

Міністерство освіти і науки України
Харківський національний
університет імені В. Н. Каразіна

КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ В НАУКОЄМНИХ ТЕХНОЛОГІЯХ



ЗБІРНИК НАУКОВИХ ПРАЦЬ МІЖНАРОДНОЇ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

(23-25 листопада 2022 року, м. Харків, Україна)

Харків
2022

Міністерство освіти і науки України
Харківський національний
університет імені В. Н. Каразіна

КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ В НАУКОЄМНИХ ТЕХНОЛОГІЯХ



ЗБІРНИК НАУКОВИХ ПРАЦЬ МІЖНАРОДНОЇ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

(23-25 листопада 2022 року, м. Харків, Україна)

Харків
2022

*Регістраційне посвідчення УкрІНТЕІ МОН України
(№ 1054 від 20 грудня 2021 року)*

*Затверджено до друку рішенням Вченої ради
Харківського національного університету імені В. Н. Каразіна
(протокол № 19 від 26 грудня 2022 року)*

Редакційна колегія:

Азаренков Н. А. (гол. редактор), д.ф.-м.н.,
академік НАН України, проф., ІВТ ХНУ імені
В. Н. Каразіна

Ванін В. А., д.т.н., проф., НТУ «ХП»

Горбенко І. Д., д.т.н., проф., ФКН ІВТ ХНУ
імені В. Н. Каразіна

Доля Г. М., проф., д.т.н., проф. ФКН ІВТ ХНУ
імені В. Н. Каразіна

Жолткевич Г. Н., д.т.н., проф., ФМІ ХНУ
імені В. Н. Каразіна

Куклін В. М., д.ф.-м.н., проф., ФКН ІВТ ХНУ
імені В. Н. Каразіна

Лазурик В. Т., д.ф.-м.н., проф., ФКН ІВТ ХНУ
імені В. Н. Каразіна

Рассомахін С. Г., д.т.н., доц., ФКН ІВТ імені
В. Н. Каразіна

Споров О. Є., к.ф.-м.н., доц. ХНУ імені
В. Н. Каразіна

Стервєдєв М. Г., к.т.н., доц., ФКН ІВТ імені
В. Н. Каразіна

Толстолузька О. Г., д.т.н., с.н.с., доц., ФКН ІВТ
ХНУ імені В. Н. Каразіна

Ткачук М. В., д.т.н., проф., ФКН ІВТ ХНУ
імені В. Н. Каразіна

Харченко В. С., д.т.н., проф., НАУ
імені М. Є. Жуковського

Шматков С. І., д.т.н., проф., ХНУ
імені В. Н. Каразіна.

Шульга М. Ф. д.ф.-м.н., акад. НАНУ, проф.,
ІНЦ ХФТІ НАНУ

Адреса редакційної колегії: 61022, м. Харків, майдан Свободи, 6,
ХНУ імені В. Н. Каразіна, к. 534.
Тел. +380 (57) 705-42-81, email: kmht@karazin.ua.

Комп'ютерне моделювання в наукоємних технологіях: збірник наукових праць
міжнародної науково-технічної конференції (23-25 листопада 2022 р., м. Харків, Україна) –
Харків: ХНУ ім. В. Н. Каразіна, 2022. –218 с.

Доповіді, що увійшли до збірника, висвітлюють такі напрямки:

- 1 секція: Математичне моделювання технологічних процесів та приладів.
- 2 секція: Моделювання інформаційних процесів у складних і розподілених системах.
- 3 секція: Системи автоматизованого збору та когнітивного представлення наукових даних.
- 4 секція: Моделювання фізичних процесів в радіаційних, плазмових та інших сучасних технологіях.
- 5 секція: Безпека інформаційних систем і технологій.
- 6 секція: Моделі процесів розробки та оцінки якості програмного забезпечення.
- 7 секція: Студентів, аспірантів та молодих вчених.

Для викладачів, наукових працівників, аспірантів, студентів вишів.
Робочі мови конференції: українська, англійська.

ЗМІСТ

ЗМІСТ	4
АБАКУМОВА В.С., ХРУСЛОВ М.М. CRITICAL INTEGRAL DIMENSIONS OF ADDRESS LOADING.	8
АВЕРКОВ О.Ю., АВЕРКОВ Ю.О. ЧИСЕЛЬНЕ МОДЕЛЮВАННЯ РЕЗИСТИВНОЇ НЕСТІЙКОСТІ ТРУБЧАСТОГО ПУЧКА ЕЛЕКТРОНІВ, ЩО РУХАЄТЬСЯ НАД ПЛАЗМОВИМ ЦИЛІНДРОМ	10
АВЕРКОВ Ю.О., АВЕРКОВ О.Ю. ОСОБЛИВОСТІ СПЕКТРУ НАПІВПРОВІДНИКОВОЇ НАНОТРУБКИ В ЗОВНІШНЬОМУ МАГНІТНОМУ ПОЛІ	14
АРТЮХ С.О., ХРУСЛОВ М.М. ОСОБЛИВОСТІ ВИКОРИСТАННЯ ЦИФРОВИХ ДВІЙНИКІВ	18
БЕЗСМЕРТНИЙ Д.Р., КОНДРАТЮК Д.С. МЕТОДИ НАВЧАННЯ РЕКУРЕНТНОЇ ТА ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ НА ОСНОВІ ПОПУЛЯЦІЙНОГО АЛГОРИТМУ ДЛЯ РОЗПІЗНАВАННЯ ОБРАЗІВ	20
БЕРЕЖНИЙ А.А. ПРОГРАМНА МОДЕЛЬ РОБОТА З ДИФЕРЕНЦІЙНИМ ПРИВОДОМ	24
БИКОВА Т.В. АНАЛІЗ ВИМОГ ДО МЕТРОЛОГІЧНИХ ХАРАКТЕРИСТИК «ІНТЕРНЕТУ РЕЧЕЙ»	28
БУРЛАЙ І.С., БАКУМЕНКО Н.С. КОМП'ЮТЕРНА МОДЕЛЬ СИСТЕМИ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЦІН НА ЕНЕРГОНОСІЇ ЗА ДОПОМОГОЮ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ.	32
ВЕРЕМІЄНКО Б.О., ЧУБ О.І. ВИКОРИСТАННЯ АФІННИХ ПЕРЕТВОРЕНЬ У ГЕОМЕТРИЧНОМУ МОДЕЛЮВАННІ ...	36
ВОЛИНСЬКИЙ В.В., МІРОШНИК М.А. РОЗРОБКА МОДЕЛІ НАУКОВО-ДОСЛІДНИЦЬКОЇ ЛАБОРАТОРІЇ	39
ГАМЗАЄВ Р.О., ШКОДА Б.В. МОДЕЛІ ТА ПРИНЦИПИ ПОБУДОВИ ЛІНІЙКИ ПРОГРАМНИХ ПРОДУКТІВ НА ПРИКЛАДІ СИСТЕМ ОБМІНУ ПОВІДОМЛЕННЯМИ	44
ГЛУХОТА А.В., ШЕВЧЕНКО Д.О., СТІЛЕЦЬ В.Є. НЕЙРОМЕРЕЖЕВІ ТЕХНОЛОГІЇ ДЛЯ ОБРОБКИ ШУМУ В ЗОБРАЖЕННЯХ	48
ГОЛОВЧЕНКО В.Б., СТІЛЕЦЬ В.Є. РЕГРЕСІЙНІ МОДЕЛІ ПРОГНОЗУВАННЯ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ	51
ДААС Т.І., ТКАЧУК М.В. ПРО ОДИН ПІДХІД ДО РОЗРОБКИ БАНКІВСЬКИХ ІНФОРМАЦІЙНИХ СИСТЕМ: ЗНАННЯ-ОРІЄНТОВАНІ МОДЕЛІ ТА МІКРОСЕРВІСНІ АРХІТЕКТУРИ	54
ДЕЙНЕГА Т.С., СВАТОВСЬКИЙ І.І. ДОСЛІДЖЕННЯ ЗАСТОСУВАННЯ АЛГОРИТМІВ ШТУЧНОГО ІНТЕЛЕКТУ В СИСТЕМАХ ВИЯВЛЕННЯ/ЗАПОБІГАННЯ ВТОРГНЕНЬ	58

ДЕРЕВ'ЯНКО Я.А., ГОРБЕНКО І.Д. ГЕНЕРАЦІЯ НЕЛІНІЙНИХ ВУЗЛІВ ЗАМІН З ВИКОРИСТАННЯМ ФАКТОРІАЛЬНОЇ СИСТЕМИ ЧИСЛЕННЯ	63
ДОЛЯ Г.М., БОНДАРЕНКО К. О. SIMULATION MODEL OF THE MULTIMODE LASER RADIATION TRANSFORMATION ON A GLASS BEADS RETROREFLECTIVE SURFACE	67
DONETS V., UGRYUMOV M., STRILETS V. A MEASURE OF COMPACTNESS FOR FUZZY CLUSTERING BASED ON ENTROPY.....	71
ДОРОШЕНКО М.І, СТРИЛЕЦЬ В.Є. МЕТОД АНАЛІЗУ ПАРАМЕТРІВ ТРЕНДУ У ТРАФІКУ КОМП'ЮТЕРНОЇ МЕРЕЖІ	75
ЄГОРОВ Є.С., ШМАТКОВ С.І. КЛАСТЕРИЗАЦІЯ ДАНИХ НА ОСНОВІ МЕТОДІВ БАГАТОВИМІРНОГО ДИСКРИМІНАНТНОГО АНАЛІЗУ ТА С-СЕРЕДНІХ	79
ЄСІНА М. В., ГОРБЕНКО І. Д., КАЧКО О. Г. РОЗРОБКА, СТАНДАРТИЗАЦІЯ ТА ВПРОВАДЖЕННЯ СТІЙКИХ ПОСТКВАНТОВИХ КРИПТОПРИМІТИВІВ НА МІЖНАРОДНОМУ ТА НАЦІОНАЛЬНОМУ РІВНЯХ	83
ЗАЦ О.Д., ШМАТКОВ С.І. МОДЕЛЮВАННЯ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ	87
ЗЕЛЕНСЬКА Н.В. АНАЛІЗ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО УПРАВЛІННЯ КОМП'ЮТЕРНИМИ ДАНИМИ	90
ІСАЕВ В.С МОДЕЛЬ УПРАВЛІННЯ РИЗИКАМИ ПРИ РЕАЛІЗАЦІЇ ІТ-ПРОЕКТУ	94
КАНДІЙ С.О. ПОРІВНЯННЯ АРГУМЕНТІВ БЕЗПЕКИ ПЕРСПЕКТИВНИХ МЕХАНІЗМІВ ІНКАПСУЛЯЦІЇ КЛЮЧІВ	97
КАПЛУН В.М., ТОЛСТОЛУЗЬКА О.Г. РОЗРОБКА КОМП'ЮТЕРНОЇ МОДЕЛІ WEB-ЗАСТОСУНКА З ОБРОБКОЮ ВЕЛИКИХ ОБСЯГІВ ДАНИХ	101
КИЛИМНИК Д.О. ОСНОВНІ ВИЗНАЧЕННЯ ТА ХАРАКТЕРИСТИКИ ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНИХ СИСТЕМ	105
KIRICHOK A.V., MISHYN O.V. ON THE DYNAMICS OF A SHORT BUNCH OF RELATIVISTIC ELECTRONS IN PLASMA	109
КОВАЛЕНКО А. В., БУЛАВІН Д. О. ВИКОРИСТАННЯ РОЗШИРЕНОГО АНАЛІЗУ ДОХОДУ ТА ВИТРАТ ДЛЯ УПРАВЛІННЯ ЯКІСТЮ ІТ-ПРОЕКТУ	113
КОНЮШЕНКО Р.В., СВАТОВСЬКИЙ І.І. ДОСЛІДЖЕННЯ ЗАХИЩЕНОСТІ КАНАЛІВ ПЕРЕДАЧІ ДАНИХ І УПРАВЛІННЯ БЕЗПЛОТНИМИ ЛІТАЛЬНИМИ АПАРАТАМИ	116

