P_{rj}^i = \langle \langle C, N' \rangle \langle C, N \rangle \rangle,$$

де C – клієнт, на якого оформлено заяву,

N – номер, у який виконується заселення,

N' – номер, з якого виконується виселення.

На рисунку 2 зображено схематичне надання заселень та переселень, які можуть входити до складу певного плану з розселення групи клієнтів.

Оптимальний план P^{opt} заселення з множини планів P може бути знайдено на базі наступної оптимізаційної моделі:

$$P^{opt} = \underset{P^{opt} \in P}{\text{extremum}} Q(P_i),$$

де P – множина планів заселення;

Q – функція оцінки якості плану заселення.

Для оцінки оптимальності плану будемо використовувати метод адитивної згортки з ваговими коефіцієнтами, отже:

$$P^{opt} = \max_{i=1, m} \sum_{j=1}^n \alpha_j \beta_j K_j(P_i),$$

де n – кількість критеріїв,

m – кількість планів,

α_j – нормований множник,

β_j – ваговий коефіцієнт,

$K_j(P_i)$ – оцінка плану P_i за критерієм K_j .

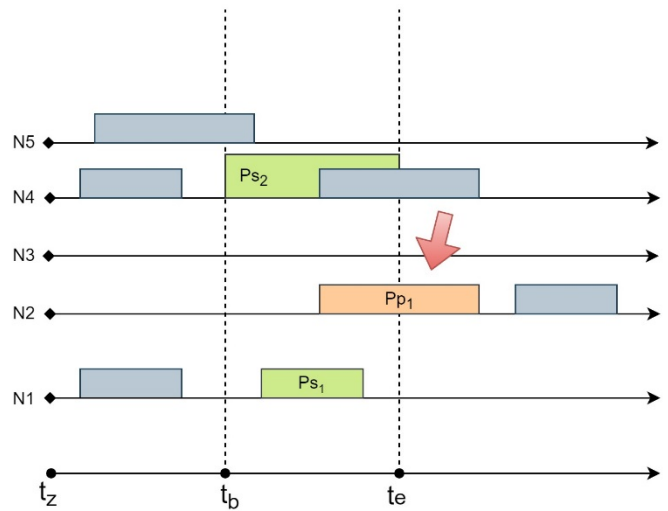


Рис. 2 – Схематичне надання заселень та переселень відповідно до плану

Сукупність критеріїв $K = \{K_j\}_{j=1}^m$ необхідна для того, щоб визначити який із запропонованих планів максимально вдовольняє вимогам та є оптимальним.

Для даної задачі були визначені наступні критерії якості:

- $K_1(P_i) = \min q$ – кількість переселень (найкращою вважається мінімальна кількість переселень),

- $K_2(P_i) = F(T)$ – вдоволення вимог за типом номерів,

$K_3(P_i) = F(L)$ – вдоволення вимог за поверхом,

$K_4(P_i) = F(U)$ – вдоволення вимог за зручностями в номерах,

$K_5(P_i) = F(O)$ – вдоволення вимог за кількістю людей у номерах.

IV РЕАЛІЗАЦІЯ СИСТЕМИ

Система управління готельним бізнесом була розроблена у вигляді веб-додатку. Для її реалізації було використано клієнт-серверну архітектуру, що є одним із архітектурних шаблонів програмного забезпечення та є домінуючою концепцією у створенні розподілених мережних додатків і передбачає взаємодію та обмін даними між ними.

Для зберігання даних, що необхідні під час планувань заселення та в цілому для управління готельним бізнесом, була розроблена база даних, основними таблицями якої є «Користувач», «Номер», «Заявка», «Вимога», «Бронювання».

Сутність «Бронювання», яка необхідна для закріплення конкретного номеру за конкретним клієнтом на певний проміжок часу, матиме наступні атрибути: унікальний ідентифікатор бронювання; ідентифікатор клієнту; номер; дата заїзду; дата виїзду; кількість людей; відмітка бронювання; сума; термін оплати; сплачено.

А сутність «Заявка», призначена для формування заявок на заселення із побажаннями клієнта та подальшому їх розгляданню матиме наступні атрибути: унікальний ідентифікатор заявки; ідентифікатор клієнту; кількість людей; бажаний поверх; дата заїзду; дата виїзду; розглянуто чи ні.

Сутність «Вимога» є складовою частиною заявки та виділяє наступні атрибути: унікальний ідентифікатор вимоги; ідентифікатор заявки; загальна кількість людей для вимоги; бажана кількість людей у номеру; бажаний поверх; бажаний рівень комфорту; зручності.

Система поділяється на окремі функціонально-незалежні модулі, серед яких модуль взаємодії з базою даних, сервіси, команди і веб-інтерфейс системи. Таке розподілення на окремі частини дозволяє застосовувати в системі не тільки вертикальне, але й горизонтальне масштабування.

Реалізація проекту відбувалася за допомогою мови програмування Java, JSONa, як форматом передачі даних, протокола HTTP, та HTML як мови розмітки інтерфейсу користувача. Під час роботи використовувалася IDE Eclipse. У якості СКБД виступає MySQL.

ВИСНОВКИ

В процесі виконання роботи був проведений аналіз методів календарного планування та розроблена математична модель планування розселення групи людей та визначена модель вибору оптимального плану розселення. Даний підхід в моделювання дозволяє не тільки врахувати потреби групи клієнтів під час розселення, а й оптимальним чином врахувати особливості обслуговування таких групових замовлень.

В ході роботи були виконані наступні завдання:

а) проведено аналіз та моделювання предметної галузі управління готельним бізнесом;

б) проведено аналіз методів планування, обрано та реалізовано метод календарного планування;

в) розроблено математичну модель задачі планування групових розселень, а також оптимізаційну модель пошуку оптимального плану;

г) програмно реалізовано веб-додаток бронювання та управління готельними номерами на основі розроблених моделей.

Розроблені в ході роботи моделі представляють собою математичний опис ситуації в готелі та вирішення задачі групового розселення. Вони можуть послужити основою для розробки більш складних варіантів задачі розселення групи людей за обраними критеріями.

REFERENCES ЛІТЕРАТУРА

- [1] Васильченко, М.Д. Планирование на предприятии. [Текст]/ М.Д. Васильченко – М.: Издательство: «Эксмо», 2007. – 100 с.
- [2] Смирнов, Н. И. JAVA 2 Enterprise. Основы практической разработки распределенных корпоративных приложений [Текст]/ Н.И. Смирнов. – М.: КУДИЦ-ОБРАЗ, 2002. – 240 с.;
- [3] В. И. Ширяев, В. И. Ширяев, “Принятие решений. Математические основы. Статические задачи”, Издательство: «Либроком», 2016.
- [4] Фишберн П. Теория полезности для принятия решений[Текст]/ П. Фишберн – М.: Наука, 1978. — 352 с.

Розробка САД-Системи для Моделювання Динаміки Штучних Біо-Структур

Катерина Газдюк
Відділ комп'ютерних технологій
Чернівецький національний університет
імені Юрія Федьковича
Чернівці, Україна
kateryna.gazdyik@gmail.com

Володимир Жихаревич
Відділ комп'ютерних технологій
Чернівецький національний університет
імені Юрія Федьковича
Чернівці, Україна
vzhikhar81@gmail.com

Сергій Остапов
Відділ комп'ютерних технологій
Чернівецький національний університет
імені Юрія Федьковича
Чернівці, Україна
sergey.ostapov@gmail.com

Денис Куц
Відділ комп'ютерних технологій
Чернівецький національний університет
імені Юрія Федьковича
Чернівці, Україна
dena.kuts1@gmail.com

CAD-System Development for Modeling Dynamics of Artificial Biostructures

Kateryna Gazdiuk
Department of computer technologies,
Yuriy Fedkovych Chernivtsi National University
Chernivtsi, Ukraine
kateryna.gazdyik@gmail.com

, Volodymyr Zhikharevich
Department of computer technologies,
Yuriy Fedkovych Chernivtsi National University
Chernivtsi, Ukraine
vzhikhar81@gmail.com

Sergiy Ostapov
Department of computer technologies,
Yuriy Fedkovych Chernivtsi National University
Chernivtsi, Ukraine
sergey.ostapov@gmail.com

Denys Kuts
Department of computer technologies,
Yuriy Fedkovych Chernivtsi National University
Chernivtsi, Ukraine
dena.kuts1@gmail.com

Анотація — Робота присвячена розробці комп'ютерної системи для моделювання динаміки біо-подібних процесів та структур. В якості методу моделювання обрано метод рухомих клітинних автоматів, що дозволяє описувати досліджувані процеси набором функцій взаємодій між сусідніми клітинами (агентами).

Abstract— The work is devoted to the development of a computer system for modeling the dynamics of bio-like processes and structures. As a method of modeling, the method of movable cellular automata is chosen, that allows describing the investigated processes by a set of interactions' functions between adjacent cells (agents).

Ключові слова— рухомі клітинні автомати, моделювання біологічних процесів, САД-система

Keywords— movable cellular automata; simulation biological proces; CAD-system

I. ВСТУП

Вивчення процесів, що відбуваються в живих організмах та моделювання відповідних штучних біо-систем набуває все більшого значення на сучасному етапі розвитку інформаційних технологій та при стрімкому зростанні обчислювальних потужностей цифрової техніки. Серед комп'ютерних моделей біо-процесів та систем можна виділити два основні типи – кількісні моделі та якісні. Якісні моделі за допомогою комп'ютерної графіки дають змогу розкривати невідомі

досі властивості складної системи: її структуру, динаміку розвитку, стійкість, цілісність тощо. Кількісні моделі мають переважно характер прогнозу майбутніх чи пояснення минулих значень змінних, що характеризують систему [1,2]. Слід зауважити, що кожна із таких моделей характеризується окремими підходом та методом до розробки. Таким чином, актуальним є створення універсального зручного програмного засобу, що надавав би дослідникам у галузі біоінженерії можливість візуалізації динаміки формування найрізноманітніших біо-подібних процесів та структур.

II. ПОПЕРЕДНІ ДОСЛІДЖЕННЯ

На сьогоднішній день методи моделювання біологічних процесів і систем широко розвиваються, та потребують більш складних обчислень і традиційні підходи стають менш ефективними. В цій області досліджуються нові способи вдосконалити класичну пряму імітацію шляхом різних підходів до еволюції та розвитку штучних організмів (непряме кодування).

інтегрує декілька математичних моделей. До них відноситься модель Cells Potts, яка може моделювати кластеризацію клітин, ріст, поділ, смерть, зчеплення та обмеження об'єму та площі поверхні; а також модуль знаходження розв'язків диференціальних рівнянь з частинними похідними для моделювання розподілу зовнішніх хімічних полів та диференціації типів клітин.

CompuCell3D надає можливість експериментувати та тестувати біологічні моделі, забезпечуючи гнучкий пакет із багатьма різними рівнями керування, що може бути розширений при потребі.

В роботах [4,5] використовується 2-мірна реалізація, хоча авторами розроблено і 3-мірну версію моделі. Пухлина у моделі подана як клітинна пласка область, що дозволяє вивчати глобальну еволюційну динаміку значно більшої кількості ракових клітин, покращуючи обчислювальні можливості. Дослідниками схожих агентоподібних моделей пухлинного росту було доведено, що еквівалентні результати можна отримати як для дво-, так і для три-мірних випадків.

Розробка Java 3D візуалізації ракових пухлин описана у [6]. При моделюванні було враховано, що форма пухлини залежить від її типу і оточуючого середовища (кровопостачання, кістка, тканина, внутрішній простір тіла тощо). Для опису пухлинних клітин було застосовано прості математичні моделі, що використовують елементарну геометрію для обчислення розташування вокселів. Це дозволило зменшити обчислювальні можливості та підвищити швидкість відображення тривимірної візуалізації пухлини. Візуалізовано 4 основні типи геометричної форми росту пухлини у сферичному, лінійному, плоскому та довільному випадках.

У роботі [7] запропоновано поняття двовимірного динамічного клітинного автомата (ДКА) з однорідною структурою, яке розширює можливості двовимірних КА і дозволяє застосовувати клітинно-автоматний підхід до завдань, пов'язаних з динамікою структури модельованого простору в двовимірному випадку. Також розроблена дворівнева ДКА-модель росту та поділу клітин ціанобактерії *Anabaena catenula*. В процесі дослідницької роботи розроблено програмний комплекс, що дозволяє користувачеві програмувати двовимірні статичні КА-моделі, двовимірні ДКА-моделі з однорідною структурою і тривимірні КА-моделі з динамічною неоднорідною структурою осередків.