КОСТЮЧЕНКО А.Д., ЧУБ О.І. ВИКОРИСТАННЯ МЕТОДІВ СТАТИСТИЧНОГО АНАЛІЗУ В РЕКОМЕНДАЦІЙНИХ СИСТЕМАХ	120
КУЛИК В.М., БАКУМЕНКО Н.С. ПРОГНОЗУВАННЯ НАВАНТАЖЕНЬ В КОМП'ЮТЕРНИХ МЕРЕЖАХ ЗА ДОПОМОГОЮ АВТОРЕГРЕСІЙНИХ МЕТОДІВ	123
ЛАГУТІН А.Г., СТРИЛЕЦЬ В.Є. СУЧАСНІ ЗАСОБИ І ТЕХНОЛОГІЇ СТВОРЕННЯ ВЕБ-САЙТІВ	127
МАКСИМЕНКО В.Ю., ТОЛСТОЛУЗЬКА О.Г. ОНЛАЙН-ТРЕНАЖЕР ДЛЯ ВИЗНАЧЕННЯ КОГНІТИВНИХ ЗДІБНОСТЕЙ ЛЮДИНИ	130
МАЛИГА І.Є. МОДЕЛІ І МЕТОДИ ОБРОБКИ ТЕКСТОВИХ ДАНИХ ЗАСОБАМИ МАШИННОГО НАВЧАННЯ	133
МАНИЦЬКИЙ С.С. АНАЛІЗ МЕТОДІВ ПОДІЛУ КАНАЛІВ У ВИРІШЕННІ ПИТАНЬ МНОЖИННОГО ДОСТУПУ В КОМП'ЮТЕРНІЙ МЕРЕЖІ	137
МЕДВЕДЕНКО В.О. АНАЛІЗ ЗАСТОСУВАННЯ КОНЦЕПЦІЇ WEB 3.0 ТА ЇЇ ВПЛИВУ НА МЕРЕЖУ ІНТЕРНЕТ	141
МІРОШНИЧЕНКО Д.О., ТОЛСТОЛУЗЬКА О.Г., ПЕТРИЧЕНКО М.О. МОДЕЛЬ АВТОМАТИЗАЦІЇ РОЗГОРТАННЯ ІНФРАСТРУКТУРИ ДЛЯ СИСТЕМИ УПРАВЛІННЯ КОНТЕНТОМ В ХМАРНОМУ СЕРЕДОВИЩІ	144
МОРОЗ Є.Р., ХРУСЛОВ М.М. РОЗРОБКА СИСТЕМИ МОНІТОРИНГУ МЕРЕЖІ ТЕЛЕФОННИХ ЛІНІЙ	149
ОРЛЕНКО О.А., МІШИН О.В. РОЗМІТКА ПОЧАТКУ ТА КІНЦЯ ТРЕКІВ В АУДІОФАЙЛІ МЕТОДОМ КЛАСТЕРИЗАЦІЇ	153
ПЕТУХОВ І.І., ЛИСИЦЯ О.Ю. CFD-МОДЕЛЮВАННЯ ТЕПЛОМАСООБМІННИХ ТА ГІДРОГАЗОДИНАМІЧНИХ ПРОЦЕСІВ У СТРУМЕНЕВИХ АПАРАТАХ	157
ПОДНАІКО V.O., RASSOMAKHIN S.G. CURRENT ALGORITHMS FOR IDENTIFICATION USING BIOMETRIC DATA	160
ПОМОГАЄВ К.О., ТОЛСТОЛУЗЬКА О.Г., АРТЮХ О. А. КОМП'ЮТЕРНА МОДЕЛЬ 3D-ВІЗУАЛІЗАЦІЇ ПРИМІЩЕНЬ ІЗ ЗАСТОСУВАННЯМ МЕТОДІВ DATA MINING	162
ПОПОВА А.О., ТОЛСТОЛУЗЬКА О.Г. АНАЛІЗ ПІДХОДІВ ДО ОБРОБКИ ДАНИХ, ЩО ПРИХОДЯТЬ ВІД РЕЧЕЙ. СИСТЕМИ ПОБУДОВИ ІНТЕРНЕТУ РЕЧЕЙ	167
RADOUTSKYI K. E., RADOUTSKA A. K. ADVANTAGES AND DISADVANTAGES OF AGILE PROJECT MANAGEMENT METHODOLOGY	171

СКОБ Ю.О. МОДЕЛЮВАННЯ НАСЛІДКІВ ВПЛИВУ ТЕПЛОВОГО ВИПРОМІНЮВАННЯ НА ДОВКІЛЛЯ	173
СУХОВЕРСЬКИЙ Д.О., ДОЛЯ Г.М. МОДЕЛЬ РОБОТИ ВОЛОКОННО-ОПТИЧНОЇ ЛІНІЇ ЗВ'ЯЗКУ ІЗ АКУСТООПТИЧНИМ МОДУЛЯТОРОМ СВІТЛА	176
ТЕЛЕЖЕНКО Д.О., ТОЛСТОЛУЗЬКА О.Г. ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ АРХІТЕКТУРИ ВІРТУАЛЬНИХ РОЗПОДІЛЕНИХ СИСТЕМ	179
ТОВСТОКОРЕНКО О.Ю. ПРОБЛЕМИ ПІДТРИМКИ ПРОЦЕСІВ КОНФІГУРУВАННЯ СЦЕНАРІЇВ АВТОМАТИЗАЦІЇ У РОЗРОБЦІ ТА СУПРОВОДІ СИСТЕМ «РОЗУМНИЙ БУДИНОК»	182
ТОЛСТОЛУЗЬКИЙ Є.Д., МІРОШНИК М.А. МОДЕЛЬ МЕРЕЖЕВОГО ПЛАНУВАННЯ ІЗ ЗАСТОСУВАННЯМ МЕТОДІВ МУЛЬТИПАРАЛЕЛЬНОЇ ОБРОБКИ ІНФОРМАЦІЇ	185
УРСУ Р.Є., ЛАБЕНКО Д.П. МОДЕЛЬ УПРАВЛІННЯ ЯКІСТЮ ПРОДУКЦІЇ	188
ХРУСЛОВ М. М., КІПРЕНКО Є. В. РОЗРОБКА ТА РЕАЛІЗАЦІЯ АЛГОРИТМІВ АНАЛІЗУ ТА ОБРОБКИ ДАНИХ НАТУРНИХ ТА ЧИСЕЛЬНИХ ЕКСПЕРИМЕНТІВ З ВИКОРИСТАННЯМ СУЧАСНИХ БАЗ ДАНИХ	191
ХРУСЛОВ М.М., ЗОЛУХІНА О.В., ВЕКЛИЧ Д.А. ПРОГРАМНЕ РІШЕННЯ ОПТИМІЗАЦІЇ УПРАВЛІННЯ ПЕРСОНАЛОМ НА ПРИКЛАДІ РЕСТОРАННОГО БІЗНЕСУ	193
ЧЕПАК А.А., СТІЛЕЦЬ В.Є. МАШИННЕ НАВЧАННЯ В ЗАДАЧАХ АНАЛІЗУ І ВДОСКОНАЛЕННЯ ВЕБ-САЙТІВ	196
СНУВ О.І., NOVOZHYLOVA M.V. SOFTWARE CARPENTRY ENVIRONMENT AS A MODEL OF MODERN EDUCATIONAL AND RESEARCH COMMUNITY	200
ШЕВЧЕНКО Д.О., ЛИХАЧ О.Ю., УГРЮМОВ М.Л. МЕТОДИ ВИЯВЛЕННЯ ВИКИДІВ В СИСТЕМАХ АВТОМАТИЗОВАНОГО ЗБОРУ ДАНИХ	204
ШМАТКОВ С. І., ТИТАРЕНКО Т. О. МЕТОД АДАПТИВНОЇ КЛАСТЕРИЗАЦІЇ ДАНИХ НА ОСНОВІ МЕТЕТІЧЕСКОГО АЛГОРИТМУ	210
ЮЩЕНКО В.С. ДАТЧИКИ ВИМІРЮВАННЯ ВІДСТАНЕЙ ЯК ОСНОВА 3D-СКАНУВАННЯ	213

АБАКУМОВА В.С., ХРУСЛОВ М.М.

CRITICAL INTEGRAL DIMENSIONS OF ADDRESS LOADING

The theme of the work: Critical integral dimensions of address loading.

The purpose of the work: Identify the boundary values of the indicators for transmission of the HTTP request over the network.

Research methods: the method of comparison and measurement.

The results: The current state shows that web browsers (clients) have a significant enough URL limit to allow requests to be sent to server. In turn, critical integral dimension depends on the configuration of the server itself to receive and process requests (with header and/or body) within the limit specified in the configuration. Such a limit can apply to both the size of the header and the body.

Conclusions: each server has its configuration of header and request body size constraints based on the resources used. Moreover, requests more significant than the MTU are handled by the fragmentation mechanism, which in turn carries the risk of the server processing the request in the case of reassembling.

Key words: HTTP Request, HTTP Protocol, TCP/IP, REST API, Maximum Transmission Unit.

HTTP Requests

The Internet is already well-established in most areas of human activity. The primary use of the Internet is to exchange information, i.e. to receive and send it with the possibility of making changes. The HTTP protocol can be considered the basis for large-scale use.

HTTP was designed to be usable as an interface to distributed object systems. It is the backbone of all web-based communication and a client-server protocol. The HTTP protocol is known to rely on the TCP standard, meaning that it is connection-based.

Therefore, a client and a server must establish a TCP connection to exchange an HTTP request/response pair. HTTP defines a set of request methods to indicate the desired action to be performed for a given resource. The standard methods for HTTP/1.1 specifying CRUD operation, presence of headers and body are defined below in Table 1 [1, 2].

Table1 Characteristics of HTTP requests

Method	CRUD	Accepts Request Header	Accepts Request Body
GET	Read	Yes	Optional [1, 2]
POST	Create	Yes	Yes
PUT	Update	Yes	Yes
DELETE	Delete	Yes	Optional [1, 2]
PATCH	Update	Yes	Yes
HEAD	Read	Yes	Optional [1, 2]
TRACE	Read	Yes	No
OPTIONS	Read	Yes	Optional [1, 2]
CONNECT	-	Yes	Optional [1, 2]

Defining REST APIs

Based on Roy Fielding's dissertation – Architectural Styles and the Design of Network-based Software Architectures, which was published in 2000, we can assume that The Representational State Transfer (REST) style is an abstraction of the architectural elements within a distributed hypermedia system [3].

In this case, REST does not use any concrete technology. The most common solution is choosing HTTP protocol and JSON as representation formats. REST API is an application programming interface that conforms to the constraints of REST architectural style and allows for interaction with RESTful web services [4]. From this, it follows that HTTP methods are a means to manipulate resources.

Therefore, the design and development of the API should follow design patterns and best practices. In today's distributed and microservice systems, APIs have even become a separate product, so attention must be paid to indicators such as client/server traffic load and errors during network transmission that cause the server to respond 'incorrectly'.

To build a proper API Ecosystem, it is necessary to perform testing – by applying a standard testing pyramid. It is also essential to conduct performance testing, which should include a

comprehensive view of the product, not the network alone – the test should include the network, parsing, responding to the query and returning a response [5].

Critical integral dimension

Resources are essential in the construction and functioning of the API ecosystem. On this basis, the server has limitations on receiving and processing capabilities for any number of requests from the client. API design should focus on limiting the backend calls and minimising the size of the data returned. Also, consideration must be given to content attack protection (CAP) [6].

The question thus becomes what the limits to the functioning of the ecosystem are. HTTP does not have specific length limitations for many of its protocol elements because the lengths that might be appropriate will vary widely, depending on the deployment context and purpose of the implementation [4].

The REST API supports Uniform Resource Locators (URLs) with a length of up to 6000 characters (approximately 5.859 KB). The HTTP protocol does not limit the size of requests, but each browser and server have a data transfer limit. It is recommended to use an average of 512-1024 KB or a maximum of 5 parameters to create one request. Requests will usually be limited by either the web server or the programming technology used to process the form submission. The following maximum URL lengths apply to commonly used browsers. Since all of these browsers have a substantial market share, website operators have to stick to the lowest common denominator and limit their URLs to 2,083 characters in length. For example, Microsoft Edge: has 2 083 characters, Google Chrome: has 32 779 characters, Mozilla Firefox: has more than 64 000 characters, and Apple Safari: has more than 64 000 characters.

The network's constraints will impose the need to be considered in this way. Digital data exchange within a computer network uses frames and packets. The maximum payload of a TCP segment is 65,495 bytes. The entire TCP segment must fit in the 65,515-byte payload field of an IP packet. Since the TCP header is a minimum of 20 bytes, only 65,495 bytes are left for TCP data.

Fragmentation is a normal part of the Internet Protocol (IP). In essence, each type of network potentially has a different Maximum Transfer Unit (MTU) – the maximum frame size (size of the data limit) a network will transport based on hardware in a network. MTU is limited in size and can be no more than 1500 bytes. The maximum size of a TCP packet is 64KB (65535 bytes), and it will require at least 44 frames [7]. If the packet is too big to travel between two routing devices, it gets broken into fragments. These fragments look like IP packets in their own right and can traverse the network. They are reassembled when they reach their destination.

Thus, each server has its configuration of header and request body size constraints based on the resources used. Moreover, requests more significant than the MTU are handled by the fragmentation mechanism, which in turn carries the risk of the server processing the request in the case of reassembling.

REFERENCES

1. Fielding R. RFC 9110: HTTP Semantics / R. Fielding, M. Nottingham, J. Reschke – URL: <https://www.rfc-editor.org/rfc/rfc9110.html> (Last accessed: 01.11.2022).
2. Fielding R. Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content / J. Reschke – URL: <https://www.rfc-editor.org/rfc/rfc7231> (Last accessed: 01.11.2022).
3. Fielding R. Architectural Styles and the Design of Network-based Software Architectures / Fielding Roy Thomas, 2000.
4. Continuous API Management / M. Medjaoui, E. Medjaoui, R. Mitra, M. Amundsen., 2022.
5. Gough J. Mastering API Architecture / J. Gough, D. Bryant, M. Auburn., 2021.
6. Subramanian H. Hands-On RESTful API Design Patterns and Best Practices / H. Subramanian, P. Raj.
7. Fall, Kevin R. TCP/IP illustrated. – 2nd ed. / Kevin R. Fall, W. Richard Stevens.

ХРУСЛОВ Максим Михайлович – к. ф.-м.наук, доцент кафедри електроніки і управляючих систем ХНУ імені В.Н.Каразіна; e-mail: maksym.khruslov@karazin.ua ; <https://orcid.org/0000-0001-9639-9340>

АБАКУМОВА Вікторія Сергіївна – студентка кафедри моделювання систем і технологій; xa11867750@student.karazin.ua; <https://orcid.org/0000-0002-1352-9902>

Наукові інтереси: – *Аналіз даних, Машинне навчання, Інтернет речей.*

АВЕРКОВ О.Ю., АВЕРКОВ Ю.О.

ЧИСЕЛЬНЕ МОДЕЛЮВАННЯ РЕЗИСТИВНОЇ НЕСТІЙКОСТІ ТРУБЧАСТОГО ПУЧКА ЕЛЕКТРОНІВ, ЩО РУХАЄТЬСЯ НАД ПЛАЗМОВИМ ЦИЛІНДРОМ

Topic of the article. The topic of the work is theoretical study of nonlinear interaction of charged particles with solid-state plasma media of cylindrical configuration.

Purpose of the work. The purpose of the work is to establish a physical mechanism, which is the basis of nonlinear stabilization of the resistive instability of a tubular electron beam that moves along the solid-state plasma cylinder.

Research methods. It has been assumed that the electron collision frequency in the plasma cylinder is much higher than the frequency of plasma eigenmode. The beam is assumed to be nonrelativistic, and, thus, the problem is solved in the electrostatic approximation. With the use of the Maxwell equations and equations of plasma electron motions by the Krylov–Bogolyubov averaging method, a system of self-averaged equations for amplitude and phase of the wave, which change slowly in time, as well as for coordinates and speed of the beam electrons has been obtained. The numerical analysis of this system is performed by the method of macroparticles.

Results. The growth of the wave amplitude is shown to be stabilized nonlinearly because of the self-trapping of the beam electrons by the field of the electrostatic wave excited in the beam itself. The saturation time of instability and the maximum amplitude of the excited wave are shown to depend on the radius of the plasma cylinder. It has been demonstrated that the larger the radius of the plasma cylinder, the later the nonlinear stage of instability begins and the larger the maximum amplitude of the excited wave.

Conclusions. In this work, the physical mechanism underlying the resistive instability effect has been described in detail. This allows a deeper understanding of the physical nature of the interaction of charged particle beams with the solid-state plasma media.

Keywords: beam instability, nonlinear interaction, Krylov–Bogolyubov averaging method, Vavilov–Cherenkov resonance, self-trapping effect.

Вступ

Однією із актуальних проблем сучасних радіофізики та електроніки є визначення механізмів генерації електромагнітних хвиль внаслідок взаємодії заряджених частинок, що рухаються, з різними електродинамічними структурами. Особлива увага приділяється дослідженням пучкових нестійкостей в сильнозбурхувальній твердотільній плазмі, коли частота зіткнень носіїв заряду є набагато більшою частоти електромагнітної хвилі, зокрема електростатичної хвилі. Наприклад, у роботі [1] описано ефект резистивної нестійкості, яка виникає внаслідок поглинання енергії електронного потоку діелектричним середовищем, в якому він поширюється. У нашій роботі в електростатичному наближенні з використанням рівнянь Пуассона та руху електронів побудовано нелінійну теорію резистивної нестійкості трубчастого електронного пучка, що обдуває твердотільний плазмений циліндр. Метою є виявлення стабілізації зростання амплітуди електростатичної хвилі пучка внаслідок самозахоплення електронів пучка її полем.

Постановка задачі й результати моделювання

Нехай твердотільний плазмений циліндр, радіус якого є ρ_0 , займає область простору $0 \leq \rho \leq \rho_0$, $0 \leq \varphi \leq 2\pi$ і $|z| < \infty$ (вісь z циліндричної системи координат спрямована вздовж вісі симетрії циліндра). Трубочастий пучок заряджених частинок з рівноважною густиною n_0 та швидкістю $v_0 \ll c$ (де c — швидкість світла у вакуумі) транспортується уздовж поверхні циліндра. Взаємодія електронного пучка з власними хвилями (коливаннями) діелектричного циліндра описується системою рівнянь, що складається з рівнянь електростатики, матеріального рівняння та рівняння руху електронів плазми:

$$\operatorname{rot} \mathbf{E}(\mathbf{r}, t) = 0; \operatorname{div} \mathbf{D}(\mathbf{r}, t) = 4\pi e n(\mathbf{r}, t), \quad (1)$$

$$\mathbf{D}(\mathbf{r}, t) = \varepsilon_0 \mathbf{E}(\mathbf{r}, t) + 4\pi e N_0 \int_{-\infty}^t \mathbf{u}(\mathbf{r}, t') dt', \quad (2)$$

$$\frac{\partial \mathbf{u}(\mathbf{r}, t)}{\partial t} = \frac{e}{m} \mathbf{E}(\mathbf{r}, t) - \mathbf{u}(\mathbf{r}, t), \quad (3)$$

де $\mathbf{E}(\mathbf{r}, t)$ і $\mathbf{D}(\mathbf{r}, t)$ — вектори напруженості та індукції електричного поля в точці з радіус-вектором $\mathbf{r}$ в момент часу t ; e і m — заряд і ефективна маса електрона; $n(\mathbf{r}, t)$ — густина електронів пучка; ε_0 , N_0 , ν і $\mathbf{u}(\mathbf{r}, t)$ — діелектрична стала кристалічної ґратки, рівноважна концентрація електронів, частота зіткнень і швидкість електронів твердотільної плазми відповідно. Густина $n(\mathbf{r}, t)$ визначається як $n(\mathbf{r}, t) = n(z, t) \delta(\rho - \rho_0)$, де $n(z, t)$ — поверхнева густина електронів пучка; $\delta(x)$ — дельта-функція Дірака.

Введемо потенціал електричного поля $\psi(\rho, z, t)$, який внаслідок статичного характеру поля задовольняє умові $\mathbf{E}(\rho, z, t) = -\nabla \psi(\rho, z, t)$. На поверхні циліндра $\rho = \rho_0$ потенціал $\psi(\rho, z, t)$ є безперервним, а його похідна за радіальною координатою $\partial \psi(\rho, z, t) / \partial \rho$ стрибкоподібно змінюється внаслідок наявності поверхневого заряду пучка $en(z, t)$. Надалі обмежимося розглядом азимутально-симетричних електростатичних хвиль і представимо потенціал $\psi(\rho, z, t)$ у вигляді $\psi(\rho, z, t) = \psi_0(\rho, z, t) \psi_A(t) \exp\{i[q_z z - \omega t + \alpha(t)]\}$, де q_z і ω — хвильовий вектор і частота власної електростатичної хвилі циліндра, які у разі резонансної вавілова–черенковської взаємодії між пучком та хвилею зв'язані зі швидкістю електронів пучка умовою $\omega = q_z v_0$. Величини $\psi_A(t)$ і $\alpha(t)$ являють собою амплітуду і фазу, що повільно змінюються у часі, електростатичної хвилі. Відповідними умовами «повільності» у разі, коли виконується умова сильного зіткнення $\nu |\mathbf{u}(\mathbf{r}, t)| \gg |\partial \mathbf{u}(\mathbf{r}, t) / \partial t|$, є

$$\frac{1}{\psi_A(t)} \frac{\partial \psi_A(t)}{\partial t} \ll \omega, \quad \frac{1}{|\alpha(t)|} \left| \frac{\partial \alpha(t)}{\partial t} \right| \ll \omega. \quad (4)$$

Беручи до уваги обране зображення для $\psi(\rho, z, t)$ із рівнянь (2) і (3) у разі сильних зіткнень маємо $\mathbf{D}(\mathbf{r}, \omega) = \varepsilon(\omega) \mathbf{E}(\mathbf{r}, \omega)$, де $\varepsilon(\omega) = \varepsilon_0 + 4\pi i \sigma / \omega$, $\sigma = e^2 N_0 / m \nu$ — провідність твердотільної плазми циліндра. Згідно роботі [2], поверхнева густина $n(z, t)$ може бути зображена як

$$n(z, t) = \frac{2\pi n_0 a}{qM} \sum_{p=1}^M \delta[z - z_p(z_0, v_0, t)], \quad (5)$$

де $q = |q_z|$. Координата $z_p(z_0, v_0, t)$ відображає положення окремої p -ої макрочастинки. Вираз (5) зображує той факт, що безперервний трубчастий потік електронів має вигляд набору макрочастинок (у нашому випадку — заряджених кілець), кількість яких на довжині хвилі дорівнює M . Дотримуючись методик роботи [2], отримуємо наступну самоузгоджену систему рівнянь:

$$\begin{cases} \frac{d\phi}{d\tau} = -A\phi - \frac{B}{M} \sum_{p=1}^M \sin(\xi_p + \alpha), \\ \frac{d\alpha}{d\tau} = 1 - \frac{B}{\phi M} \sum_{p=1}^M \cos(\xi_p + \alpha), \\ \frac{d\xi_p}{d\tau} = \eta_p, \\ \frac{d\eta_p}{d\tau} = C\phi \sum_{p=1}^M \sin(\xi_p + \alpha), \end{cases} \quad (6)$$

де $\phi = \psi_A(t)/\psi_{\max}$, $\psi_{\max}$ — максимальне значення потенціалу поля хвилі, при якому здійснюється захоплення електронів пучка в потенційну яму хвилі; $\tau = \omega t$; $\xi_p = q_z \tilde{z}_p(t)$; $\eta_p = \tilde{v}_{zp}/v_0$; $\tilde{z}_p(t)$ і $\tilde{v}_{zp}(t)$ — збурення координати та швидкості p -ої макрочастинки в системі відліку пучка; $A = \omega_p^2 \Delta_1 / \nu \omega$; $B = \sqrt{1 + A^2 / \Delta_0^2}$; $C = a q \omega_b^2 / \omega^2 \sqrt{A^2 + \Delta_0^2}$; $\omega_p = \sqrt{4\pi e^2 N_0 / m}$ — частота ленгмюрівських коливань в середовищі циліндра; $\omega_b = \sqrt{4\pi e^2 n_0 / m_0}$ — плазмова частота пучка, рівноважна концентрація електронів якого дорівнює n_0 ; m_0 — маса вільного електрона; $\Delta_0 = \varepsilon_0 I'_0(q\rho_0)/I_0(q\rho_0) - K'_0(q\rho_0)/K_0(q\rho_0)$; $\Delta_1 = I'_0(q\rho_0)/I_0(q\rho_0)$; $I_0(q\rho)$ і $K_0(q\rho)$ — модифіковані функції Бесселя нульового порядку першого та другого роду (функції Інфельда та Макдональда) відповідно. «Штрих» вказує на похідну функції за аргументом.

Чисельний аналіз системи рівнянь (6) проведемо для електродинамічної системи із циліндричного напівпровідника GaAs з параметрами: $\varepsilon_0 = 13.2$; $m = 0.063 m_0$; $\nu = 10^{13} \text{ s}^{-1}$ і $N_0 = 10^{13} \text{ cm}^{-3}$ та трубчастого електронного пучка, що обдуває циліндр, з параметрами: $n_0 = 10^{10} \text{ cm}^{-3}$, $a = 0.01 \text{ cm}$ і $v_0 = 0.1 c$.

Рис. 1 відображає залежності $\phi \equiv \phi(\tau)$ в електродинамічних системах з радіусами $\rho_0 = 0.5 \text{ cm}$ (залежність 1), $\rho_0 = 1 \text{ cm}$ (залежність 2) і $\rho_0 \rightarrow \infty$ (залежність 3). Відмітимо, що залежність 3 відповідає системі з плоскою межею розподілу середовищ (тобто системі, в якій стрічковий електронний пучок транспортується уздовж плоскої поверхні напівпровідникового півпростору).