Усі згадані вище моделі володіють як і перевагами так і недоліками. Основою для їх створення слугувала конкретна задача, а отже, і результат націлений на спеціалістів у конкретній області (медицина, генетика, тощо). У даній роботі пропонується розробка універсальної комп'ютерної системи, яка б дозволяла дослідникам візуалізувати модель не лише за допомогою набору вбудованих функцій взаємодій, а й створення власних.

III. ФУНКЦІОНАЛЬНЕ ПРЕДСТАВЛЕННЯ САД-СИСТЕМИ

Універсальність та зручність програмного засобу моделювання обумовлює необхідність його побудови на

основі досить простих та універсальних методів. Самоорганізація в природних системах (фізична, біологічна, екологічна, соціальна) часто опирається на міради ідентичних агентів і, по суті, стохастичну динаміку. Нетривіальні шаблони і колективна поведінка можуть виникати з відносно простих правил для окремих агентів – факт, який характеризує складні системи. Складність вивчення цих механізмів полягає в тому, що в лабораторних умовах майже неможливо відстежити динаміку окремих частинок, які входять до складу цих механізмів. Також слід врахувати зворотній зв'язок між цими самоорганізаційними механізмами, і, безпосередньо, самими процесами, що відбуваються в організмі (наприклад, рух, ріст, поділ клітини, та ін.).

Асинхронні стохастичні клітинно-автоматні моделі є найбільш вживаними при моделюванні біологічних систем, оскільки відображають не детермінованість, а стохастичність більшості природних процесів. Тому, для дослідження пов'язаного з динамікою структури модельованого простору, обрано метод рухомих клітинних автоматів (РКА), що дозволяє використовувати концепцію сусідства та описувати процеси, що відбуваються у біо-структурах, набором функцій взаємодій між сусідніми агентами (клітинами).

Органічні молекулярні комплекси методом РКА можуть бути змодельовані аналогічно методам молекулярної динаміки, одноклітинні організми можуть бути представлені також на псевдомолекулярному рівні (цитоскелет у вигляді ланцюгів та мереж з РКА, мембрана як множина РКА у вигляді шару із різноманітними включеннями тощо), а багатоклітинні мікроорганізми природним чином слід представляти множинами РКА, що імітують відповідні клітини та їх динаміку.

У роботі [8] описано алгоритм визначення сусідніх клітин, який надалі використовується при моделюванні. Він дає можливість ефективно обчислювати сусідство на двовимірному неоднорідному масиві клітин і не вимагає значних витрат обчислювальних ресурсів. Основною перевагою такого підходу є організований спеціальним чином масив, який дозволяє частково змінювати структуру модельованої поверхні, не перебудовуючи при цьому всю поверхню цілком.

Основним модулем, запропонованого програмного засобу є модуль, який містить класи, що визначають базові структури даних для представлення моделей на основі рухомих клітинних автоматів (РКА-моделей) і набір основних алгоритмів їх взаємодії. Модель є конкретною реалізацією РКА-моделі та правил взаємодії. Вона складається із базових конструкцій і алгоритмів, описаних в Основному модулі. Далі пропонується опис методу та підходу для побудови моделей та алгоритмів основного модуля.

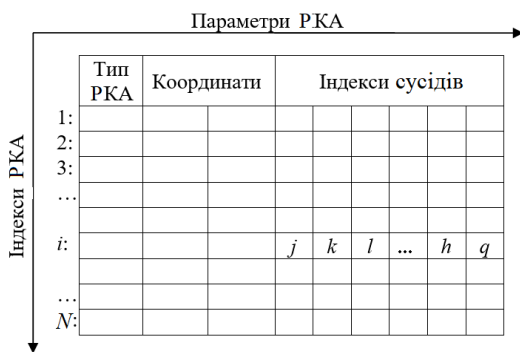
Як відомо, основна відмінність РКА від класичних клітинних автоматів полягає в здатності РКА приймати довільні неперервні значення координат в просторі. В зв'язку з цим виникає необхідність виділення додаткових параметрів у структурах даних, що зберігають інформацію про стан всього комплексу РКА. Ми використовували динамічний індексний масив, що дозволяє не тільки

ідентифікувати будь-який РКА у системі, але й містить індекси його найближчих сусідів (рис.1).

Основна суть роботи алгоритму – ітераційна модифікація вмісту індексного масиву. Для цього було реалізовано асинхронний підхід, при якому рівномірнішим чином був вибраний один з N РКА і один з M його найближчих сусідів:

```
for i = random (1 ... N)
j = MCA [i, P - random (1 ... M) + 1]
endfor
```

Тут i – індекс обраного РКА; j – індекс його сусіда; МСА (Movable Cellular Automata) – масив, що зберігає індекси сусідів. За індексами двох обраних РКА визначаються їх типи: $a_i^s = \text{MSA}[i; s]$; $a_j^s = \text{MSA}[j; s]$, де s – тип (стан) i -го j -го РКА.



N ; P] де N – кількість РКА, P – кількість параметрів

Вибір асинхронного підходу при розробці алгоритму РКА обумовлений тим, що дозволяє уникнути колізій, тобто задовольнити критерій коректності (не буде жодної спроби змінити стан однієї і тієї ж клітини більш ніж один раз в один і той самий момент часу t).

Клітинно-автоматна модель допускає k можливих станів рухомих клітинних автоматів, які описуються за допомогою алфавіту $A = \{a_1; a_2; \dots; a_k\}$ і мають різне призначення і функціонал, також для диференціації призначення можна використовувати різний радіус сфер, що є геометричним представленням РКА.

Описана типізація наслідок прагнення максимально-можливого спрощення моделі, виділення в ній найсуттєвішого. Після вибору двох сусідніх РКА реалізується функція взаємодій F . Ця функція залежить від типів взаємодіючих РКА. При цьому в деяких випадках, взаємодії здійснюються за участю більше двох РКА. Кожна з функцій F композиція окремих елементарних операцій, що здійснюються при взаємодіях відповідних типів. Призначення функцій можуть бути різноманітними. Наприклад: імітація теплових коливань; відштовхування двох РКА при їх взаємопроникненні; відштовхування при зближенні і притягування при видаленні (наприклад при моделюванні червякоподібного руху[9]); вирівнювання РКА (наприклад для моделювання пружних властивостей ниток цитоскелету при моделюванні амебоподібного руху); перетворення РКА

одного типу у РКА іншого типу; утворення нових РКА. При цьому слід також враховувати і певні обмеження, що виникають із властивостей модельованого середовища, наприклад обмеження на кількість сусідів, тощо.

ВИСНОВКИ

Розглядається сучасний стан проблеми інформаційних технологій в біоінженерії. Наводяться приклади створення відповідних математичних і комп'ютерних моделей та програм в різних країнах. Подано опис програмного засобу для комп'ютерного моделювання динаміки біологічних процесів та систем.

Розробка програмного засобу комп'ютерного моделювання дозволить користувачу досліджувати динамічні біологічні процеси та структури. Більш чітке розуміння механізмів, які керують розвитком біологічних систем, сприятиме створенню нових ефективних вакцин в боротьбі з бактеріальними інфекціями, пухлинами та іншими хворобами, покращить продуктивність при використанні біологічного матеріалу в сфері діяльності людини і може дати багато ідей та архітектурних рішень в кібернетиці і робототехніці. Допомагає на практиці перевірити гіпотези та теоретичні припущення, які допомагають у вивченні різних біологічних процесів. Також комп'ютерна модель дає змогу виявити нові явища в поведінці досліджуваних систем.

REFERENCES ЛІТЕРАТУРА

- [1] J. F. Li, J. Lowengrub. The effects of cell compressibility, motility and contact inhibition on the growth of tumor cell clusters using the Cellular Potts Model. *Journal of Theoretical Biology*, Vol. 343, 2014., pp. 79–91.
- [2] M. Scianna, L. Preziosi, L. Preziosi; Ed. by M. Chaplain, J. Batzel, M. Bachar. *New Challenges for Cancer Systems Biomedicine. Hybrid cellular Potts model for solid tumor growth*, Springer, 2012, pp. 205–224.
- [3] Multicompartment cell-based modeling of confined migration: regulation by cell intrinsic and extrinsic factors - Sandeep Kumar, Alakesh Das and Shamik Sen, *Molecular Biology of the Cell*, DOI:10.1091/mbc.E17-05-0313, (2018).
- [4] Sottoriva A., Verhoeff J.J., Borovski T., McWeeney S.K., Naumov L., et al. (2010) Cancer stem cell tumor model reveals invasive morphology and increased phenotypical heterogeneity. *Cancer Res* 70: 46–56.
- [5] J Anderson A.R., Weaver A.M., Cummings P.T., Quaranta V. (2006) Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell* 127: 905–915.
- [6] Chen Yin Jie, Daniel C. Doolan, Sabin Tabirca. *Geometrically based 3D Cancerous Cell Growth. Seventh Irish Workshop on Computer Graphics (2006) Eurographics Irish Chapter*.
- [7] Vitvitsky A. A. *Razrabotka i issledovaniye metodov kletochno-avtomatnogo modelirovaniya rosta i deleniya kletok zhivyykh organizmov. Diss. na soiskaniye uchenoy stepeni kandidata fiziko-matematicheskikh nauk po spetsialnosti 05.13.18 – matematicheskoye modelirovaniye. chislennyye metody i kompleksy programm*.
- [8] V.V. Zhykharevych, K.P. Hazdiuk, Alhorytm vyznachennia susidnikh elementiv mnozhyny rukhomykh klitynykh avtomativ za umov fiksovanoi kilkosti susidiv. *Visnyk Natsionalnoho tekhnichnoho universytetu "Kharkivskiy politekhnichnyi instytut". Zbirnyk naukovykh prats. Seriya: Informatyka ta modeliuвання, Kharkiv: NTU "KhPI", № 33 (1142), 2015, s.75-82.*
- [9] V. Zhikharevich, K. Gazdiuk, S. Ostapov, "Simulation of the dynamics of worm-like organisms using the movable cellular automata", VI-th International Conference of Students, PhD-Students and Young Scientists "Engineer of XXI Century", 02 Dec. 2016, pp. 469-475

The Expert System for Monitoring and Preventing Natural Emergency Situations on Automobile Highways

Leonid Nefedov
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
nefedovli@i.ua

Kononuhin Aleksandr
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
makonon@i.ua

Fil Nataliya
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
fil_nu@i.ua

Binkovskaya Angela
Automation and Computer-Integrated Technologies
Kharkiv National Automobile and Highway University
Kharkiv, Ukraine
ab.binkovskaya@gmail.com

Експертна Система Моніторингу і Попередження Природних Надзвичайних Ситуацій на Автомобільних Дорогах

Леонід Нефьодов,
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
*Харківський національний автомобільно-дорожній
університет*
Харків, Україна
nefedovli@i.ua

Кононіхін Олександр
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
*Харківський національний автомобільно-дорожній
університет*
Харків, Україна
makonon@i.ua

Філь Наталія
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
*Харківський національний автомобільно-дорожній
університет*
Харків, Україна
fil_nu@i.ua

Бінковська Анжела
*кафедра автоматизації та комп'ютерно-інтегрованих
технологій*
*Харківський національний автомобільно-дорожній
університет*
Харків, Україна
ab.binkovskaya@gmail.com

Abstract—*The article describes an expert system for monitoring and preventing natural emergency situations on automobile roads. A method for generating fuzzy SQL queries, based on standard language constructs, is developed. An example of constructing a fuzzy logic based selection of projects on preventing natural emergency situations on automobile roads is considered. The direction for further work is related to the development of an intelligent interface for the expert system for monitoring and preventing natural emergency situations on automobile highways.*

Keywords—*natural emergency situations, fuzzy logic, fuzzy set, membership function*

I. INTRODUCTION

Modern automobile highways (AHs) are the driver of development of the Ukrainian economy. To increase the volume of transportation and improve the investment climate, it is necessary to create a modern network of AHs.

However, the condition of a vast majority of AHs with respect to their use for traffic is unsatisfactory. Moreover, 97% of all AHs are completely worn down and require capital repair or maintenance. Ukraine ranks 132nd in the

world in terms of the level of development of AHs. In addition, more than 9.6 thousand bridges in terms of technical parameters do not meet the requirements of modern standards and actual load, 1923 bridges require urgent repairs, and 86 bridges are in unsafe condition. Such a position of public roads and bridges does not ensure to the full extent a speedy, comfortable, economical, and safe transport of passengers and freight, development of transit traffic; creates social tension; reduces the competitiveness of the Ukrainian economy; hinders the further socio-economic development of Ukraine and its integration into the European community. To solve the problem, the Government of Ukraine approved the Concept of the state target economic program for the development of public roads for 2018-2022 [1].

In the last decade, there observed a steady trend of increasing anthropogenic pressures on the environment, land degradation, global climate change. All these causes activation and intensive development of dangerous natural processes and phenomena, increases the vulnerability of the population and territories to natural emergency situations (NES). The insufficient efficiency and development of

systems for monitoring potentially dangerous natural processes and phenomena, absence or poor condition of protective engineering structures are also important factors that affect material losses, and in some cases also human ones [2].

The implementation of preventive measures, projects, programs and strategies for protecting the population and territories reduces potentially negative consequences of emergency situations. The strategy for preventing NESs is aimed at solving the following tasks: prevention of causes of NESs; prevention of NESs and maximum mitigation of their consequences. The unstable economic situation in Ukraine makes it difficult to fully protect the population and territories from NESs. In addition, all existing projects on preventing NESs are very expensive and require significant financial, material, technical and human resources.

Particular attention should be paid to monitoring and preventing NESs. The problem of monitoring and predicting potentially dangerous sources of possible NESs can be solved by conducting multi-criteria analysis of various situations and making scientifically based effective decisions.

The main tasks of the system for monitoring the threat of NESs on AHs are [2]: conducting observations of sources of NESs and AHs; collecting, processing and analyzing information on sources of NESs and AHs; creating a data bank on sources of NESs and condition of AHs; predicting NESs on AHs; providing government bodies and other state organizations subordinated to the Government of Ukraine, local executive and regulatory bodies with information on the threat or occurrence of NESs on AHs [2].

II. PROBLEM STATEMENT

Based on the monitoring and expertise of sources of NESs on AHs and condition of the roads, there identified the AR sections and roadside facilities that require implementation of preventive measures and works considered as a set of possible projects on preventing NESs on AHs.

All of these projects define a strategy for preventing NESs on AHs. The strategy for preventing NESs is aimed at creating sustainable development conditions and ensuring a significant mitigation of risk and decrease of losses from NESs.

Due to the lack of financial resources for the implementation of the whole range of measures to minimize the risk of occurrence of NESs, the task of optimal distribution of the allocated financial, material, technical and human resources among all possible projects on preventing NESs on AHs comes to the fore.

For this purpose, the evaluation and ranking of possible projects on preventing NESs on AHs is performed, with regard to multi-criteria analysis of all possible projects on preventing NESs on AHs, based on expert information, i.e., their advantages and disadvantages are considered from the system point of view. The evaluation of possible projects on preventing NESs on AHs can be carried out according to the following criteria: frequency of occurrence of NESs on the given section of the AR; damage in the event of NES on the given section of the AR; capital costs for the implementation

of works to prevent NESs on the given section of the AR, etc.

For carrying out multi-criteria evaluation of possible projects on preventing NESs on AHs, it is suggested to use the hierarchy analysis method, which, based on expert data, allows ranking projects on preventing NES by their significance based on the given criteria.

The prediction and prevention of NESs is currently being implemented largely on the basis of poorly formalizable and completely non-formalizable knowledge, which are the result of years of observations, experience and intuition of various specialists [3]. Under these conditions, expert systems (ESs) are an important tool to solve problems of modeling and predicting dangerous events.

ESs for monitoring and preventing NESs on AHs are sophisticated software packages intended for collecting, processing and generalizing poorly formalizable knowledge of experts. Therefore, when developing a knowledge base for ESs, it is necessary to use the fuzzy logic technique that will allow formalizing the procedure for evaluating sources of NESs on AHs and make scientifically based decisions to prevent NESs on AHs. A fuzzy ES for monitoring and preventing NESs on AHs uses knowledge representation in the form of fuzzy products and linguistic variables [3].

Figure 1 presents a fuzzy logic structural model of an ES for monitoring and preventing NESs on ARs [3].

III. THE METHOD FOR FUZZY SELECTION OF PROJECTS ON PREVENTING NESs ON AHs

One of the elements of the system under consideration is the knowledge base. It is necessary to develop means to form a strategy, which contains projects on preventing NESs on AHs, based on fuzzy logic inference in the relational database. Fuzzy logic search and inference is based on the theory of fuzzy sets and is implemented through fuzzy SQL queries [4]–[7].

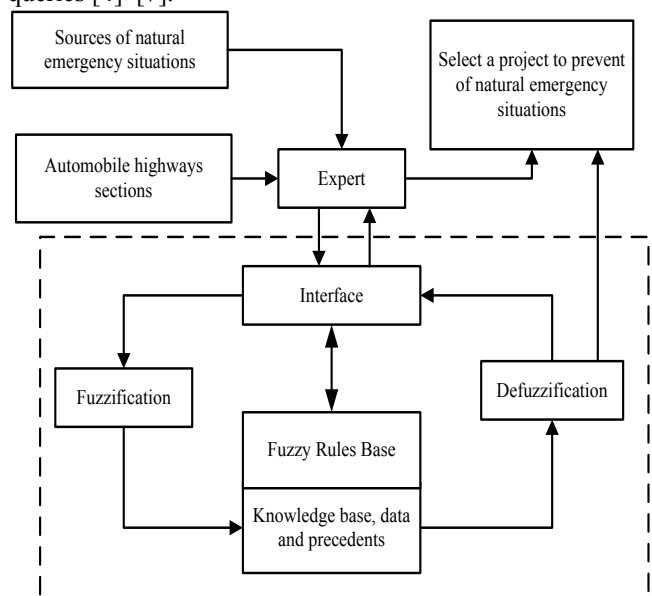


Fig. 1 The fuzzy logic structural model of an ES for monitoring and preventing NESs on ARs

The final decision-making takes into account the whole set of criteria for the implementation of measures, even, at

first glance, insignificant ones, which makes the decision most suitable to the reality.

After the synthesis of resulting (generalized) estimates and the ranking of projects, it is necessary to formulate a strategy for preventing NESs on AHs for the future, representing the portfolio of projects.

Based on the received portfolio of projects on preventing NESs on AHs, it is necessary to form a portfolio of projects on preventing NESs for the current period, which requires more detailed and accurate information.

In work [2] the information and reference support for the technique for monitoring and preventing NESs on AHs was developed. Information on sources of possible NESs, operating conditions is stored in relational databases. There is also a database of possible projects on preventing NESs on AHs. When developing a strategy for preventing NESs on AHs and a program for preventing NESs on AHs, it is necessary to select from the database of possible projects those that have the best value/cost ratio. This choice is supposed to be carried out using fuzzy queries.

The method of fuzzy data selection is based on the theory of fuzzy sets, which was first proposed by Lotfi Zadeh [9]-[11]. To implement fuzzy queries, it is necessary to formally describe the sets of values of terms in a logical expression. The model of such structures is the linguistic variable (LV) corresponding to the used fuzzy concepts of “low cost”, “high value” [4]. In this case, there should be defined universal sets of these variables, sets of basic terms, and membership functions for each term, taking into account the requirements imposed on them [4].

Let us consider an example of a fuzzy selection of projects on preventing NESs on AHs. Fuzzy data selection in the relational database is performed through SQL queries. All information about possible projects on preventing NESs on AHs is stored in the database in the table TProject (Tbl. 1), which has the following columns: id – the number of the projects on preventing NESs on.

AHs, TWorth – the value of the projects on preventing NESs on AHs, TValue - cost (UAH ths).

TABLE I. INFORMATION ON PROJECTS ON PREVENTING NESs ON AHs

id	TWorth	TValue
1	0,20	60
2	0,17	35
3	0,16	50
4	0,11	18
5	0,1	45
6	0,07	18
7	0,06	30
8	0,05	6
9	0,04	25
10	0,03	25

To implement fuzzy queries, we formalize two linguistic variables: “Cost” and “Value”. The linguistic variable “Value” has the following form:

To implement fuzzy queries, we formalize two linguistic variables: “Cost” and “Value”. The linguistic variable “Value” has the following form:

$$\langle \beta, \Psi, T, G, M \rangle, \quad (1)$$

where β - “Value”; $\Psi = [0.03; 0.20]$ - the scope of definition of the linguistic variable; $T = \{“high”, “medium”, “low”\}$; G - the syntactic procedure for the formation of new terms with the help of the connectives “AND”, “OR” and the modifiers “NO”, “VERY”, “MORE OR LESS”; M - the procedure for setting on the set Ψ the fuzzy sets $A = “low”, B = “medium”, C = “high”,$ and fuzzy sets for terms according to the rules for translating fuzzy connectives and modifiers and other operations on fuzzy sets.

Let us define trapezoidal membership functions for the variable “Value”:

$$\mu_A(x) = \begin{cases} 1, & x \in [0, 0,03] \\ (0,06 - x) / 0.03, & x \in (0,03, 0,06] \\ 0, & x > 0,06 \end{cases} \quad (2)$$

$$\mu_B(x) = \begin{cases} 0, & x < 0,05; x > 0,12 \\ (x - 0,05) / 0.03, & x \in [0,05, 0,08] \\ 1, & x \in [0,08, 0,1] \\ (0,12 - x) / 0.02, & x \in (0,1, 0,12] \end{cases} \quad (3)$$

$$\mu_C(x) = \begin{cases} 0, & x < 0,11 \\ (x - 0,11) / 0.04, & x \in [0,11, 0,15] \\ 1, & x \in [0,15, 0,2] \end{cases} \quad (4)$$

The graphs of the membership functions are shown in Figure 2.

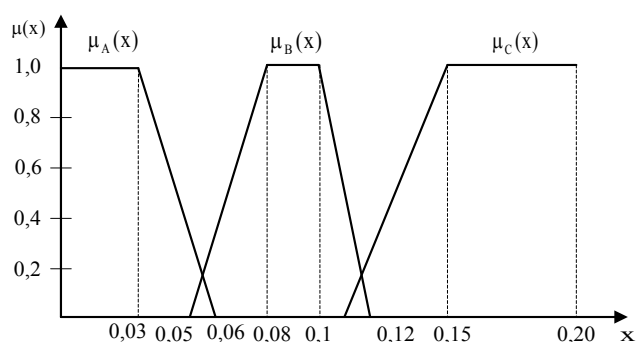


Fig. 2. The membership function of terms of the variable “Value

The linguistic variable “Cost” has the form:

$$\langle \gamma, \Gamma, T, G, M \rangle, \quad (5)$$

where γ - “Cost”; $\Gamma = [6; 60]$; $T = \{“low”, “medium”, “high”\}$; G - the syntactic procedure for the formation of new terms with the help of the connectives “AND”, “OR” and the modifiers “NO”, “VERY”, “MORE OR LESS”; M - the procedure for setting on the set Γ the fuzzy sets $D = “low”, E = “medium”, F = “high”,$ and fuzzy sets for terms $G(T)$ according to the rules for translating fuzzy connectives and modifiers and other operations on fuzzy sets.

Let us define trapezoidal membership functions for the variable “Cost”:

$$\mu_D(y) = \begin{cases} 1, & y \in [6;15] \\ (20 - y)/5, & y \in (15;20] \\ 0, & x > 20 \end{cases} \quad (6)$$

$$\mu_E(y) = \begin{cases} 0, & y < 18; y > 40 \\ (y-18)/4, & y \in [18;22]; \\ 1, & y \in [22;35] \\ (40 - y)/5, & y \in (35;40] \end{cases} \quad (7)$$

$$\mu_F(y) = \begin{cases} 0, & y < 38 \\ (y-38)/7, & y \in [38;45]; \\ 1, & y \in [45;60] \end{cases} \quad (8)$$

The graphs of the membership functions are shown in Figure 3.