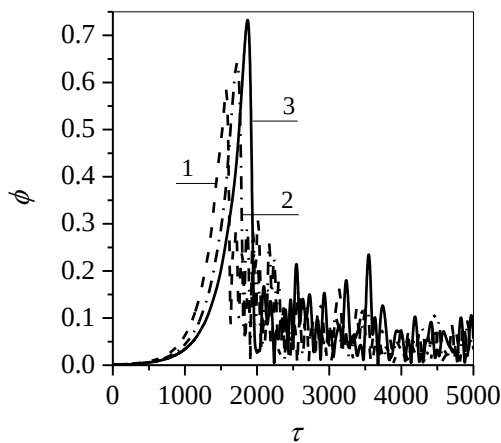


Рис. 1. Залежності безрозмірної амплітуди ϕ потенційної хвилі від безрозмірного часу τ в електродинамічних системах з плазмовими циліндрами, радіуси яких $\rho_0 = 0.5 \text{ cm}$ (1), $\rho_0 = 1 \text{ cm}$ (2) і $\rho_0 \rightarrow \infty$ (3) відповідно

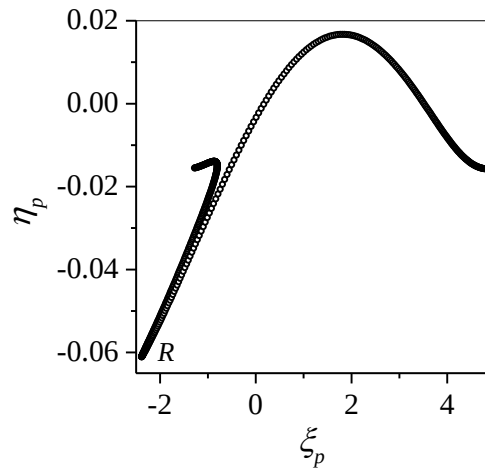


Рис. 2. Фазовий портрет електродинамічної системи, електронний пучок якої складається з M макрочастинок, у момент часу, коли настає насичення нестійкості ($\tau = 1570$)

Чисельний розв'язок системи рівнянь (6) було виконано з застосуванням методу Рунге–Кутти. Пучок електронів представлявся у вигляді набору макрочастинок (у формі кілець), які в початковий момент часу були рівномірно розподілені у проміжку $0 \leq \xi_p \leq 2\pi$. Загальна кількість макрочастинок $M = 500$. Початкові значення безрозмірної амплітуди ϕ покладалось рівним $\phi_0 = 10^{-4}$, а повільної фази α та її похідної $d\alpha/d\tau$ — рівними нулю.

З рис. 1 випливає, що чим більший радіус електродинамічної системи, тим більше значення максимальної амплітуди потенціалу $\phi_{\max}$ поля азимутально-симетричної електростатичної хвилі, у момент часу якого настає насиченість резонансної нестійкості електронного пучка. В електродинамічній системі з $\rho_0 = 0.5 \text{ cm}$ маємо $\phi_{\max} \approx 0.58$, а з $\rho_0 = 1 \text{ cm}$

— $\phi_{\max} = 0.64$. Але в системі з меншим радіусом повільна амплітуда потенціалу поля ϕ досягає свого максимального значення $\phi_{\max}$ раніше (коли $\tau \equiv \tau_{\max} \approx 1569$, що відповідає часу $t \equiv t_{\max} \approx 16\gamma^{-1}$, де $\gamma \approx 10^8 \text{ s}^{-1}$ — інкремент нестійкості), ніж в системі з більшим радіусом (коли $\tau \equiv \tau_{\max} \approx 1718$, тобто $t \equiv t_{\max} \approx 17\gamma^{-1}$). Крім того, зі зростанням радіусу електродинамічної системи ρ_0 залежності $\phi(\tau)$ наближаються до залежності 3, яка відповідає випадку плоско-поверхневої межі стрічковий пучок–твердотільна плазма. Також, коли $\tau > \tau_{\max}$, значення потенціалу ϕ поля поверхневої електростатичної хвилі циліндричного твердотільно-плазмового хвилеводу осциляційно змінюються навколо свого середнього значення. Ці флуктуації потенціалу ϕ мають суттєво нерегулярний (хаотичний) характер. Це зумовлено тим, що на момент захоплення електронів пучка полем електростатичної хвилі кільцеві макрочастинки не встигають зібратися у згустки внаслідок розкиду за швидкостями та координатами.

Розглянута електродинамічна система, до складу якої входить твердотільна плазма, принципово відрізняється від системи роботи [2] тим, що сильнозіштовхувальна плазма не підтримує власні хвилі (коливання) циліндра. У зв'язку з цим захоплення електронів пучка здійснюється полем електростатичної хвилі самого пучка. На рис. 2 фазовий портрет електродинамічної системи з радіусом $\rho_0 = 0.5 \text{ см}$ демонструє ефект самозахоплення електронів пучка полем пучкової хвилі та їх хаотизацію. Звернемо увагу, що група відбитих макрочастинок, яка позначена буквою R, з'являється на момент насичення нестійкості. Відбиття макрочастинок здійснюється електричним полем потенційної хвилі самого пучка.

Висновки

У роботі виконано чисельне моделювання методом макрочастинок нелінійної стабілізації нестійкості трубчастого електронного пучка у разі його транспортування уздовж бокової поверхні твердотільного плазмового циліндра. Показано, що нелінійна стабілізація резистивної нестійкості, що виникає в електродинамічній системі, відбувається внаслідок самозахоплення електронів полем пучкової хвилі. Встановлено, що нелінійна стадія нестійкості починається пізніше у системі з більшим радіусом твердотільно-плазмового циліндра.

ЛІТЕРАТУРА

1. Лопухин В. М., Веденов А. А. Усилитель на поглощении. *Успехи физических наук*. 1954. Т. 53, Вып. 1. С. 69–86.
2. Иванов А. А. Физика сильно неравновесной плазмы : монография. Москва : Атомиздат, 1977. 352 с.

АВЕРКОВ Олег Юрійович – студент 4-го курсу факультету комп'ютерних наук ХНУ імені В.Н. Каразіна; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: xa12850341@student.karazin.ua; olegaverkov072@gmail.com; ORCID: 0000-0001-6055-015X.

Наукові інтереси:

– Алгоритми доказу та верифікації доказу транзакцій блокчейн.

АВЕРКОВ Юрій Олегович – д. ф.-м. н.; професор кафедри моделювання систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: yuriyaverkov@karazin.ua; yuriyaverkov@gmail.com; ORCID: [0000-0002-1169-9393](https://orcid.org/0000-0002-1169-9393).

Наукові інтереси:

– взаємодія заряджених часток та електромагнітних полів в неоднорідних провідних середовищах.

АВЕРКОВ Ю.О., АВЕРКОВ О.Ю.

ОСОБЛИВОСТІ СПЕКТРУ НАПІВПРОВІДНИКОВОЇ НАНОТРУБКИ В ЗОВНІШНЬОМУ МАГНІТНОМУ ПОЛІ

Topic of the article. The topic of the work is theoretical study of the eigenmode spectra of a semiconductor nanotube with dielectric filling placed in an external magnetic field.

Purpose of the work. The purpose of the work is to reveal an influence of the quantum nature of the conducting properties of the semiconductor nanotubes on the eigenmode spectra of a semiconductor nanotube with dielectric filling placed in an external magnetic field.

Research methods. It is assumed that the nanotube is infinitely thin in the radial direction and its electron gas is degenerated one. To derive the dispersion equation of the structure under study we use the methods of classical electrodynamic for solving the Maxwell equations with the corresponding boundary conditions on the lateral surface of the nanotube. To obtain the formula that gives the connection of the equilibrium surface electron density in the nanotube with the Fermi energy and the energies of the filled subzones we use methods of quantum mechanics.

Results. The dispersion equation for the eigenmodes spectra of the structure under study has been derived and numerically analyzed. The dispersion curves of the hybrid eigenmodes is shown to have parts with negative dispersion and the number of dispersion branches as well as the equilibrium surface electron density in the nanotube oscillate with growth of the number of magnetic flux quanta through the nanotube.

Conclusions. It has been demonstrated that the surface currents in the nanotube have an important bearing on the spectrum of eigenmodes in the structure under study. The presence of the parts of the dispersion curves of the hybrid eigenmodes with negative dispersion signifies the semiconductor nanotubes can be used as the delaying medium in nanoscale oscillators of electromagnetic radiation. The revealed oscillations of the number of dispersion branches and the equilibrium surface electron density with growth of the number of magnetic flux quanta through the nanotube are the manifestation of the Aharonov-Bohm effect.

Keywords: Aharonov-Bohm effect, semiconductor, nanotubes, eigenmodes spectra, hybrid modes

Вступ

Напівпровідникові нанотрубки створюються з атомів кремнію, а елементарна комірка таких структур подібна до бджолиних стільників. Напівпровідникові нанотрубки є перспективним матеріалом для створення нової елементної бази наноелектроніки завдяки своїй здатності підтримувати поширення поверхневих електромагнітних хвиль – плазмон-поляритонів [1]. У цій роботі буде проведено чисельне моделювання спектру напівпровідникової нанотрубки, яка заповнена немагнітним діелектриком і знаходиться в постійному магнітному полі.

Постановка задачі й результати моделювання

Припускаємо, що немагнітний діелектричний циліндр радіусу ρ_c займає область $0 \leq \rho \leq \rho_c$, $0 \leq \varphi \leq 2\pi$, та $|z| < \infty$ (де вісь z спрямована уздовж осі циліндру). Циліндр виготовлений з ізотропного немагнітного матеріалу з реальною діелектричною проникністю ε . На бічній поверхні циліндру знаходиться нескінченно тонкий напівпровідниковий шар. Хвилевід розташований у вакуумі в зовнішньому постійному магнітному полі, вектор, якого $\mathbf{H}_0$ спрямований паралельно до його осі симетрії. Система рівнянь, що описує таку електродинамічну модель є система рівнянь Максвелла, яка доповнена рівнянням неперервності для густини поверхневого заряду та густини поверхневого струму

$$\operatorname{rot} \mathbf{H}(\mathbf{r}, t) - \frac{1}{c} \frac{\partial}{\partial t} \mathbf{D}(\mathbf{r}, t) = \frac{4\pi}{c} \mathbf{j}(\mathbf{r}, t); \quad \operatorname{div} \mathbf{D}(\mathbf{r}, t) = 4\pi e N(\mathbf{r}, t); \quad (1)$$

$$\operatorname{rot} \mathbf{E}(\mathbf{r}, t) + \frac{1}{c} \frac{\partial}{\partial t} \mathbf{H}(\mathbf{r}, t) = 0; \quad \operatorname{div} \mathbf{H}(\mathbf{r}, t) = 0, \quad (2)$$

$$e \frac{\partial n^{2D}(\varphi, z, t)}{\partial t} + \operatorname{div} \mathbf{j}^{2D}(\varphi, z, t) = 0. \quad (3)$$

де e – заряд електрону, c – швидкість світла у вакуумі, $\mathbf{H}(\mathbf{r}, t)$ – напруженість магнітного поля, $N(\mathbf{r}, t) = n^{2D}(\varphi, z, t) \delta(\rho - \rho_c)$, $n^{2D}(\varphi)$ – густина поверхневого заряду в нанотрубці,

$\mathbf{j}(\mathbf{r}, t) = (0, \mathbf{j}_r(\mathbf{r}, t))$, $\mathbf{j}_r(\mathbf{r}, t) = (j_\varphi(\mathbf{r}, t), j_z(\mathbf{r}, t))$, $\mathbf{j}_r(\mathbf{r}, t) = \mathbf{j}_r^{2D}(\varphi, z, t)\delta(\rho - \rho_c)$, $\mathbf{j}_r^{2D}(\varphi, z, t)$ – густина поверхневого струму уздовж нанотрубки, $\delta(x)$ – дельта-функція Дірака. Вектор електричної індукції $\mathbf{D}(\mathbf{r}, t)$ зв'язаний з електричним вектором $\mathbf{E}(\mathbf{r}, t)$ наступним рівнянням:

$$D_q(\mathbf{r}, t) = \int_{-\infty}^t \hat{\varepsilon}(t-t')E_p(\mathbf{r}, t')dt', \quad (4)$$

де $\hat{\varepsilon}(t-t')$ – функція впливу. Для однорідного та ізотропного діелектричного циліндру маємо $\hat{\varepsilon}(t-t') = \varepsilon\delta_{qp}\delta(t-t')$, де δ_{qp} символ Кронекера.