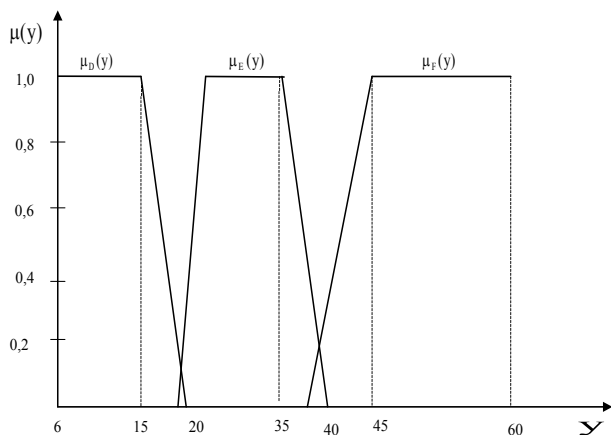


Figure 3 - The membership function of terms of the variable “Cost”

Table 2 presents the projects and the calculated degrees of membership.

TABLE II.
PROJECTS AND THE CALCULATED DEGREES OF MEMBERSHIP

id	TWorth	TValue	TWorth			TValue		
			μ_A	μ_B	μ_C	μ_D	μ_E	μ_F
1	0,20	60	0	0	1	0	0	1
2	0,17	35	0	0	1	0	1	0
3	0,16	50	0	0	1	0	0	1
4	0,11	18	0	0,5	0	0,4	0	0
5	0,1	45	0	1,0	0	0	0	1
6	0,07	18	0	0,67	0	0,4	0	0
7	0,06	30	0	0,33	0	0	1	0
8	0,05	6	0,33	0	0	1,0	0	0
9	0,04	25	0,67	0	0	0	1,0	0
10	0,03	25	1,00	0	0	0	1,0	0

Let us consider a fuzzy selection in the form of the SQL query “to find all projects of medium value and low cost”:

```
select * from TProject where (TWorth = "средняя" and
TValue = "низкая");
```

The fuzzy selection result is presented in Tabl. 3.

In this query, the two parameters are determined by the degree of membership and the operator “and” is defined as the Cartesian product of fuzzy sets.

Fuzzy queries can be used effectively to select information from relational databases using qualitative criteria and fuzzy conditions/

TABLE III THE RESULT OF A FUZZY SELECTION OF PROJECTS ON PREVENTING NESS ON AHs

id	TWorth	TValue	TWorth			TValue		
			μ_A	μ_B	μ_C	μ_D	μ_E	μ_F
4	0,11	18	0	0,5	0	0,4	0	0
6	0,07	18	0	0,67	0	0,4	0	0

CONCLUSION

In the article, a structural model of the expert system for monitoring and preventing natural emergency situations on automobile highways is developed, which, unlike known ones, is based on fuzzy logic. The method for fuzzy selection of projects in the expert system for monitoring and preventing natural emergency situations on automobile highways is elaborated, which, unlike known ones, is based on the value/cost fuzzy evaluation of possible projects on preventing natural emergency situations on automobile highways. This allows to adequately assess sources of natural emergency situations and condition of automobile highways and on this basis choose effective projects. An example of constructing a fuzzy query for the selection of projects on preventing natural emergency situations on automobile highways is considered. The processing of fuzzy queries results in an intelligent search, when the exact request attributes are fuzzy or undefined. This significantly simplifies the work with large volumes of various data for experts. The direction of further work is related to the development of an intelligent interface for the expert system.

REFERENCES ЛІТЕРАТУРА

- [1] Pro shkvalennia Kontseptsii Derzhavnoi tsilovoi ekonomichnoi prohramy rozvytku avtomobilnykh dorih zahalnoho korystuvannia derzhavnoho znachennia na 2018-2022 roky. Available at: <http://zakon3.rada.gov.ua/laws/show/34-2018-p>. Accessed: 2 March 2018.
- [2] Nefedov, L. I., Fil', N. Yu., Gubin, Yu. L., Mel'nichenko Ye. M. (2011). Modeli i metody upravleniya chrezvychnymi situatsiyami na magistral'nykh avtomobil'nykh dorogach. KhNADU, 136.
- [3] Drobnohod, N. I. (2001). Ustoychivoye ekologicheskii bezopasnoye razvitiye: ukrainskiy kontekst. *Zerkalo nedeli*, 21, 15. Accessed: 2 March 2018.
- [4] Shtovba, S. D. Vvedeniye v teoriyu nechetkikh mnozhestv i nechetkuyu logiku. [Introduction to the theory of fuzzy sets and fuzzy logic.] Available at: <http://matlab.exponenta.ru/fuzzylogic/book1>. Accessed: 2 March 2018.
- [5] Paklin, N. B. Nchetkiye zaprosy k relyatsionnym bazam dannykh. [Fuzzy demands to relational databases.] Available at: <http://www.basegroup.ru>. Accessed: 2 March 2018.
- [6] Chuykova, E. N., Molchanov, A. A. Formirovaniye nechetkikh zaprosov k baze dannykh. [Assignment of fuzzy demands to the database.] Sbornik konkursnykh rabot Vserossiyskogo smotrakonkursa nauchno-tekhnikeskogo tvorchestva studentov vysshikh uchebnykh zavedeniy «Evrka-2007». [Collection of entries of the All-Russian competition of sci and tech creativity of university students “Eureka-2007”.] Novocherkassk : Oniks+, 2007, pp. 38–40.
- [7] Moiseyenko, S. I., Levchenko, V. A. Translyatsiya operatorov relyatsionnoy algebry v SQLoperatoru. [Translation of relation algebra statements to SQL – statements.] *Vestnik of DSTU*, 2005, vol. 1, no. 1, pp. 691–696.
- [8] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning - I. *Information Sciences*, 1975, vol. 8, iss. 3, pp. 199–249.
- [9] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning - II. *Information Sciences*, 1975, vol. 8, iss. 4, pp. 301–357.
- [10] Zadeh, L. A. The concept of a linguistic variable and its application to approximate reasoning - III. *Information Sciences*, 1975, vol. 9, iss. 1, pp. 43–80.

Automated System for Light Therapy with Frequency Scanning Irradiated Areas

Aleksandr Krushets
Radio Engineering Faculty
NTUU "Igor Sikorsky Kyiv Polytechnic
Institute
Kyiv, Ukraine
aleksandrkrushets@gmail.com

Aleksey Yanenko
Radio Engineering Faculty
NTUU "Igor Sikorsky Kyiv
Polytechnic Institute"
Kyiv, Ukraine
op291@meta.ua

Автоматизована Система для Світлотерапії з Частотним Скануванням Ділянки Опромінення

Олександр Крушець
Радіотехнічний факультет
НТУУ "Київський політехнічний
інститут ім. І. Сікорського"
Київ, Україна
aleksandrkrushets@gmail.com

Василь Кузь
кафедра Електричної інженерії
ТНТУ Тернопільський національний
технічний університет ім. І. Пулюя
Тернопіль, Україна
vasylkuz1992@gmail.com

Олексій Яненко
Радіотехнічний факультет
НТУУ "Київський політехнічний
інститут ім. І. Сікорського"
Київ, Україна
op291@meta.ua

Abstract— In the work the author proposed an automated microprocessor system for light therapy with frequency scanning of the irradiated areas, which provides determination frequencies of maximum absorption, which provides an increase in the effectiveness of treatment. The structural scheme of the system is given and the principle of its work is described as well as experimental patterns of frequency scanning of some areas of the human body.

Keywords—light therapy, microprocessor system, frequency scanning.

Анотація—В роботі розглянуто запропоновану авторами автоматизовану мікропроцесорну систему для світлотерапії з частотним скануванням ділянки опромінення, яка забезпечує визначення частот максимального поглинання, що забезпечує підвищення ефективності лікування. Приведена структурна схема системи та описано принцип її роботи а також експериментальні графіки частотного сканування деяких ділянок тіла людини.

Ключові слова—світлотерапія, мікропроцесорна система, частотне сканування.

I. INTRODUCTION

In our time, the use of light therapy is quite widespread. This therapy is based on the influence of light energy of the corresponding intensity (usually low intensity) and the wavelength on certain areas of the patient's body [1]. These can be active zones or biologically active points of the human body.

An important role in this case is to choose the frequency of modulation of the light flux, which according to [2] is within the range of 0.1-100 Hz. The specified range is biologically active with maximum sensitivity for brain tissue in the range of 6-20 Hz. In addition, pulsed modulation leads to clear biological effects, which are confirmed

experimentally and can reduce the level of electromagnetic effects while preserving its effectiveness.

The maximum effect of light therapy arises from the largest (or resonant) absorption of the energy of the illuminating signal. In [3] variants of matrix radiation of the selected body region at a fixed frequency are considered without determining the absorption level. The authors analyzed the main photobiological methods of the influence of optical radiation of the ultraviolet range of waves (250-400 nm), which are used to irradiate the affected skin. Since the ultraviolet range, whose signals cause mainly superficial action, is considered, the task of the work was to achieve a uniform distribution of power, which was achieved through the use of LED dies.

A similar solution is used in the range of visible light in the matrices and apparatus "Barva" Korobova [4]. The use of light waves with longer wavelengths that penetrate more deeply into the biofilm requires a different approach.

The authors [5] developed a device that allows to determine the absorption, but because of the functional construction of the circuit of the device with switching modulation conversion of probes and received signals and the presence of sample elements, such measurement can be carried out only at a fixed modulation frequency.

The indicated disadvantages considerably limit the possibilities and effectiveness of these schemes of light therapy.

The authors of this work proposed a system for light therapy, which in an automated mode provides the possibility of scanning of modulation frequencies with the parallel definition of optimal (with maximum absorption) frequencies for therapy in the selected area of the body.

II. MAIN BODY

An automated system for light therapy, which provides the ability to automatically modulate the frequency of modulation in a range close to the human biorhythm frequencies with the simultaneous measurement of the level of absorbed light flux by the biological object, is presented in Fig. 1. The scheme includes elements for generating the light flux, its modulation, an optoelectronic transducer consisting of two lenses and a measuring sensor (photocell) and a control unit (microcontroller).

The measurement results are processed by a personal computer.

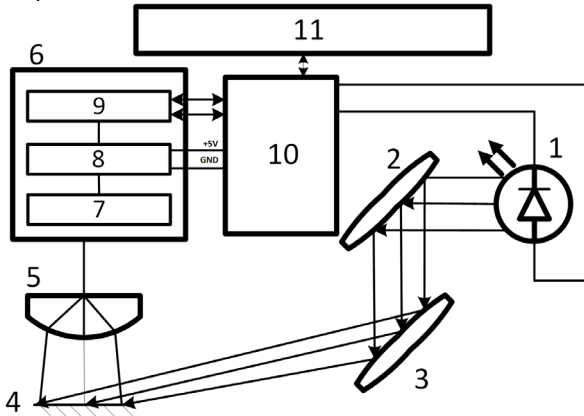


Fig. 1. Scheme of automated system for light therapy

Symbols on Fig. 1: 1 LED, the authors selected red (wavelength ~ 660 nm) because it provides penetration into the body of the human body to a depth of about 8-10 mm, 2.3 - a system of focusing lenses, 4 - a biological object, 5 - Cleaning lens, 6 - Optical sensor, 10 - Microcontroller, 11 - PC.

The scheme operates as follows. The supply voltage of red light diode 1 is modulated by human biorhythm frequencies, 1-100 Hz, using a sufficiently accessible microcontroller *Arduino Uno* 10. Modulated light flux through a system of focusing lenses 2, 3 is directed to a specific area or BAT of a biological object 4. Part of the light is absorbed by the bio-object, and the part is reflected.

The reflected light with the help of a collecting lens 5 enters the optical sensor 6. In turn, the optical sensor 6 with high frequency (kHz) carries surface images and independently treats them, since it contains three main blocks: the *Serial port* 9, the system of reception and image processing of *Image Processor* 7, consisting of a DSP processor and a built-in light signal receiver (IAS), and a *Voltage Regulator and Power Control* 8 power control system.

On the basis of the analysis of the received sequential images (representing a square matrix of pixels of different brightness in grayscale, from 0 - "complete black" to 63 - "complete white"), the integrated DSP processor 7 calculates the resulting values, that is, the value of the brightness of the entire matrix at certain moments time. Due to the possibility of programming the sensor, we obtain (expect) the average

values of the result indicators at the required flashing frequency of the LED.

Next, the transfer of the results of the DSP processor 7, through the microcontroller 10, to the personal computer 11 using serial port Serial Port 9. And as a result, at the output, we obtain the value of relative brightness (a measure indicating the amount absorbed by the biological object light) at each of the required frequencies within 1-100 Hz. That will determine the most effective frequencies for the patient's light therapy.

B. Sensor description

As an optical sensor, the AVAGO TECHNOLOGIES chip [6] *ADNS-2610* was used. It has a resolution of up to 400 cpi (*countsperinch*), and an image capture rate of 1500 frames per second.

Although the sensor and "sees" the image black and white, however, the sensitivity of the matrix depends on the wavelength. On Fig. 2 shows the characteristic of the selected sensor.

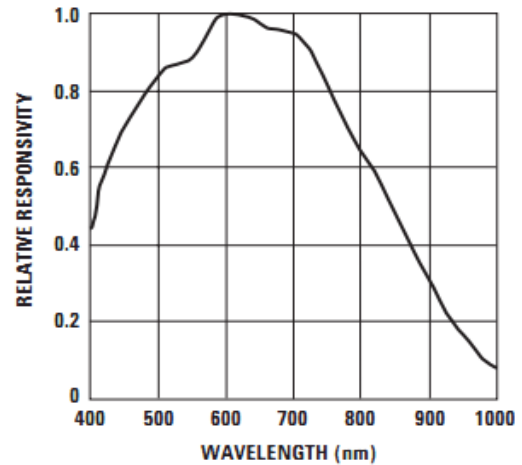


Fig. 2. Dependence of sensitivity of the sensor on the wavelength

As can be seen from the characteristic, the sensitivity of the greatest sensitivity is at wavelengths equal to 600-700 nm. Therefore, the LED with the largest wavelength (660 nm) was chosen. Because, when using the LEDs of another wave of a wave, it is necessary to carry out additional amplification of the reflected signals or to introduce a digital correction of the nonlinearity of the graph, since the sensitivity can decrease by 2-3 times.

The sensor summarizes the brightness of all 324 pixels in the current frame. Then programmed the calculation of the average brightness, which is made by the formula:

$$A = X * 128/324 \quad (1)$$

where: A- the average value of the brightness of the pixels, X - the current value of the register (multiplied by 128 because of the present number shift relative to the register on 2^7).

On Fig. 3 example image as the sensor sees the surface that it scans. It is a square matrix of pixels of different brightness in grayscale of 18×18 .

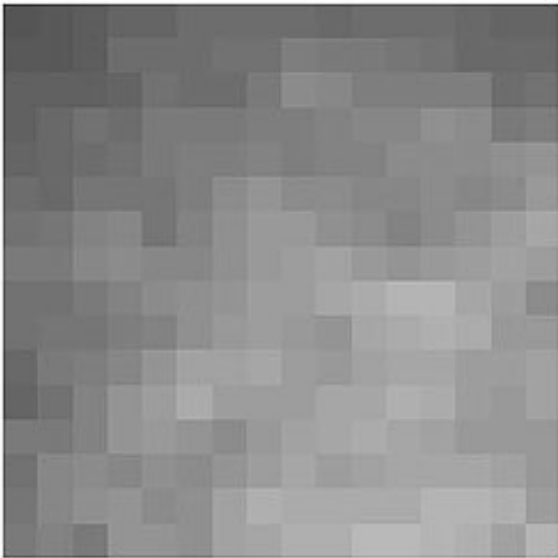


Fig. 3. An example image obtained by the sensor (matrix pixels 18x18)

III. ALGORITHM OF THE INFORMATION PROCESS PROGRAM

The program for managing and processing metering information is developed in the *Arduino IDE* environment. The program functions as follows

First, the Timer One library is introduced. This library allows you to use a 16-bit timer in the AVR microcontroller installed on the Arduino boards. It is used for PIM in the analog Write function. Also introduced functions that describe the operation of the sensor. Then global variables are announced to represent the array with frequencies that will integrate our measurements of the relative light absorption intensity and reserve a place under the two dimensional array for recording the measurement results.

In the main body of the program, the speed of the serial port is declared, and the timer initialization subroutine is started at intervals of 1 s (equal to 1 Hz) and the PWM signals are set to the LED output with a fill factor of 50%. A subroutine measuring 5 values of relative brightness at a given frequency is called, which is due to reading information from the sensor register at 0x07 Pixel_sum (which is the count average pixel value). After that, the first cycle begins, in which there is a change of frequency from 1 to 100 Hz, and each time a subroutine for measuring the values at each of the frequencies is started.

Next, a second cycle begins in which of the measured values at each frequency is the average value and their output is. The processing of the results of the five measurements increases the accuracy of the determination of the intensity on the frequency coordinate grid. Although it is possible to increase the number of measurements at each of the frequencies, but this is not a good idea, as it will greatly increase the time to obtain the resulting performance.

The algorithm of the program in the form of a block diagram is depicted in Fig. 4.

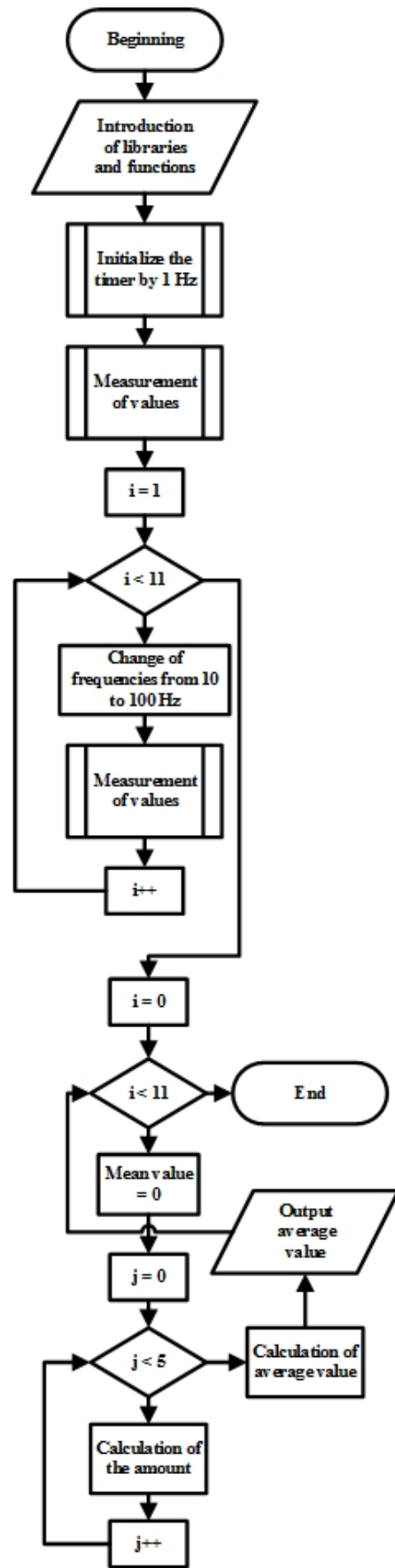


Fig. 4. The algorithm of the program

IV. EXPERIMENTAL STUDIES

A. Results of work

After the development of the scheme and layout of the device, experimental research was carried out on various biological objects. As a biological object, the reverse side of the palm is selected - Fig. 5, soft handkerchiefs - Fig. 6, and soft tissue of the thigh - Fig. 7. Experiments are carried out at a change in frequency from 1 to 100 Hz, which corresponds to most biological rhythms of the human body.

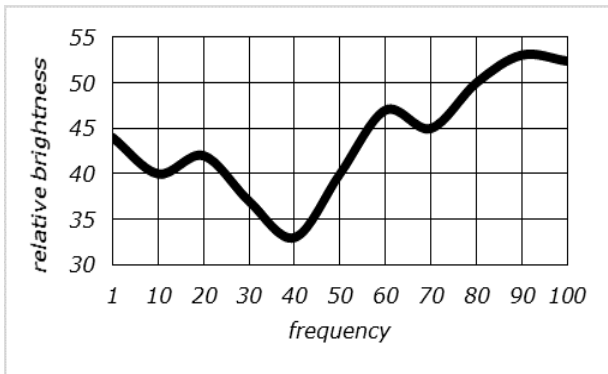


Fig. 5. Back side of the palm with a small amount of muscle tissue

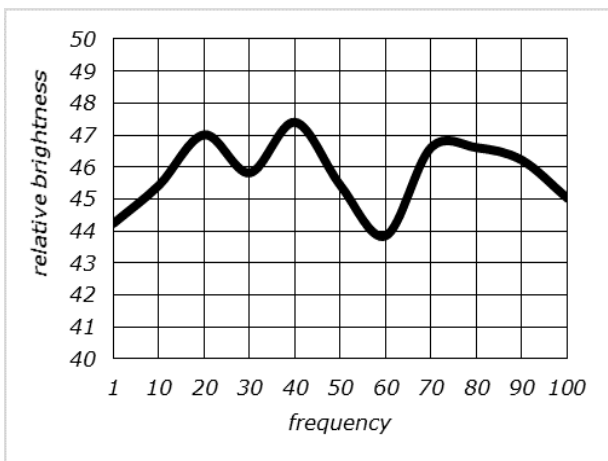


Fig. 6. Soft tissues of the hand

Fig. 7. Soft tissue hip

From the drawings it is seen that the greatest absorption capacity of human body biotransformation is within

35-65 Hz. Increasing the modulation frequency to 100 Hz results in a smooth growth of the reflecting component.

Examination of the absorption capacity of the swine fat tissue sample Fig. 8 showed that in addition to the first maximum absorption (resonance), there is still a similar absorption area in the frequency range of 75-85 Hz, but there is also a frequency modulation frequency range in the range of 60-75 Hz where the maximum reflection of the global flow occurs. Experimental measurements were performed on the body parts of a young healthy respondent without a fatty layer of skin.

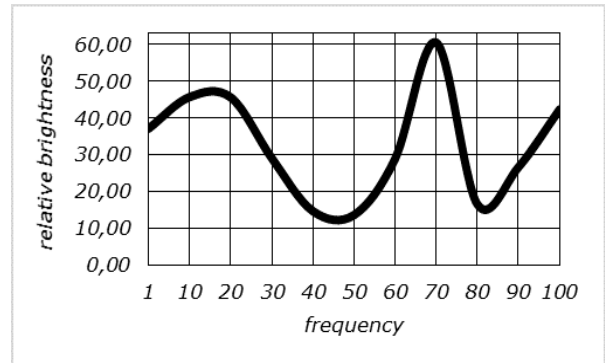


Fig. 8. Porcine fatty tissue

CONCLUSIONS

1) The developed system allows, in practically real-time, to determine the frequencies of maximum absorption of the selected area of the patient's body. The time taken for the measurement and construction of the absorption schedule is 4-5 minutes.

2) Experimental verification showed that the maximum absorption of the power of the light flux is at frequencies close to 30-60 Hz, which corresponds to the smallest value of the relative brightness (0 - completely black, 63 - completely white) on the sensor, since the reflected power at these frequencies is the smallest.

3) The presence of the fatty layer can lead to the appearance and shift of the maximal absorption frequencies.

REFERENCES ЛІТЕРАТУРА

- [1] Rojas J. C. Low-level light therapy of the eye and brain / J. C. Rojas, F. Gonzalez-Lima // Eye and Brain. – 2011. – No 3. – P. 49–67
- [2] Adey W. Frequency and Power Windowing in Tissue Interaction with Wear Electromagnetic Fields // IEEE - Vol.68.- 1/-1980 - p.140-147
- [3] Яненко О.П. Дослідження рівня опромінення пристрою світлотерапії/ О.П. Яненко, В.І. Кузь// Вісник КПІ, серія «Приладобудування»- вип.52(2). 2016- с.108-112
- [4] Коробов, А.М. Фототерапевтические аппараты Коробова серии «Барва» / А.М. Коробов, В.А. Коробов, Т.А. Лесная. – Харьков, ХНУ им. В.Н. Каразина. – 2008. – 175 с.
- [5] Патент на корисну модель, №91337, Україна., МПК А61Н 39/00, G1J 1/00, Пристрій для світло терапії / Яненко О.П., Михайленко С.В, винахідники, заявка № u201401890 від 25.02.2014; опубл. 25.06.2014, Бюл. № 12
- [6] ADNS-2610 (2008). Optical sensor. Available: <https://media.digikey.com/pdf/Data%20Sheets/Avago%20PDFs/ADNS-2610.pdf>. Last accessed 4th Sep 2018

Інтелектуальна Інформаційна Система Аналізу Вмісту Кошика Покупця Супермаркету

Іванна Сенейко
Національний університет
«Львівська політехніка»
Львів, Україна
ivannasenyko@gmail.com

Наталія Едуардівна Кунанець
Національний університет
«Львівська політехніка»
Львів, Україна

Антоній Валентинович Ржеуський
Національний університет
«Львівська політехніка»
Львів, Україна
antonii.v.rzheuskiy@lpnu.ua

Intelligent Information System of Content Analysis of Supermarket Shopping Cart

Ivanna Seneyko
Lviv Polytechnic National University
Lviv, Ukraine
ivannasenyko@gmail.com

Nataliia Kunanets
Lviv Polytechnic National University
Lviv, Ukraine
nek.lviv@gmail.com

Antonii Rzheuskiy
Lviv Polytechnic National University
Lviv, Ukraine
antonii.v.rzheuskiy@lpnu.ua

Анотація—Розглянуто особливості розроблення інтелектуальної інформаційної системи аналізу вмісту кошика покупця супермаркету.