Щоб знайти рівняння дисперсії хвиль в системі, що досліджується, необхідно скористатися граничними умовами для електричних і магнітних полів при $\rho = \rho_c$. Це наступні умови: при $\rho = \rho_c$ тангенціальні компоненти електричного поля неперервні, тангенціальні компоненти магнітного поля зазнають розриву, що обумовлений поверхневим струмом у нанотрубці. Зауважимо, що при $\rho = \rho_c$ нормальна компонента магнітної індукції залишається неперервною, а нормальна компонента електричної індукції зазнає розриву, що обумовлений збуренням зарядом провідного шару.

Ми представляємо усі змінні як набір просторово-часових гармонік, наприклад:

$$\mathbf{E}(\mathbf{r}, t) = \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{E}_n(\rho, q_z, \omega) \exp[i(q_z z + n\varphi - \omega t)] dq_z d\omega, \quad (5)$$

де ω, q_z і n – частота, поздовжнє хвильове число та номер просторової гармоніки (співпадає з азимутальним модовим індексом), відповідно.

Задовольняючи вказаним вище граничним умовам на поверхні циліндру з урахуванням нерелятивістської природи руху носіїв заряду в нанотрубці, отримуємо наступне дисперсійне рівняння для власного спектру напівпровідникової нанотрубки з діелектричним заповненням:

$$\Delta(n, q_z, \omega) = \frac{4\pi i}{\omega \rho_c} \left[\frac{n^2}{\rho_c^2 q^2} \sigma_{\varphi\varphi}(n, q_z, \omega) + \sigma_{zz}(n, q_z, \omega) \right] \Delta^H(n, q_z, \omega), \quad (6)$$

де $\sigma_{\alpha\beta}$ ($\alpha, \beta = \varphi, z$) – компоненти тензора діелектричної проникності напівпровідникової нанотрубки, які наведені в роботі [2],

$$\Delta(n, q_z, \omega) = \Delta^E(n, q_z, \omega) \Delta^H(n, q_z, \omega) - \left[\frac{n q_z \omega (\varepsilon - 1)}{q_1^2 |q_2|^2 \rho_c^2 c} \right]^2; \quad (7)$$

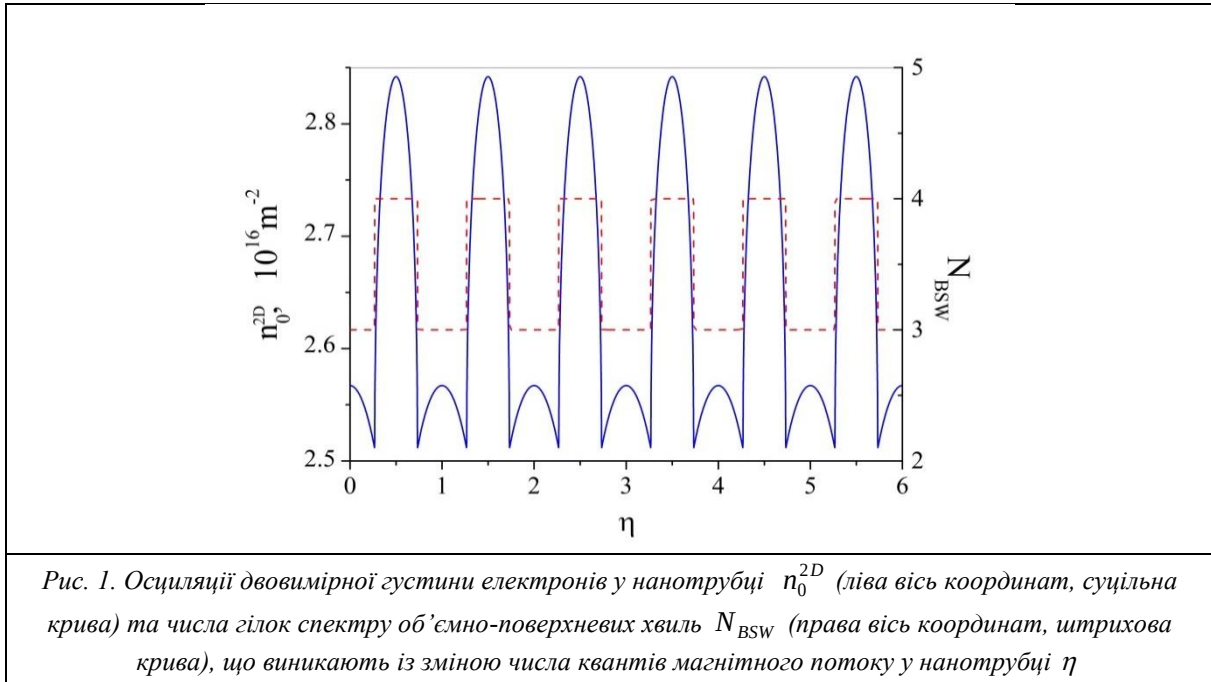
$$\Delta^E(n, q_z, \omega) = \frac{1}{|q_2| \rho_0} \frac{K'_n(|q_2| \rho_c)}{K_n(|q_2| \rho_c)} + \frac{\varepsilon}{q_1 \rho_c} \frac{J'_n(q_1 \rho_c)}{J_n(q_1 \rho_c)}; \quad (8)$$

$$\Delta^H(n, q_z, \omega) = \frac{1}{|q_2| \rho_0} \frac{K'_n(|q_2| \rho_c)}{K_n(|q_2| \rho_c)} + \frac{1}{q_1 \rho_c} \frac{J'_n(q_1 \rho_c)}{J_n(q_1 \rho_c)}; \quad (9)$$

Зауважимо, що рівняння $\Delta(n, q_z, \omega) = 0$ описує власний спектр діелектричного циліндру.

Оскільки компоненти тензора провідності нанотрубки зазнають коливання зі зміною числа квантів магнітного потоку у нанотрубці (див. [2]), права сторона дисперсійного рівняння (6) також зазнає подібні коливання. Тому, для заданого значення хвильового числа q_z розв'язки дисперсійного рівняння (6) відносно частоти також будуть зазнавати подібні осциляції. Чисельний аналіз дисперсійного рівняння (6) показує, що зі зміною числа квантів магнітного потоку в нанотрубці η мають місце не тільки осциляції частоти власних мод, але й числа дисперсійних гілок спектру об'ємно-поверхневих хвиль N_{BSW} одночасно з двовимірною концентрацією електронів в нанотрубці n_0^{2D} . Ці осциляції демонструє рис. 1 побудований для

наступних параметрів структури: $\rho_c = 30 \text{ nm}$, $\varepsilon = 9.8$, у якості матеріалу напівпровідника обраний напівпровідник GaAs з енергією Фермі $\varepsilon_F = 6.47 \text{ meV}$.



Двовимірна концентрація електронів в нанотрубці n_0^{2D} обчислюється за наступною формулою:

$$n_0^{2D} = \frac{\sqrt{2m^*}}{2\pi^2\hbar\rho_c} \sum_m \sqrt{\varepsilon_F - \varepsilon_m}, \quad (10)$$

де m^* – ефективна маса електронів у напівпровіднику, $\hbar$ – постійна Планка, $\varepsilon_m = \varepsilon_0(m + \eta)^2$ – енергія електронів в підзоні з номером $m = 0, \pm 1, \pm 2, \dots$, ε_0 – ротаційний квант.

З рис. 1 можна побачити, що величини n_0^{2D} і N_{BSW} осцилюють із зміною величини η з періодом, що дорівнює одиниці. Такі осциляції можна вважати проявом ефекту Ааронова-Бома. Наявність коливань частоти власних хвиль спектру досліджуваної структури можна наочно продемонструвати для азимутально симетричних хвиль ($n = 0$) у випадку виродженого електронного газу, коли задовольняються наступні умови:

$$q_z v_F \ll \omega, \quad \varepsilon_F \gg \varepsilon_0, \quad (11)$$

де v_F – швидкість Фермі. Якщо обмежитись розглядом тільки імовірної частини компоненти тензора провідності нанотрубки $\sigma_{zz}(n, q_z, \omega)$ (див. [2]), отримаємо наступне дисперсійне рівняння:

$$\Delta^E(0, q_z, \omega) = -\frac{4\pi e^2 n_0^{2D}}{m\omega^2 \rho_c} \left[1 + \frac{2}{\pi^2} \left(\frac{\varepsilon_0}{\varepsilon_F} \right)^{3/4} \sum_{\ell=1}^{\infty} \frac{1}{\ell^{3/2}} \sin \left(2\pi\ell \sqrt{\frac{\varepsilon_F}{\varepsilon_0}} \frac{\pi}{4} \right) \cos(2\pi\ell\eta) \right]. \quad (12)$$

З виразу (11) можна побачити, що права частина дисперсійного рівняння осцилює із зміною числа квантів магнітного потоку у нанотрубці η .

У квазістаціонарному наближенні (коли $c \rightarrow \infty$) та для $\varepsilon = 1$ дисперсійне рівняння (6) набирає наступного вигляду:

$$\frac{4\pi i}{\omega \rho_c} I_n(|q_z| \rho_c) K_n(|q_z| \rho_c) \left[n^2 \sigma_{\varphi\varphi}(n, q_z, \omega) + \rho_c^2 q_z^2 \sigma_{\varphi\varphi}(n, q_z, \omega) \right] = -1 \quad (13)$$

(де $I_n(x)$ – модифікована функція Бесселя n -го порядку першого роду (функція Інфельда)) та описує повільні гібридні поверхневі хвилі в напівпровідниковій нанотрубці. Зауважимо, що поклавши в рівнянні (6) $n=0$ та $\hbar \rightarrow 0$ (тобто нехтуючи квантовими ефектами), ми отримаємо відоме дисперсійне рівняння для симетричних поверхневих мод нанотрубки:

$$\omega^2 = \frac{4\pi e^2 n_0^{2D} \rho_c q_z^2}{m^*} I_0(|q_z \rho_c|) K_0(|q_z \rho_c|), \quad (14)$$

детальний аналіз якого виконаний у роботі [3].

Висновки

У цій роботі були розглянуті власні електромагнітні моди напівпровідникової нанотрубки із діелектричним заповненням. Було отримано дисперсійне рівняння для власного спектру досліджуваної структури з урахуванням ефекту запізнення. Це рівняння описує дисперсійні властивості як поверхневих, так і об'ємно-поверхневих хвиль. У випадку коли азимутальне хвильове число відмінно від нуля, відповідні власні хвилі є гібридними модами, тобто містять усі компоненти електричних і магнітних полів. За допомогою чисельно аналізу встановлено, що частоти власних гібридних хвиль спектру, число гілок спектру а також двовимірна густина електронів у нанотрубці зазнають осциляції із зміною числа квантів магнітного потоку. Ці осциляції можна вважати проявом ефекту Ааронова-Бома.

ЛІТЕРАТУРА

- McNeish T., Gumbs G, Balassis A. Model of plasmon excitations in a bundle and a two-dimensional array of nanotubes. *Physical Review B*. 2008. Vol. 77, Issue 23. P. 235440(11).
- Ermolaev A. M., Kofanov S. V., Rashba G. I. Electron gas dynamic conductivity tensor on the nanotube surface in magnetic field. *Advances in Condensed Matter Physics*. 2011. Vol. 2011, P. 1–7.
- Ведерников А. И., Говоров А. О., Чаплик А. В. Плазменные колебаний в нанотрубках и эффект Ааронова-Бома для плазмонов. *Журнал экспериментальной и теоретической физики*. 2001. т. 120, Вып. 4(10). С. 979–985.

АВЕРКОВ Юрій Олегович – д. ф.-м. н.; професор кафедри моделювання систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: yuriyaverkov@karazin.ua; yuriyaverkov@gmail.com; ORCID: [0000-0002-1169-9393](https://orcid.org/0000-0002-1169-9393).

Наукові інтереси:

– взаємодія заряджених часток та електромагнітних полів в неоднорідних провідних середовищах.

АВЕРКОВ Олег Юрійович – студент 4-го курсу факультету комп'ютерних наук ХНУ імені В.Н. Каразіна; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: xa12850341@student.karazin.ua; olegaverkov072@gmail.com; ORCID: 0000-0001-6055-015X.

Наукові інтереси:

– Алгоритми доказу та верифікації доказу транзакцій блокчейн.

ОСОБЛИВОСТІ ВИКОРИСТАННЯ ЦИФРОВИХ ДВІЙНИКІВ

Появі цифрових двійників сприяла потреба інженерів/розробників здійснювати постійний цифровий/аналоговий моніторинг та функціональне вдосконалення взаємопов'язаних модулів систем. Переваги вертикальної та горизонтальної масштабованості в розробці складних систем націлені на впровадження цифрових двійників.

Ключові слова: цифровий двійник, розробка систем, моделювання, віртуалізація.

Вступ

Програмні продукти еволюціонують до кіберфізичних систем та кіберфізичних виробничих систем, які мають у своєму складі багато механічних, електричних, цифрових, програмних та комунікаційних заводських модулів, які під'єднані до зовнішнього середовища, із яким вони взаємодіють [3]. Концепція 4-ї промислової революції висуває ряд вимог до горизонтальної та вертикальної інтеграції, такі аналіз, захист та швидку обробку великих масивів даних, отриманих від смарт датчиків, виконавчих механізмів pl модулів. Цифрові технології та постійне оновлення мов програмування стають все більш популярною темою для досліджень та дають змогу розробити програмні прототипи заводських модулів, щоб зрозуміти природу поведінки та можливості, що виходять за кордони експлуатаційних характеристик, як у традиційному технічному обслуговуванні, плановому ремонті, капітальному ремонті або виробництві в промисловості.

Сучасний стан технологій

Простим прикладом використання цифрового двійника (ЦД) є моніторинг поведінки комп'ютерної моделі заводського смарт датчика температури або електричного струму. Цифровий двійник включає характеристики та максимально наближену поведінку реального продукту через дані та моделі [3]. Поведінка реального продукту може бути змодельована за допомогою цифрового двійника в наданому програмному середовищі. Ця симуляція є незалежною від часу та місця фізичного продукту. Основною перевагою двійника є незалежна від місця і часу симуляція поведінки реального продукту. За допомогою ЦД можна скоротити експерименти з реальними продуктами і, отже, заощадити час і фінансові витрати в процесі розробки продукту [2]. ЦД визначається як цифрове представлення діючого продукту, послуги, процесу або виробничої системи, що характеризується певними властивостями або умовами, які використовуються з метою аналізу, розуміння та вдосконалення продукту, системи обслуговування продукту або виробництва.

Представлення цифрового двійника має наступні аспекти:

- маніпуляція різними ступенями функцій інформаційних технологій такі, як тумані, хмарні або периферійні обчислення або обчислювальні середовища штучного інтелекту із метою безпроблемної інтеграції нових двійників;
- вихідні дані надають значущі об'єми інформації, які за допомогою інтелектуальної обробки можуть використовуватись у процесах виробництва, операції із програмними продуктами, тощо;
- роль цифрового двійника полягає у розкритті відповідної семантики з метою аналізу, управління та вдосконалення технічних систем від простих рівнів аналізу та управління до автономних взаємодій між підсистемами [2].

Важливе значення в розробці ЦД мають бізнес-моделі, що стимулюють зусилля компаній, які часто зосереджуються на розвитку та вдосконаленні послуг, наприклад, у сфері хмарних обчислень. Одним із основних факторів, що сприяє використанню цифрових двійників, є потреба у віртуалізації, включаючи відповідні властивості зберігання та обчислення, що пропонуються у вигляді програмних продуктів, наприклад, таких як, Microsoft Azure, Amazon AWS.

Інженери та проектувальники можуть використовувати ЦД у наступних ситуаціях:

- як орієнтир у постановці цілей при розробці абсолютно нового програмного продукту на базі існуючого двійника;
- розширення знань про поведінку двійника в екстремальних або за подібних умов експлуатації;
- для модернізації та розвитку двійника, як продукту або послуги, додаючи новий функціонал.

Збір та обробка даних про життєвий цикл допомагає зрозуміти поведінку компонента та адаптувати характеристики аналогічних компонентів на майбутніх етапах розробки програмного продукту. ЦД дають змогу пов'язати реальні впливи навколишнього середовища зі змінами своїх робочих характеристик. У майбутньому це дозволяє підвищити якість та надійність виробів, а також продовжити термін їх експлуатації. На основі даних, зібраних протягом життєвого циклу системи, можна визначити нестандартний випадок використання та розробити майбутні продукти із оновленим функціоналом відповідно до потреб замовника [1].

Підвищення стабільності системи

Цифровий двійник має середовище для проведення оцінки свого життєвого циклу на основі отриманих емульованих даних. Замість статичної оцінки життєвого циклу, індикатори можуть бути розроблені та інтегровані в ЦД. Ці індикаторів забезпечують специфічну для продукту та актуальну оцінку ресурсів відповідно до поведінки фізичного двійника [3]. За допомогою імітаційного моделювання в цифровому двійнику користувач може приймати рішення щодо внесення налаштувань до майбутньої поведінки фізичного двійника. Поведінка фізичного двійника адаптується ЦД відповідно до налаштувань користувача.

Висновок

Цифровий двійник це цифрове утілення фізичного продукту, яке включає в себе його поведінку за допомоги імітаційних моделей. Результати моделювання представляються кінцевому користувачу, що дозволяє реалізувати адаптацію фізичного продукту в систему або виробництво через ЦД. Використання двійників в індустрії інформаційних технологій дозволяє підвищити стійкість та надійність системи у цілому.

ЛІТЕРАТУРА

1. G.L.Knapp Building blocks for a digital twin of additive manufacturing / G.L.Knapp, T.Mukherjee, J.S.Zuback, H.L.Wei, T.A.Palmer, A.De, T.DebRoy // Acta Materialia — Vol. 135. — 15 August 2015. — P.390-399.
2. Cordula Czwick Cyber-physical twins - definition, conception and benefit / Cordula Czwick, Reiner Anderl // Procedia CIRP — Vol. 90. — 2020. — P.584-588.
3. Rainer Stark Development and operation of Digital Twins for technical systems and services / Rainer Stark, Carina Freseemann, Kai Lindow // CIRP Annals — Vol. 68. — Is. 1 — 2019. — P.129-132.

АРТЮХ Сергій Олегович – аспірант 3-го курсу факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: s.artukh96@gmail.com; ORCID: 0000-0002-5856-3741.

Наукові інтереси:

- *автоматизовані системи керування технологічними процесами та програмне забезпечення для моделювання окремих модулів цих систем.*

ХРУСЛОВ Максим Михайлович – виконуючий обов'язки завідувача кафедри електроніки і управляючих систем факультету комп'ютерних наук, доцент, кандидат фізико-математичних наук; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: maksym.khruslov@karazin.ua; ORCID: 0000-0001-9639-9340.

Наукові інтереси:

- *автоматизовані системи керування технологічними процесами та програмне забезпечення для моделювання окремих модулів цих систем.*

БЕЗСМЕРТНИЙ Д.Р., КОНДРАТЮК Д.С.

МЕТОДИ НАВЧАННЯ РЕКУРЕНТНОЇ ТА ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ НА ОСНОВІ ПОПУЛЯЦІЙНОГО АЛГОРИТМУ ДЛЯ РОЗПІЗНАВАННЯ ОБРАЗІВ

Тема роботи. Методи навчання рекурентної та згорткової нейронної мережі на основі популяційного алгоритму для розпізнавання образів.

Мета роботи. Метою дослідження є синергія популяційного алгоритму з рекурентною та згортковою нейронною мережею. Та розробка цих програмних моделей.

Методи дослідження. Методи побудови нейронних мереж, методи розпізнавання образів, основи популяційних алгоритмів.

Результати. Результатом виконання роботи є дві програмні моделі рекурентної та згорткової нейронної мережі на основі популяційного алгоритму для розпізнавання образів.

Висновки. У роботі була розглянута загальна схема популяційного алгоритму, особливості розпізнавання образів, та було розроблено дві програмні моделі нейронної мережі на основі популяційного алгоритму для розпізнавання образів. Особливість розробки була у синергії популяційного алгоритму та нейронної мережі, популяційний алгоритм допомагає навчатися нейронній мережі. У випадку рекурентної нейронної мережі популяційний алгоритм допомагає налаштувати ваги, а у випадку згорткової допомагає знаходити чотири параметри нейронної мережі. Також додатково для порівняння було розроблено дві моделі розпізнавання образів на основі генетичного алгоритму. Вже на цьому етапі результати показали, що моделі на основі популяційного алгоритму мають перевагу у якості розпізнавання.

Ключові слова: навчання нейронної мережі, рекурентна нейронна мережа, згорткова нейронна мережа, розпізнавання, популяційний алгоритм.

Є багато методів навчання згорткової та рекурентної нейронної мережі для розпізнавання образів. В наш час розпізнавання образів набирає обертів, основна суть поставленого завдання є допомога навчитися нейронній мережі за допомогою популяційного алгоритму. Основна проблема цього завдання полягає в самій ідеї поєднати ці два алгоритми, тому що готові патерни були знайдені з іншими нейронними мережами та з іншим популяційними алгоритмами. Після аналізу знайденої інформації пропонується в рекурентній нейронній мережі за допомогою популяційного алгоритму налаштовувати ваги LSTM шару, а в згортковій нейронній мережі налаштовувати один з шарів. Також для порівняння кожного алгоритму було розроблено дві моделі на основі генетичного алгоритму. В рекурентній та згортковій нейронній мережі генетичний алгоритм допомагає побудувати саму структуру шару.

```
Population Initialization Done: [128-64, 32-512-256-32, 64-512, 64-512-256-512-32-64-max-128-32, 128-128]
Generation 0
Evaluating Population fitness
[128-64]
Epoch 1/5
750/750 [=====] - 18s 12ms/step - loss: 2.7984 - accuracy: 0.9178 - val_loss: 0.1416 - val_accuracy: 0.9745
Epoch 2/5
750/750 [=====] - 9s 11ms/step - loss: 0.0727 - accuracy: 0.9802 - val_loss: 0.0867 - val_accuracy: 0.9781
Epoch 3/5
750/750 [=====] - 9s 12ms/step - loss: 0.0379 - accuracy: 0.9877 - val_loss: 0.0810 - val_accuracy: 0.9797
Epoch 4/5
750/750 [=====] - 9s 12ms/step - loss: 0.0245 - accuracy: 0.9923 - val_loss: 0.0781 - val_accuracy: 0.9804
Epoch 5/5
750/750 [=====] - 9s 11ms/step - loss: 0.0220 - accuracy: 0.9927 - val_loss: 0.1044 - val_accuracy: 0.9794
157/157 [=====] - 1s 5ms/step - loss: 0.0824 - accuracy: 0.9802
128-64 0.9801999926567078
```

Рис. 1 Навчання з різною структурою шарів згорткової нейронної мережі на основі генетичного алгоритму


```

lstm: (None, 28, 160)
lstm: (None, 124)
dense: (None, 42)
dense: (None, 207)
dense: (None, 173)
last_dense (None, 10)
Training set shape (60000, 28, 28)
1688/1688 [=====] - 194s 110ms/step - loss: 0.4916 - accuracy: 0.8954 - val_loss: 0.2133 - val_accuracy: 0.9638
The first 1 Dai di 1 The fitness of chromosomes is 0.960300
This chromosome is: [2, 3, 160, 124, 42, 207, 173]
lstm: (None, 28, 204)
lstm: (None, 217)
dense: (None, 226)
dense: (None, 191)
dense: (None, 143)
last_dense (None, 10)
Training set shape (60000, 28, 28)
1688/1688 [=====] - 325s 190ms/step - loss: 0.5980 - accuracy: 0.8965 - val_loss: 0.2521 - val_accuracy: 0.9613

```

Рис. 2 Навчання з різною структурою шарів рекурентної нейронної мережі на основі генетичного алгоритму

У випадку популяційного алгоритму, він знаходить коефіцієнти для однієї окремої моделі, тому можна контролювати результати та корегувати параметри для більш кращих результатів.

Особливості розпізнавання образів

Класифікація зображень досягається шляхом диференціації зображення на встановлену категорію на основі змісту бачення. Класифікація зображень є одним із напрямів машинного навчання. Машинне навчання складається з модуля вилучення особливостей, який вилучає такі важливі ознаки, як ребра, текстури тощо, та модуля класифікації, який класифікує на основі вилучених особливостей. Основне обмеження машинного навчання полягає в тому, що, відокремлюючи, воно може витягувати лише певний набір ознак на зображеннях і не в змозі виділити диференційовані ознаки з навчального набору даних. Цей недолік виправляється за допомогою глибокого навчання.