Ключові слова— інтелектуальна інформаційна система, кошик покупця супермаркету, клієнт, асоціативні правила.

I. ВСТУП

Протягом останніх років основною тенденцією розвитку оптово-роздрібної торгівлі є створення великих торговельних підприємств. Споживачі почали надавати перевагу супермаркетам та гіпермаркетам.

Мета статті – проаналізувати функціональні характеристики інтелектуальної інформаційної системи аналізу вмісту кошика покупця супермаркету.

II. ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Великі торговельні підприємства характеризуються великою кількістю товарів, нижчим рівнем цін, вигідними маркетинговими пропозиціями, зручним розміщенням та наявністю автостоянок. Їх перевагою є: ефективно організований процес, наявність власних торговельних марок, просторі приміщення, широкий асортимент товарів, орієнтація на кожного споживача [1].

Супермаркет – це великий магазин самообслуговування, спеціалізацією якого є торгівля товарами повсякденного попиту (переважно продовольчими). Він є одним із сучасних форм концентрації торгівлі [2]. Внаслідок стрімкого розвитку ринкових відносин і посилення конкурентних процесів між торговельними мережами актуалізувалась необхідність вивчення і розуміння психології споживача та причин прийняття ним рішень щодо придбання товарів. У конкурентній економічній системі для виживання і розвитку керівництву підприємства необхідне точне розуміння поведінки споживача: як він

купує, чому купує, де купує, що саме купує і т. д. Підприємці вимушені розробляти все нові методи залучення нових клієнтів та утримання існуючих. Для збереження клієнтів необхідно враховувати їх інтереси. Але враховувати інтереси кожного клієнта є метою, яку важко досягнути [3]. Альтернативним засобом розв'язання цієї проблеми є застосування аналізу вмісту кошика покупця супермаркету. Даний підхід допомагає вивчати смаки клієнтів та особливості їх поведінки, що сприяє створенню відповідного асортименту, плану розташування товарів і проведенні акцій, які сприятимуть збільшенню як кількості покупців, так і прибутку. Метою аналізу ринкового кошика є виявлення асоціацій між різними подіями, тобто знаходження правил для кількісного опису взаємного зв'язку між двома або більше подіями. Такі правила називаються асоціативними правилами (association rules). Базовим поняттям у теорії асоціативних правил є транзакція – деяка множина подій, що відбуваються спільно. Типова транзакція – покупка клієнтом товару в супермаркеті. У переважній більшості випадків клієнт купує не один товар, а набір товарів, що називається ринковим кошиком. При цьому виникає питання: чи є покупка одного товару в кошику наслідком або причиною покупки іншого товару, тобто, чи пов'язані дані події? Цей зв'язок і встановлюють асоціативні правила. Наприклад, може бути виявлене асоціативне правило, котре стверджує, що клієнт, який купив молоко, з імовірністю 75 % купить і хліб. Для характеристики асоціативних правил використовуються такі показники:

- значення достовірності (confidence), яке показує ймовірність, з якою, в разі покупки предмета А також буде придбаний предмет Б;
- величина підтримки (support), визначається як відношення кількості транзакцій, в яких знаходиться предмет чи група предметів до загальної кількості усіх транзакцій.

Аналіз ринкового кошика – це аналіз наборів даних для певної комбінації товарів, пов'язаних між собою.

Іншими словами, виконується пошук товарів, присутність яких у транзакції впливає на ймовірність наявності інших товарів або комбінацій товарів. Сучасні касові апарати в супермаркетах дозволяють збирати інформацію про покупки, що може зберігатися в базі даних. Потім накопичені дані можуть використовуватися для побудови систем пошуку асоціативних правил [4]. Саме для автоматизованого аналізу цих даних розроблена система аналізу вмісту кошика покупця супермаркету (Рис. 1). Основними функціями системи є: пошук асоціативних правил, створення акційних пропозицій, ABC та XYZ аналіз. Для пошуку

асоціативних правил та популярних наборів було реалізовано алгоритм Apriori. Користувач може налаштувати параметри пошуку через інтерфейс (мінімальна та максимальна підтримка, мінімальна та максимальна достовірність, максимальна потужність шуканих множин, які часто зустрічаються). Створення вигідних акційних пропозицій та пошук асоціативних правил здійснюється на основі знайдених популярних наборів. За допомогою ABC-аналізу можна аналізувати внесок товару в результат роботи підприємства, а XYZ-аналіз дає інформацію про стабільність продажів певного товару.

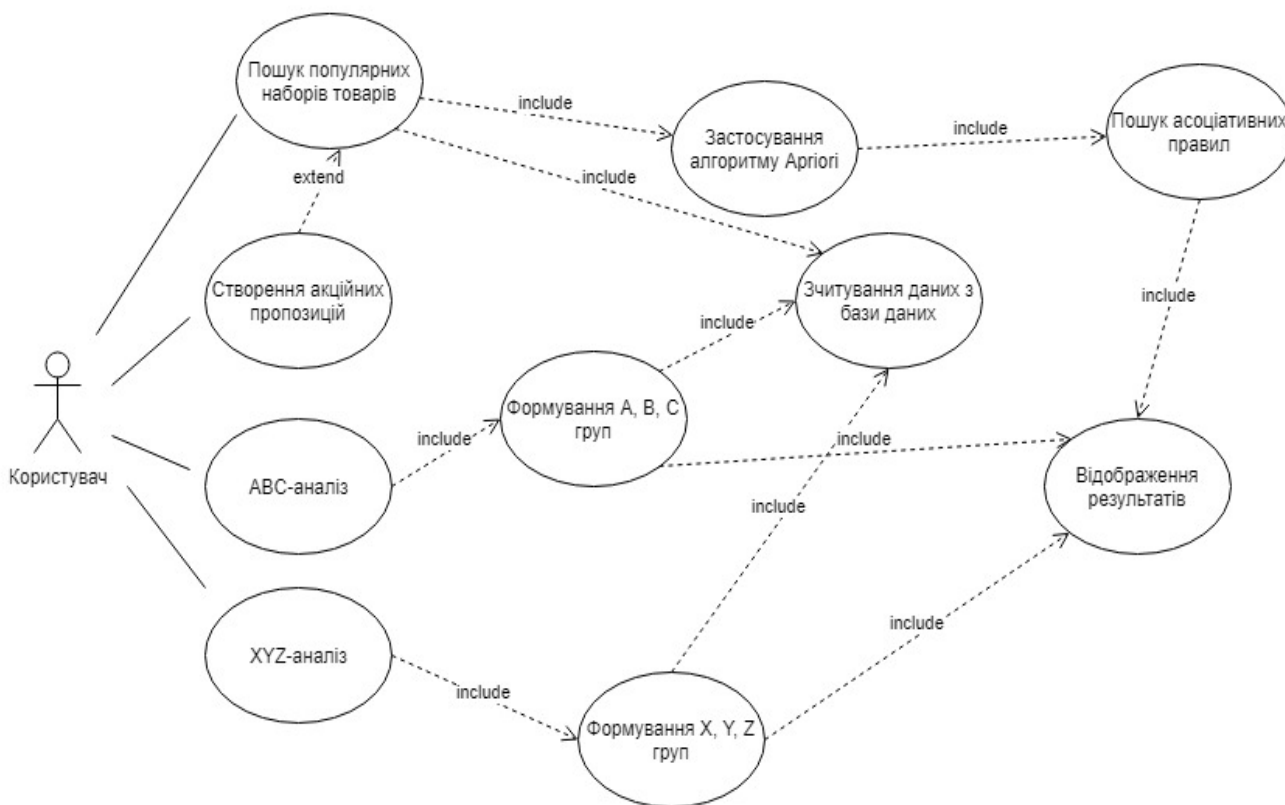


Рис. 1. UML діаграма системи аналізу вмісту кошика покупця супермаркету.

Для реалізації даного програмного продукту були застосовані такі технології: Angular 5, HTML, CSS, Bootstrap, C#, ASP.NET WebApi.

ВИСНОВКИ

В наш час супермаркети є важливим осередком товарообігу. Внаслідок своєї масштабності часто залишаються з проблемами зайвих або недостатніх закупівель, що в свою чергу є причиною зменшення їхньої рентабельності. З огляду на те, актуальним завданням є аналіз кошика покупця супермаркету, що сприяє збільшенню товарообігу, враховуючи потреби споживачів, та зменшує матеріальні збитки.

REFERENCES ЛІТЕРАТУРА

- [1] Які основні тенденції розвитку оптово-роздрібної торгівлі? Режим доступу до ресурсу: http://m.pidruchniki.com/1219101052382/marketing/yaki_osnovni_tendentsiyi_rozvitku_optovo-rozdribnoyi_torgivli.
- [2] Hamlett J. Regulating UK supermarkets: an oral-history perspective. History & Policy, 2008.
- [3] Базюк О.Д., Михалевич В.М. Використання CRMсистеми для управління взаємовідносинами з клієнтами. Режим доступу: <http://nauka.kushnir.mk.ua/?p=74955>.
- [4] Лесна Н. С. Методи аналізу поведінки покупця інтернет магазину для підвищення рівня продажів. Scientific Journal «ScienceRise», 2015, №2, с. 6–9.

Мультимедійні Технології Навчання Студентів з Особливими Освітніми Потребами

Валерій Семенець
Харківський національний
університет радіоелектроніки
Харків, Україна
valery.semenets@nure.ua

Петро Подпружніков
Харківський національний
університет радіоелектроніки
Харків, Україна
petro.podpruzhnykov@nure.ua

Аліна Овченко, Сергій Позняков
Харківський національний
університет радіоелектроніки
Харків, Україна
cntm@ukr.net

Multimedia Technologies for Teaching Students with Special Educational Needs

Valerii Semenets
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
valery.semenets@nure.ua

Petro Podpruzhnikov
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
petro.podpruzhnykov@nure.ua

Alina Ovchenko, Sergey Pozniakov
Kharkiv National University
of Radio Electronics
Kharkiv, Ukraine
cntm@ukr.net

Анотація—Описані нові можливості в навчанні осіб з обмеженими можливостями, які здійсненні завдяки використанню мультимедійних систем і технологій.

Abstract—The paper describes new opportunities for training disabled people through multimedia systems and technologies.

Ключові слова— мультимедійність, доступність, дошка інтерактивна, студенти з особливими потребами.

Keywords— multimedia, accessibility, board interactive, students with special needs.

I. ВСТУП

Мультимедійні технології знаходять своє застосування в різних предметних областях допомагаючи кращому засвоєнню як окремих тем, так і дисциплін, що вивчаються в цілому. Особи із захворюваннями органів зору, слуху, опорно-рухового апарату в силу специфіки ускладнень обмежені в можливостях отримання інформації в звичайному розумінні. Основною метою інтегрованої освіти є здійснення комплексної психологічної і соціальної адаптації студентів з особливими освітніми потребами в умовах професійної підготовки. Мультимедіа технології - спосіб підготовки електронних документів, що включають візуальні і звукові ефекти, мультипрограмування різних ситуацій, для поліпшення сприйняття інформації і її доступності [1].

Інклюзивна освіта являє собою таку організацію процесу навчання, при якій всі учні, незалежно від їх фізичних, психічних, інтелектуальних, культурно-етнічних, мовних та інших особливостей, навчаються разом зі своїми однолітками в освітніх установах, які

враховують їх особливі освітні потреби і надають необхідну підтримку. Тому сьогодні головними завданнями освіти є підготовка учнів до швидкого сприйняття і обробки великих обсягів інформації в рівних умовах [2].