Розпізнавання було проведено на стандартному датасеті MNIST (Modified National Institute of Standards and Technology database)

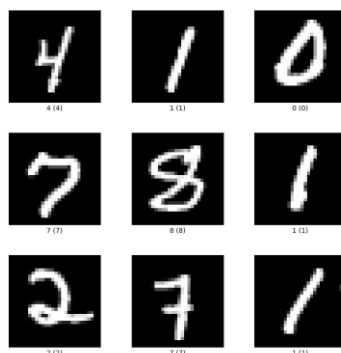


Рис. 3 Датасет MNIST

Моделі розпізнавання образів

На цьому етапі результати роботи цих двох моделей є доволі перспективними, так як вони розпізнають образи з точністю 0,95-0,99. За попередньою оцінкою можна зробити висновок у користь моделі з популяційним алгоритмом який буде в подальшому виконуватись на більш складних датасетах.

```

current config: [15, 7, 2, 4] val: 0.9900000095367432
current config: [12, 5, 3, 3] val: 0.991100013256073

] print('Test loss: ', bp)
  print('Test accuracy: ', value, v)

Test loss: [12, 5, 3, 3]
Test accuracy: 0.9860000014305115 0.991100013256073

```

Рис. 4 Результати моделі згорткової нейронної мережі

Новая наилучшая частица найдена на итерации #496 со среднеквадратической ошибкой: 0.06867739746741808

	precision	recall	f1-score	support
0	0.98	0.90	0.94	52
1	0.00	0.00	0.00	41
2	0.86	0.71	0.78	45
3	0.00	0.00	0.00	44
4	0.92	0.75	0.83	48
5	0.00	0.00	0.00	41
6	0.95	0.81	0.88	48
7	0.00	0.00	0.00	41
8	0.00	0.00	0.00	49
9	0.00	0.00	0.00	41
micro avg	0.93	0.34	0.50	450
macro avg	0.37	0.32	0.34	450
weighted avg	0.40	0.34	0.37	450
samples avg	0.34	0.34	0.34	450

Рис. 5 Результати моделі рекурентної нейронної мережі

Так як додатково до кожної моделі був розроблений алгоритм з генетичним алгоритмом можна провести порівняльний аналіз.

	Модель рекурентної нейронної мережі	Модель згорткової нейронної мережі
Популяційний алгоритм	0,97	0,99
Генетичний алгоритм	0,95	0,97
Середнє значення	0,96	0,98

Висновок

У роботі була розглянута загальна схема популяційного алгоритму, особливості розпізнавання образів, та було розроблено дві програмні моделі нейронної мережі на основі популяційного алгоритму для розпізнавання образів. Особливість розробки була у синергії популяційного алгоритму та нейронної мережі, популяційний алгоритм допомагає навчатися нейронній мережі. У випадку рекурентної нейронної мережі популяційний алгоритм допомагає налаштувати ваги, а у випадку згорткової допомагає знаходити чотири параметри нейронної мережі. Також додатково для порівняння було розроблено дві моделі розпізнавання образів на

основі генетичного алгоритму. Вже на цьому етапі результати показали, що моделі на основі популяційного алгоритму мають перевагу у якості розпізнавання.

ЛІТЕРАТУРА

1. Cervantes A, Galván IM, Isasi P. AMPSO: A New Particle Swarm Method for Nearest Neighborhood Classification. IEEE Trans Syst Man Cybern B Cybern. 2009; 39(5):1082–91.
2. Murphy KP. Machine Learning: A Probabilistic Perspective. Adaptive computation and machine learning series. Cambridge: The MIT Press; 2012.
3. Лосев Ю. І., Руккас К. М., Шматков С. І. Комп'ютерні мережі: навч. посіб. / за редакцією Ю. І. Лосева. Харків : ХНУ імені В. Н. Каразіна, 2013. 248 с.
4. Brunelli R., Poggio T. Face recognition through geometrical features – Computer Vision—ECCV'92. – Springer Berlin/Heidelberg, 1992.
5. Gacnn: training deep convolutional neural networks with genetic algorithm. URL: <https://arxiv.org/pdf/1909.13354.pdf> (Last accessed: 14.11.2022).
6. Particle Swarm Optimization: A survey of historical and recent developments with hybridization perspectives. URL: <https://arxiv.org/ftp/arxiv/papers/1804/1804.05319.pdf> (Last accessed: 14.11.2022).

БЕЗСМЕРТНИЙ Денис Русланович – бакалавр; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: ha11867824@karazin.ua; gerbyboyz@gmail.com; ORCID: 0000-0002-3424-9539.

Наукові інтереси:

– *нейронні мережі*.

КОНДРАТЮК Діана Сергіївна – бакалавр; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: kondrdi2000@gmail.com; ORCID: 0000-0003-0843-544X.

Наукові інтереси:

– *нейронні мережі*.

БЕРЕЖНИЙ А.А.

ПРОГРАМНА МОДЕЛЬ РОБОТА З ДИФЕРЕНЦІЙНИМ ПРИВОДОМ

Актуальність: розвиток робототехніки дуже важливий сучасний напрямок, а особливо платформ, що рухаються, але проблема складності їх реалізації залишається до сих пір. Тому виникає потреба у вдосконалення методів їх реалізації.

Мета: метою роботи є побудова математичної та програмної моделі.

Методи дослідження: у роботі застосовувалися методи диференціального обчислення та програмування у середовищі Wolfram Mathematica.

Результати: побудова програмної моделі робота та візуалізація отриманих даних графічними засобами.

Висновки: головним результатом роботи є розроблені математична та програмна моделі робота, які дозволили розрахувати траєкторію руху двоколісного робота з диференційним приводом..

Ключові слова: кінематична модель, математична модель, диференційний привід, модель робота.

Кінематична модель

В ході виконання роботи, було припущено, що робот складається з ідеальної твердої основи, на яку встановлена співвісна колісна пара з диференціальним приводом. Також вважаємо, що колеса, що управляють, мають точковий контакт з плоскою поверхнею, по якій вони рухаються без проковзування. Стан робота (двоколісної платформи) на площині визначається трьома змінними: декартовими координатами x і y базової точки A , яка розташована на осі обертання коліс і рівновіддалена від них обох, а також кутом між вектором швидкості базової точки і позитивним напрямом осі абсцис (див. Рис.1).

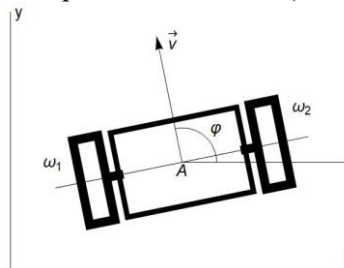


Рис.1. Кінематична схема двоколісного робота

Управління роботом здійснюється зміною кутових швидкостей обертання і, що у свою чергу призводить до зміни лінійних швидкостей руху коліс.

Після вказаних припущень та простих математичних викладок отримуємо наступну кінематичну модель робота

$$\begin{cases} \dot{x} = \frac{\omega_1(t) + \omega_2(t)}{2} r \cdot \cos \varphi(t) \\ \dot{y} = \frac{\omega_1(t) + \omega_2(t)}{2} r \cdot \sin \varphi(t) , \\ \dot{\varphi} = \frac{r}{L} (\omega_2(t) - \omega_1(t)) \end{cases} \quad (1)$$

де r – радіус коліс;

L – відстань між центрами коліс;

ω_1 , ω_2 – кутові швидкості;

φ – напрям вектору швидкості.

Побудова математичної моделі

Після вдосконалення математичних викладок, було виявлено, що знаючи залежність швидкостей $\omega_1(t)$ і $\omega_2(t)$ від часу визначається траєкторія руху.

По-перше зауважимо, що кутові швидкості $\omega_1(t)$ і $\omega_2(t)$ можуть бути як позитивними так і негативними. Знаки вибиратимемо наступним чином: якщо обертання колеса таке, що воно призводить до руху робота вперед, то вважаємо $\omega_i > 0$, в іншому випадку $\omega_i < 0$. Також якщо $\omega_2 > \omega_1$, то $\Omega > 0$, що відповідає повороту всієї платформи проти годинникової стрілки навколо центру повороту, точки C .

Прямолінійний рух робота

Як перший приклад розглянемо обертання коліс у позитивному напрямку з однаковою швидкістю $\omega_2 = \omega_1 = \omega = \text{const}$, яка не залежить від часу. В цьому випадку робот рухається прямолінійно з постійною швидкістю як за модулем, так і за напрямком, в ході математичних викладок отримуємо формулу

$$\begin{cases} x = \int \dot{x} dt = \int \omega r \cdot \cos \varphi_0 dt = \omega r \cos \varphi_0 \cdot t + x_0 \\ y = \int \dot{y} dt = \int \omega r \cdot \sin \varphi_0 dt = \omega r \sin \varphi_0 \cdot t + y_0 \end{cases}, \quad (2)$$

де x_0, y_0 – початкові (постійні) координати точки A ;

φ_0 – початковий напрямок вектору швидкості.

Обертання коліс у позитивному напрямку з постійним кутовим прискоренням ε . Причому нехай кутові швидкості обертання коліс є $\omega_2(t) = \omega_1(t) = \omega = \varepsilon t + \omega_0$.

Зауважимо, що значення кутового прискорення ε визначається експериментально виходячи з конструкційних особливостей робота, а початкова кутова швидкість для робота в стані спокою $\omega_0 = 0$.

$$\begin{cases} x(t) = \frac{a_x t^2}{2} + v_{0x} t + x_0 \\ y(t) = \frac{a_y t^2}{2} + v_{0y} t + y_0 \end{cases}, \quad (3)$$

де $a_x = r\varepsilon \cos \varphi_0 = a \cos \varphi_0$, $a_y = r\varepsilon \sin \varphi_0 = a \sin \varphi_0$ – компоненти прискорення вздовж відповідних координат;

$v_{0x} = r\omega_0 \cos \varphi_0 = v_0 \cos \varphi_0$, $v_{0y} = r\omega_0 \sin \varphi_0 = v_0 \sin \varphi_0$ – швидкості вздовж відповідних координат;

Круговий рух робота

Тепер розглянемо рухи робота у разі обертання коліс у різних напрямках з однаковою постійною кутовою швидкістю ω . Нехай для визначеності $\omega_2 = \omega$, $\omega_1 = -\omega$. У цьому випадку робот обертається навколо нерухомої точки A з позитивною (тобто проти годинникової стрілки) постійною кутовою швидкістю. Зауважимо, що у разі $\omega_2 = -\omega$, $\omega_1 = \omega$ обертання здійснювалося б за годинниковою стрілкою, а $\Omega < 0$.

$$\begin{cases} x = \int \dot{x} dt = \int \frac{dx}{dt} dt = \int dx = x_0 \\ y = \int \dot{y} dt = \int \frac{dy}{dt} dt = \int dy = y_0 \\ \varphi = \int \dot{\varphi} dt = \int 2 \frac{r}{L} \omega dt = 2 \frac{r}{L} \omega \cdot t + \varphi_0 \end{cases}. \quad (4)$$

Щоб однозначно поставити робота, нам потрібно або задати координати точки A , або координати центру обертання. Нехай $C_x = 0, C_y = 0$, тоді і початкове положення робота та його траєкторія відповідають такій моделі

$$R = \frac{\omega_+ r}{\Omega} = \frac{\omega_1 + \omega_2}{2} r \frac{L}{r(\omega_2 - \omega_1)} = \frac{L}{2} \frac{\omega_1 + \omega_2}{\omega_2 - \omega_1}, \quad (5)$$

$$\begin{cases} x(t) = R \sin(\Omega \cdot t + \varphi_0) + x_0 - R \sin \varphi_0 \\ y(t) = -R \cos(\Omega \cdot t + \varphi_0) + y_0 + R \cos \varphi_0 \\ \varphi = \int \dot{\varphi} dt = \int \frac{r}{L} (\omega_2 - \omega_1) dt = \frac{r}{L} (\omega_2 - \omega_1) \cdot t + \varphi_0 = \Omega \cdot t + \varphi_0 \end{cases}, \quad (6)$$

де R – радіус повороту;

де $\Omega = \frac{r}{L} (\omega_2 - \omega_1)$ – постійна кутова швидкість обертання робота як цілого.

Рух роботу по спіралі

Цей приклад навіяно такими фізичними міркуваннями. Нехай робот їде вперед прямолінійно із постійною швидкістю. Це відповідає обертанню коліс у позитивному напрямку з рівними та постійними кутовими швидкостями. Попереду намічається плавний поворот радіусом R . Завдання ставиться так, щоб пройти цей поворот з тією ж лінійною швидкістю, що і при русі вздовж прямолінійної ділянки траєкторії. Для побудови даної моделі випростовувався інтеграл Френеля $C(z), S(z)$

$$\begin{aligned} C(z) &= \int_0^z \cos\left(\frac{\pi t^2}{2}\right) dt \\ S(z) &= \int_0^z \sin\left(\frac{\pi t^2}{2}\right) dt \end{aligned}. \quad (7)$$

За допомогою даного інтегралу модель руху має вигляд

$$\begin{cases} x = x_0 + \omega_0 r \sqrt{\frac{\pi}{\xi}} \left[C\left(t \sqrt{\frac{\xi}{\pi}}\right) \cos(\varphi_0) - S\left(t \sqrt{\frac{\xi}{\pi}}\right) \sin(\varphi_0) \right] \\ y = y_0 + \omega_0 r \sqrt{\frac{\pi}{\xi}} \left[S\left(t \sqrt{\frac{\xi}{\pi}}\right) \cos(\varphi_0) + C\left(t \sqrt{\frac{\xi}{\pi}}\right) \sin(\varphi_0) \right] \\ \varphi = \frac{\xi t^2}{2} + \varphi_0 \end{cases}, \quad (8)$$

де $\xi = 2r\varepsilon/L$.