Оцінку потреб у відповідних послугах і продуктах, а також вибір шляхів подолання інформаційних бар'єрів стосовно студентам з особливими освітніми потребами необхідно здійснювати, спираючись на Міжнародну класифікацію функціонування, обмеження життєдіяльності та здоров'я, прийняту Всесвітньою організацією охорони здоров'я (ICF).

У процесі навчання студентів з обмеженими можливостями здоров'я психологи рекомендують розвивати сприйняття, яке пов'язане з функціонуванням різних сенсорних систем: візуальної, аудіальної та особливо кінестетичної. Саме кінестетична система пов'язана з моторною пам'яттю і допомагає довести навички до автоматизму, перевести на рівень підсвідомості.

На сьогоднішній день актуальним є забезпечення всіх студентів рівними можливостями отримання якісної освіти.

II. МОЖЛИВОСТІ ТА ПЕРЕВАГИ МУЛЬТИМЕДІЙНИХ ТЕХНОЛОГІЙ ДЛЯ ІНКЛЮЗИВНОЇ ОСВИТИ

Мультимедійні системи - системи, які впливають оразу по декількох каналах: відео, аудіо, текст, а також часто дають можливість інтерактивної взаємодії, наприклад, в процесі гри при навчанні. Більш складним рівнем мультимедіа вважаються інтелектуальні системи, які не

тільки транслюють інформацію, а й можуть централізовано керувати нею за допомогою центрального комп'ютера [3].

Інтерактивні дошки дозволяють вивести мультимедійність на якісно новий рівень, включаючи в процес сприйняття інформації не однієї людини, а весь колектив учнів для подальшого спільного обговорення навчального матеріалу. Інтерактивні дошки дозволяють виводити зображення з комп'ютера, відеомагнітофона або DVD-програвача, виконувати будь-які кольорові графічні коментарі. За допомогою спеціальної програми все написане на великому екрані можна передати на персональні комп'ютери, планшети студентів або роздрукувати лекцію.

Інтерактивні дошки дозволяють одночасно утримувати увагу всіх учнів. Це підвищує можливості обміну інформацією. Крім того, дозволяють постійно підкріплювати і доповнювати пояснюється матеріал візуальними засобами, забезпечуючи ще одним сенсорним каналом, який допомагає засвоювати інформацію.

Функції дошки дозволяють виконувати імітаційне моделювання реальних об'єктів, процесів, явищ. Студенти можуть колективно взаємодіяти з модельованими процесами і об'єктами, обговорюючи отримані результати.

Завдяки використанню мультимедійних систем, можна адаптувати матеріал для учнів зі слабким зором, змінювати розмір тексту або графічних зображень і коригувати яскравість і контрастність в міру необхідності. Є можливість відтворювати відеоматеріали, збільшуючи зображення, щоб учні могли розглянути деталі, побачити інформацію, яка була б недоступною, а також взаємодіяти з екранними об'єктами, приймаючи на практиці більш повноцінну участь [4].

Переваги використання мультимедійних технологій:

- робить заняття цікавими і розвиває мотивацію;
- надає більше можливостей для участі в колективній роботі, розвитку особистих і соціальних навичок;
- студенти легше сприймають і засвоюють складні питання в результаті більш ясною, ефективною та динамічної подачі матеріалу;
- дозволяє використовувати різні стилі навчання, викладачі можуть звертатися до всіляких ресурсів, пристосовуючись до певних потреб студентів;
- студенти починають працювати більш творчо і стають впевненими в собі;
- надання студентам з обмеженими можливостями здоров'я рівних можливостей освіти поряд з їх здоровими однолітками;
- уникнути появи відчуття ізоляції шляхом створення умов в аудиторії, максимальної доступності;
- підвищення мотивації учнів і максимального залучення їх в навчальний процес шляхом забезпечення можливості організації всіх форм навчання: відеолекції, відеосемінари, практичні роботи в інтерактивному режимі, онлайн контроль знань.

За допомогою об'єднання бездротовою мережею центрального комп'ютера з портативними гаджетами та інтерактивною дошкою з'являється можливість створення для студентів з обмеженими можливостями автоматизованого навчального місця. Даний підхід забезпечує можливість отримання інформації, нових навичок і повноцінне, безперешкодне спілкування з одногрупниками без нозологічних особливостей. Автоматизоване навчальне місце, що складається з настільного комп'ютера, оснащеного монітором, клавіатурою і програмним забезпеченням екранного доступу, є функціональним рішенням для забезпечення інклюзивної освіти.

Автоматизоване навчальне місце для студентів з порушенням опорно-рухового апарату та ДЦП дозволить учням зі слабким зором здійснювати швидку навігацію по мережі Інтернет, працювати в середовищі Windows і з офісними додатками.

Використовуючи спеціалізовані програми, можна розширити географію аудиторії і проводити навчання дистанційно із застосуванням онлайн-передачі даних. Учасники навчання можуть маніпулювати інформацією зі своїх моніторів, активно брати участь в обговоренні матеріалів [5].

Програмні продукти, які випускаються серійно для інтерактивних дошок, можна адаптувати для навчання студентів з різними нозологіями. Можливість об'єднання мережею інтернет основного комп'ютера і портативних гаджетів дозволяє готувати досить складні завдання з декількома варіантами відповідей та забезпечити студентам з обмеженими можливостями довільні введення і вибір необхідної інформації.

ВИСНОВОК

Мультимедійні системи і технології надають широкий вибір підходів до навчання студентів з обмеженими можливостями. Дозволяють уникнути відчуття ізоляції шляхом створення умов максимальної доступності інформації в аудиторії, досягти компенсаторного ефекту при сприйнятті ними навчального матеріалу, включаючи в процес сприйняття інформації не однієї людини, а всього колективу учнів.

ЛІТЕРАТУРА REFERENCES

- [1] Александрова Н.А. Об использовании технологических средств и информационно-коммуникационных технологий в инклюзивном образовании студентов / Н.А. Александрова // Информационные технологии в образовании. – Саратов: ООО «Издательский центр «Наука», 2014. – С. 61-63.
- [2] Гаврилова Е.А. Из опыта инклюзивного обучения средствами ИКТ / Е.А. Гаврилова // Преподавание информационных технологий. – Санкт-Петербург: СПбГУ, 2016. – С.187-188.
- [3] Ярмошук І. На шляху до інклюзивного навчання / І Ярмошук // Математика в сучасній школі. – 2010. – № 1-2. – С. 51 – 54.
- [4] Электронные интерактивные доски SmartBoard, новые технологии в образовании, Москва, 2006г. / [http:// www.smartboard.ru](http://www.smartboard.ru).
- [5] Горюнова М. А., Семенова Т. В. Интерактивные доски и их использование в учебном процессе / М. А. Горюнова, Т. В. Семенова // Информатика и ИКТ: Петербург, 2013. – 336 с.

Reddit Sentiment Analysis to Improve Election Predictions

Bryant Hwang

Department of Computer Science

Korea International School

Gyeonggi-do, South Korea

bryanthwang0712@gmail.com

Abstract—Despite high rates of past success, the majority of polling agencies were unable to predict the outcome of the 2016 United States presidential election between Hillary Clinton and Donald Trump. Further research has hypothesized that these agencies underestimated the amount of support for Trump and overestimated the representation of recent college graduates. Sentiment analysis of posts on popular social media sites presents promising possibilities for better gauging public opinion. In our research, we analyze positive and negative sentiment in Reddit posts from key dates during the 2016 election that mention either candidate. The results of our study match the results of a poll conducted by the New York Times, which showed that public opinion for Trump tended to stay consistent while Clintons positive sentiment decreased over time. Our study also shows that public opinion expressed on Reddit changed drastically in accordance to major events, like debates and email controversies. Alone, our work in sentiment analysis does not produce statistically significant enough results to predict the final outcome of an election, but having the ability to gauge public opinion through social networks is a powerful way that polling agencies and researchers can strengthen their predictions in real time.

Index Terms—sentiment analysis, election prediction, social media, Reddit, Aylie

I. INTRODUCTION

Despite prior consistent success in predicting the outcomes of presidential elections, only fifteen out of 150 polls predicted Donald Trumps victory in the 2016 race [19]. This is a remarkable decrease from the 80% of pollsters that accurately predicted Barack Obamas victory against Mitt Romney in 2012 [19]. The American Association of Public Opinion Research conducted a study on this seemingly irrational result that narrowed down the the cause to three main reasons: the nonresponse bias, Shy Trumper hypothesis, and social desirability bias [16].

The nonresponse bias is the idea that people continually refuse to respond to surveys, which caused pollsters to overestimate the number of college graduate voters [16]. The Shy Trumper hypothesis states that Trump advocates hide their preference for Trump and is closely related to the social desirability bias, which is the idea that people vote for the candidate that is more favorably viewed by society [16]. Combined, the Shy Trumper hypothesis and the social desirability bias caused pollsters to underestimate the number of Trump supporters.

The presence of social media is quickly becoming influential in major elections. Both Barack Obama and Mitt Romney used social media sites like YouTube, Instagram, Facebook, Twitter, Tumblr, Pinterest, and Spotify to reach their younger constituents, but Obama had a much higher follower count: compare 18.5 million to 850 thousand followers [17]. More recently, Donald Trump used his Twitter feed to reach his 15.4 million followers.

In this study, we aim to evaluate the use of social media to predict the outcome of presidential elections. We use Google BigQuery and SQL to analyze sentiment on Reddit, a popular online forum. Reddit contains threads with diverse topics and is a far-reaching social media site, which gives us a variety and abundance of natural language data.

II. RELATED WORKS

The relationship between the sentiment analysis of social media posts and the offline political landscape has been a controversial matter. A few political analysts claim that the Twittersphere shows political opinion to a certain extent, but others fear that the majority of comments on social media compose a pointless bubble. Tumasjan et al., however, have found that the frequency a political party is mentioned online has high correlation with the actual offline political landscape; most of the time, these numbers were found to be similar to the traditional election poll results and were sometimes found to have a higher correspondence to the true results than the traditional election polls.

Previous research done by Tumasjan et al. involved acquiring Tweets and checking if online Tweets mirrored offline political sentiment [23]. They used the Linguistic Inquiry and Word Count text analysis to determine if the sentiment of each message was positive, negative, or neutral towards a certain political party or politician. They concluded that Twitter was indeed used for political deliberation and reflected the offline political landscape.

We combine these two methods to analyze Reddit data: the volume of data and a sentiment analysis algorithm used by Gayo-Avello et al.

III. METHODOLOGY

We analyze the frequency of certain occurrences and the sentiment of politically-related posts on Reddit, a popular

internet forum. We use Reddit because it has high traffic and organizes its conversations by topic, which would give us a large amount of data about public opinion.

A. Dataset

We gather original posts (OPs) and responses to OPs that contain any combination of Donald, Trump, Hillary, or Clinton (case insensitive). We also restrict our data to a few key dates from the 2016 presidential election, presented in Table 1.

TABLE I
KEY DATES FROM THE PRESIDENTIAL ELECTION

Date	Event
July 18-21, 2016	Republican National Convention
July 25-28, 2016	Democratic National Convention
September 9, 2016	Clinton calls Trump voters ignorant deplorables
September 26, 2016	First presidential debate
October 7, 2016	Trumps grab her by the pussy remarks surface
October 9, 2016	Second presidential debate
October 28, 2016	FBI reviewed new evidence Clinton emails
November 7, 2016	Day before the election

B. Volume Analysis

We first analyzed the Reddit traffic each candidate received and then compared it to the actual election. In a simple meta-analysis, we check if the number of times a candidate is mentioned online, without analyzing for sentiment, can predict the outcome of an election.

C. Sentiment Analysis

We use Aylie to classify each post as containing positive, negative, or neutral sentiment towards Clinton or Trump. We discard the neutral posts. The probability of candidate c_1 winning the election is proposed to be

$$p(c_1) = \frac{pos(c_1) + neg(c_1)}{pos(c_1) + neg(c_1) + pos(c_2) + neg(c_2)}$$

where the positive/negative value of a candidate is the percentage of data for that candidate with that sentiment multiplied by the number of Reddit posts that candidate has.

IV. RESULTS AND DISCUSSION

A. Volume Analysis

In the entire dataset we collected, Clinton and Trump were mentioned approximately equally as many times, with Trumps name being more frequent on July 18th-21st and October 7th and Clintons name being more frequent on September 9th and October 28th. These results, however, had a mean absolute error (MAE) of 18.5%. With such a high MAE, we conclude that volume analysis alone is not enough to predict who will win the election and move on to sentiment analysis.

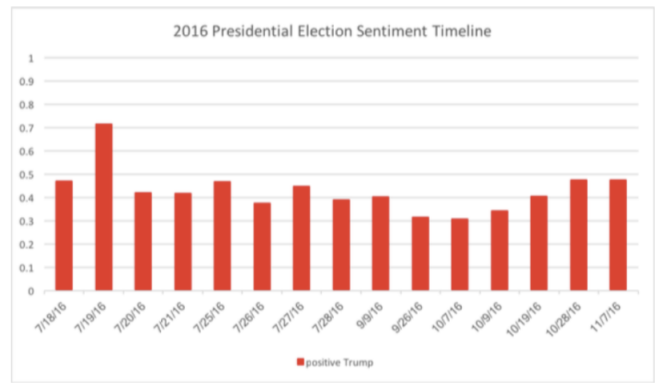


Fig. 1. Positive Sentiment for Donald Trump

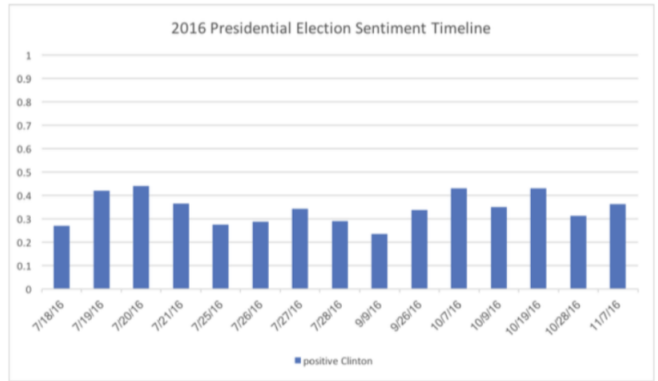


Fig. 2. Positive Sentiment for Hillary Clinton

B. Sentiment Analysis

According to our data, the trend for people who positively viewed Trump was very consistent with only a few variances near the end of the campaigning period. However, the trend for people who positively viewed Clinton depreciated as time went by. We hypothesize that this is because of the email scandal that she suffered.

Figures 1 and 2 above show the positive sentiment for both candidates. The average positive sentiment was 42.8% for Trump and 34.3% for Clinton; interestingly, the majority of users seemed to dislike both candidates. The results in Figures 1 and 2 correlate with the results from a New York Times exit poll: 73% of the people who voted for Trump believed that Clinton was neither honest nor trustworthy, and 71% of Clinton voters thought the same of Trump [25]. In the same study, when asked if Clinton was more qualified to serve as president, 89% of Trump voters responded that she was not, with similar results for Clinton voters [25].

The data from Reddit reflects a similar trend, showing that both candidates had mostly negative comments written about them. The graph presented in Figure 2 demonstrates that Trump had the least positive comments on October 7th, 2016, the same day that the video of Trump bragging about groping women surfaced on the internet. At this point of the presidential election, FiveThirtyEight gave Trump an 18.2%

chance of winning the election. After this day, Trumps positive sentiment rises until election day.

In Figure 2, CLinton hits her lowest point in positive sentiment on September 9th, which was when Clinton called Trump supporters deplorables. On the other hand, her highest point was on July 20th, the day Trump was announced as the Republican candidate at the RNC, even though she had not been officially announced as the Democratic candidate yet. However, Clintons positive sentiment declined when the DNC chair Wasserman Schultz resigned after an email leak on July 25th; afterwards, Clinton was unable to achieve increasing or constant positive sentiment. In the end, she trailed behind Trump by 11.1% the day before the election.

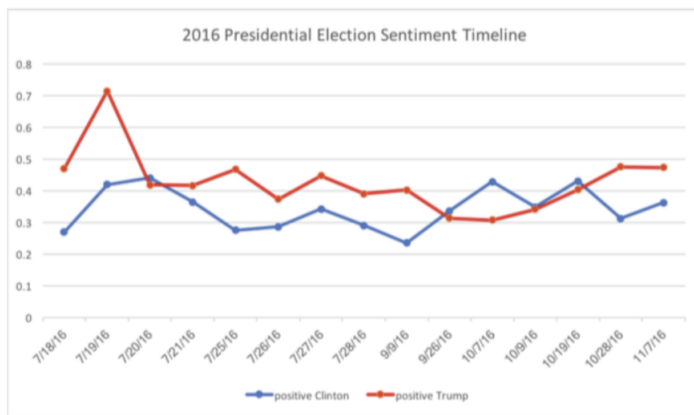


Fig. 3. Combined Positive Sentiment for Both Candidates

Figure 3 shows the combined results for positive sentiment. Trumps positive sentiment was higher than Clintons for the majority of the campaign season, especially in the month before election day. Trumps steady upwards trend in positive sentiment may have been a crucial factor in his victory. This analysis was conducted after the election had ended, but analyzing sentiment during the campaign season might have been a powerful tool to help each candidate monitor their self-image and aid pollsters with their predictions.

V. CONCLUSIONS AND FURTHER STUDIES

Our research shows that the sentiment of relevant Reddit posts corresponded with results from election polls during the campaign season, which suggests that sentiment analysis is a viable avenue for predicting election results. We do, however, need to address some challenges if we are going to use sentiment analysis as a major tool for election prediction.

The first is that Reddit offers no location data, but location is an important part of the United States electoral college. Because the US does not elect presidents with the popular vote, who voted where matters. We also have no way of knowing if some geographical regions are overrepresented.

We also have to consider the prevalence of fake news, robots, and Russian interference. Fake accounts can corrupt data by misrepresenting the views of real people, and Russian interference, a heated topic during the election, can manipulate

other peoples thoughts. Russian interference was a significant problem for Twitter, which admitted that over 50,000 accounts were involved. Another source states that up to 45% of Trumps follower base could be made up of spam or fake accounts, which could have inflated his Reddit positive sentiment. Eliminating bias and filtering unwanted data is an important issue in sentiment analysis, but sentiment analysis is proving itself to be a successful way of gauging the publics view of presidential candidates. With improvement, sentiment analysis could be used during the election cycle to help understand how much the population likes a candidate, which seems to be the most important factor in winning an election.

REFERENCES

- [1] Barthel, Stocking, Holcomb, Mitchell, 2016. Reddit news users more likely to be male, young and digital in their news preferences. <http://www.journalism.org/2016/02/25/reddit-news-users-more-likely-to-be-male-young-and-digital-in-their-news-preferences/>
- [2] Beaumont, 2016. US election final results: how Trump won. <http://theconversation.com/us-election-final-results-how-trump-won-69356>
- [3] Bort, 2012. Twitter Has A Big Problem With Fake User Accounts. <http://businessinsider.com/twitter-has-a-big-problem-with-fake-user-accounts-2012-8>
- [4] Brynes, 2016. Twitter May Have Predicted the Election <https://www.technologyreview.com/s/603010/twitter-may-have-predicted-the-election/>
- [5] Ceron, Curini, Iacus, 2016. How pollsters could use social media data to improve election forecasts. https://washingtonpost.com/news/monkey-cage/wp/2016/12/21/using-these-4-steps-pollsters-could-use-social-media-data-to-improve-election-forecasts/?noredirect=onutm_term=.6f50e215d62d
- [6] Economist, 2016. The Role of Technology in Presidential Election. <https://www.economist.com/usa/2016/11/20/the-role-of-technology-in-the-presidential-election>
- [7] Economist, 2016. Where Polling Failed, Facebook Prevailed. <https://www.economist.com/graphic-detail/2016/11/30/where-polling-failed-facebook-prevailed>
- [8] Electoral Vote, 2016. Tipping-Point State. http://www.electoral-vote.com/evp2016/Pres/tipping_point.html
- [9] Evans, 2016. Can Social Media Be Used to Predict Election Results. <https://www.bbc.com/news/election-us-2016-37942842>
- [10] Gayo-Avello, Daniel, Metaxas, Panagiotis, and Eni Mustafaraj. (2011). Limits of Electoral Predictions Using Twitter. Association for the Advancement of Artificial Intelligence 5: 490-493.
- [11] Gottfried, Shearer, 2016. News Use Across Social Media Platforms 2016. <http://journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>
- [12] Goyder, Warriner, Miller (2002). Evaluating Socio-economic Status (SES) Bias in Survey Nonresponse. <http://scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/evaluating-socio-economic-status-ses-bias-in-survey-nonresponse.pdf>
- [13] Hoffman, 2018. People Have a Lot of Feelings About Melania Trumps New Cyberbullying Initiative. <http://time.com/5268499/melania-trump-cyberbullying-humor/>
- [14] Ingram, 2018. The Facebook Armageddon. https://cjr.org/special_report/facebook-media-buzzfeed.php/
- [15] Katharina Krombholz, Dieter Merkl, Edgar Weippl (2012). Fake Identities in Social Media: A Case Study on the Sustainability of the Facebook Business Model. page 177. <https://www.sba-research.org/wp-content/uploads/publications/krombholzetal2012.pdf>
- [16] Kennedy, C., Blumenthal, M., Clement, S., Clinton, J., Durand, C., Franklin, C., McGeeney, K., Miringoff, L., Olson, K., Rivers, D., Sadd, L., Witt, E., Wlezien, C., Chen, J., Engelhardt, A., Lau, A., Trussler, M. and Patricia Pena Ibarra, L. (2018). An Evaluation of 2016 Election Polls in the U.S. <https://aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx>
- [17] Pardee, 2012. Infographic: How Obama, Romney (and Friends) Are Using Social Media. <http://adage.com/article/campaign-trail/infographic-obama-romney-social-media/236798/>

- [18] Perez, 2016. Analysis of social media did a better job at predicting Trumps win than the polls. <https://techcrunch.com/2016/11/10/social-media-did-a-better-job-at-predicting-trumps-win-than-the-polls/>
- [19] RealClearPolitics, 2012. General Election: Romney vs.Obama. https://realclearpolitics.com/epolls/2012/president/us/general_election_romney_vs_obama-1171.html
- [20] Saru, 2017. What is Twitter Doing about Fake News? <http://techacute.com/twitter-fake-news/>
- [21] SimilarWeb, 2018. Reddit.com June 2018 Overview. <https://similarweb.com/website/reddit.com>
- [22] Syed, 2016. How To Become President Of The United States. https://s3.amazonaws.com/app_usa_prod_eqffnyamdzrb/pdf6099%20%283%29.pdf
- [23] Tumasjan, Andranik Oliver Sprenger, Timm Sandner, Philipp Welp, Isabell. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Word. Journal Of The International Linguistic Association*. 10.
- [24] Woolf, 2016. How to solve Facebook's fake news problem: experts pitch their ideas. <https://theguardian.com/technology/2016/nov/29/facebook-fake-news-problem-experts-pitch-ideas-algorithms>
- [25] Wortham, 2012. The Presidential Campaign on Social Media. http://archive.nytimes.com/www.nytimes.com/interactive/2012/10/08/technology/campaign-social-media.html?_r=1
- [26] Zurcher, 2016. Hillary Clinton emails - what's it all about? <https://bbc.com/news/world-us-canada-31806907>

Наукове електронне видання комбінованого використання
Можна використовувати в локальному та мережному режимах

**КОМП'ЮТЕРНІ НАУКИ, ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ
ТА СИСТЕМИ УПРАВЛІННЯ**

**Матеріали
міжнародної науково-технічної конференції
студентів, аспірантів та молодих вчених
28–30 листопада 2018 року
Івано-Франківськ, Україна**

Матеріали статей опубліковані в авторській редакції

**COMPUTER SCIENCE, INFORMATION TECHNOLOGIES
AND MANAGEMENT SYSTEMS**

**Proceedings
of the International Scientific Young Scientists Conference**

**2018, November , 28th to 30th
Ivano-Frankivsk, Ukraine**

Materials are published in author's edition

Технічний редактор к.т.н М.Л. Петришин (ПНУ)

Видавець

Прикарпатський національний університет
імені Василя Стефаника
вул. С. Бандери, 1, м. Івано-Франківськ, 76025,
тел.: 71-56-22

E-mail: vdvcit@pu.if.ua

Свідоцтво суб'єкта видавничої справи
ДК № 2718 від 12. 12. 2006

ISBN 978-966-640-448-3