За допомогою спеціальної функції Френеля, також була розроблена модель, із кутових швидкостей для параметричного завдання траєкторії у разі $\omega_1(t) = \omega_{10} + \varepsilon t$, $\omega_2(t) = \omega_{20} - \varepsilon t$, якщо $\Delta\omega = \omega_{20} - \omega_{10} > 0$

$$\begin{cases} x = x_0 + \omega_k r \sqrt{\frac{\pi}{\xi}} \left[\cos(\Phi) \left[C\left(\frac{\Omega_0}{\sqrt{\pi\xi}}\right) + C\left(t \sqrt{\frac{\xi}{\pi}} - \frac{\Omega_0}{\sqrt{\pi\xi}}\right) \right] + \sin(\Phi) \left[S\left(\frac{\Omega_0}{\sqrt{\pi\xi}}\right) + S\left(t \sqrt{\frac{\xi}{\pi}} - \frac{\Omega_0}{\sqrt{\pi\xi}}\right) \right] \right] \\ y = y_0 + \omega_k r \sqrt{\frac{\pi}{\xi}} \left[\sin(\Phi) \left[C\left(\frac{\Omega_0}{\sqrt{\pi\xi}}\right) + C\left(t \sqrt{\frac{\xi}{\pi}} - \frac{\Omega_0}{\sqrt{\pi\xi}}\right) \right] - \cos(\Phi) \left[S\left(\frac{\Omega_0}{\sqrt{\pi\xi}}\right) + S\left(t \sqrt{\frac{\xi}{\pi}} - \frac{\Omega_0}{\sqrt{\pi\xi}}\right) \right] \right] \\ \varphi = \varphi_0 + \Omega_0 t - \frac{\xi t^2}{2} \end{cases}, \quad (9)$$

де $\omega_k = (\omega_{10} + \omega_{20})/2$ - кінцева кутова швидкість обертання коліс;

$$\Omega_0 = r(\omega_{20} - \omega_{10}) / L, \quad \xi = 2r\varepsilon / L, \quad \Phi = \frac{\Omega_0^2}{2\xi} + \varphi_0 - \text{початкова кутова швидкість обертання,}$$

кутове прискорення та початкова фаза руху робота як цілого.

Програмна модель та побудова траєкторії руху

Постановка задачі заключає в собі побудову траєкторію руху робота, якщо залежність кутових швидкостей лівого та правого колеса представлені рисунку 4. Спочатку робот знаходиться в стані спокою, надалі він прискорюється, а після цього рухається з постійною лінійною швидкістю по деякій траєкторії яку необхідно знайти.

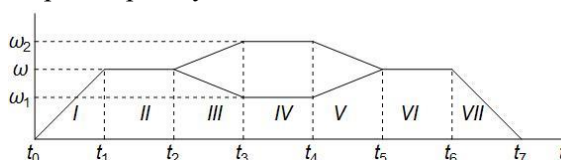


Рис.4. Залежність кутових швидкостей обертання коліс від часу

З малюнка 4 випливає, що рух робота складається із семи етапів (I, II, III, IV, V, VI, VII). Перший етап триває t_1 секунд і є рівноприскореним розгоном робота зі стану спокою. При цьому і ліве (перше) і праве (друге) колеса за час t_1 розганяються. На другому етапі два колеса обертаються із цією постійною швидкістю протягом $t_2 - t_1$ секунд. На третьому етапі перше колесо рівномірно уповільнює свою швидкість, при цьому друге колесо, за цей самий час $t_3 - t_2$ рівномірно збільшує свою швидкість. Четвертий етап триває $t_4 - t_3$ секунд і весь час колеса обертаються з постійними, але різними кутовими швидкостями, які вони придбали на третьому етапі. На п'ятому етапі за час $t_5 - t_4$ ліве колесо збільшує свою швидкість обертання, а праве зменшує. На шостому етапі обидва колеса обертаються однаковою швидкістю протягом $t_6 - t_5$ секунд. І на сьомому етапі обидва колеса поступово зменшують свою швидкість обертання до нуля за $t_7 - t_6$ секунд. Завдання зобразити яку траєкторію руху опише при цьому робот, вірніше його базова точка A .

Програмна модель робота виконана у середовищі Wolfram Mathematica, за допомогою математичної моделі. В ході виконання роботи, були набуті навички побудови математичної моделі та отримані знання з теорії спеціальних функцій, були використані навички з інтегрування. В програмній моделі був придбаний досвід з програмуванням в середовищі Wolfram Mathematica.

Головним результатом наукової роботи є програмна модель роботу з диференціальним приводом.

ЛІТЕРАТУРА

1. Stephen Wolfram (2015) An Elementary Introduction to the Wolfram Language, First Edition.
2. T Pearcey (1956) Table of the Fresnel integral to six decimal places.
3. Edward A. Bender (2000) An Introduction to Mathematical Modeling.
4. Anirudh Topiwala (2016) Modeling and Simulation of a Differential Drive Mobile Robot.
5. Henry W. Stone (1987) Kinematic Modeling, Identification, and Control of Robotic Manipulators.

БЕРЕЖНИЙ Артем Андрійович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: xa12284047@student.karazin.ua; ORCID 0000-0001-5407-9015.

Наукові інтереси:

- розробка та аналіз робототехнічних систем;
- створення пристроїв IOT та їх програмний супровід.

БИКОВА Т.В.

АНАЛІЗ ВИМОГ ДО МЕТРОЛОГІЧНИХ ХАРАКТЕРИСТИК «ІНТЕРНЕТУ РЕЧЕЙ»

Актуальність: швидке нарощення темпів впровадження технології «Інтернет речей» (IoT) призводить до того, що в найближчому часі більшість функцій моніторингу стану навколишнього середовища буде передано автоматизованим глобальним системам. Людство буде залежати від точності і достовірності інформації, яку будуть надавати ці системи. Оскільки переважна більшість цієї інформації – це результати вимірювань фізичних параметрів, то актуальною є задача оцінки метрологічних характеристик систем на базі IoT.

Мета роботи: виявити найбільш актуальні задачі метрологічного забезпечення систем на базі технології IoT і запропонувати шляхи їх вирішення.

Методи дослідження: теоретичні методи аналізу, синтезу, абстрагування та узагальнення, пояснення, класифікація, системний та функціональний аналіз.

Результати: проведено аналіз процесів функціонування вимірювальних систем, побудованих на базі IoT, та процесів їх метрологічного забезпечення. Виявлено, що найбільш актуальними і не вирішеними є проблеми дистанційного калібрування датчиків, створення вбудованих калібраторів і подовження терміну інтервалу калібрування.

Висновки: Першочерговою проблемою технології IoT є створення ретельної системи правил і стандартизації для можливості глобалізації і розширення сфер впровадження цієї технології. Необхідно вирішити питання уніфікованості датчиків, що будуть входити до складу системи. Друга проблема створення методології та апаратного забезпечення дистанційного калібрування датчиків і виконавчих механізмів з метою підвищення надійності і якості якості отриманих даних. Частково проблему калібрування можна вирішити збільшенням величини міжкалібровочного інтервала, але це призведе до удорожчання датчиків. Більш перспективним є алгоритмічний підхід для покращення якості даних, але зараз це представляє дослідницьку проблему. Хоча можна очікувати гарних результатів, якщо поєднати алгоритмічний підхід з апаратним, наприклад дистанційним калібруванням..

Ключові слова: інтернет речей, метрологічна характеристика, калібрування, метрологічне забезпечення.

Постановка проблеми та її актуальність

Нарощування розмірів вимірювальних систем на базі технології «Інтернет речей» йде дуже швидкими темпами. Це дає можливість здійснювати моніторинг стану навколишнього середовища в кожному куточку людської діяльності, але накладає ряд вимог щодо якості даних, які надають ці системи [1].

Прикладом системи, побудованої за технологією IoT є «Розумне місто». Це місто, наповнене датчиками та виконавчими механізмами, які вимірюють і контролюють споживання різних джерел енергії та параметри навколишнього середовища до саморегулювання майже без участі людини. Навіть сьогодні багато міст у всьому світі вже використовують вимірювання повітря, вимірювання трафіку, вимірювання денного світла, вимірювання споживання газу, електроенергії та води тощо. Якщо ми почнемо включати більше датчиків у наше повсякденне життя, ми зможемо замінити людське судження спеціальною вимірювальною системою та датчиками, створеними для конкретних областей.

Оскільки технологія IoT все ще є новою концепцією, її стандартизація була головним центром для багатьох організацій та урядів: наприклад, регуляторний орган США у 2009 році, Американський національний інститут стандартів або інтенсивна стандартизація та регулятивні зусилля в Європі у 2016 році (наприклад, Міжнародний Телекомунікаційний союз, Міжнародна організація стандартизації, Європейський комітет з електротехнічної стандартизації). Китайський інститут стандартизації електроніки також підкреслив важливість IoT у своєму 12-му п'ятирічному плані [2], [3]. З 2014 року було зроблено багато роботи зі стандартизації термінології, норм і правил, які тісно пов'язані з технологіями IoT, але без суттєвих покращень [4].

Однією з причин повільного розвитку IoT є те, що в його мережі існує багато різноманітних пристроїв. З плином часу кількість пристроїв, що використовуються в мережі IoT, зростає експоненціально. Оскільки все більше пристроїв починають обмінюватися даними, необхідно уникати різних стандартів країн світу. Найкращим прикладом можуть бути імперські одиниці та одиниці SI. Хоча перетворення між ними не є проблемою, воно все одно додає системі затримку.

Тому стандартизація всіх пристроїв, які підключаються в глобальну мережу є першочерговою проблемою.

З проходженням певного часу експлуатації датчика виникає потреба контролю його метрологічних характеристик. З часом ці датчики доведеться замінити або відкалібрувати заново залежно від виду датчиків. Ймовірно, більшість наших повсякденних вимірювань, таких як температура тіла або кімнатна температура й вологість, не мають бути дуже точними, тому їх, швидше за все, не доведеться повторно калібрувати. З іншого боку, датчики, такі як детектори газу та диму, як приклади спостереження безпеки, повинні бути точними. Оскільки немає простого та надійного способу перевірити, чи ці датчики працюють належним чином, без їх навмисної активації, необхідно запровадити більш зручні процедури контролю, тестування та/або самокалібрування.

Датчики, які здатні самокалібруватися, також будуть схильні до старіння, але їхні статичні та динамічні параметри мають бути набагато точнішими протягом тривалого часу. Очікувана тривалість життя цих пристроїв залежить від точності їх компонентів, що самокалібруються. Таким чином ми можемо продовжити термін служби домашнього чи промислового датчика, не збільшуючи вартість їх обслуговування, але водночас збільшуючи вартість виробництва та точність, коли це необхідно. Деякі дорогі та надійні датчики, що використовуються в екстремальних умовах, повинні бути відкалібровані навченим персоналом у певних умовах навколишнього середовища. У такому випадку необхідно взяти ці датчики та відправити їх на повторне калібрування лише в ті лабораторії, які можуть забезпечити такі умови середовища. Це було б надто дорого та створило б логістичні проблеми, якщо робити це у великих масштабах. Ось чому в останні кілька років було докладено багато зусиль, щоб уможливити повторне калібрування таких датчиків на місці в їхньому природному середовищі за допомогою дистанційного калібрування.

Датчики схильні до зносу. Дані, отримані з нових датчиків, мають обмежений термін гарантії від виробника, що показання будуть точними, а їх поведінка передбачуваною в наступний період. Після закінчення гарантійного терміну виробник датчика жодним чином не несе відповідальності за точність показань, які надає датчик.

Датчики, як і будь-який інший тип обладнання, схильні до зношування з часом, що робить їх показання менш точними та надійними. Згодом датчики потрібно відкалібрувати повторно. Визначення калібрування Міжнародним бюро мір і ваг (BIPM) таке: «Операція, яка за певних умов, на першому кроці, встановлює зв'язок між значеннями величини та похибками вимірювань, що надаються еталонами вимірювань, і відповідними показаннями з відповідним вимірюванням. невизначеності (каліброваного приладу або вторинного стандарту) і, на другому кроці, використовує цю інформацію для встановлення співвідношення для отримання результату вимірювання за індикацією [5].

В даний час у промислових застосуваннях, якщо датчик потрібно відкалібрувати вручну, зазвичай його потрібно вилучити з передбачуваного середовища, надіслати в сертифіковану акредитовану калібрувальну лабораторію, повторно відкалібрувати та надіслати назад власнику для його впровадження. Цим процесом є багато очевидних проблем: система тривалий час має працювати без датчика, середовище калібрування може не відповідати робочому середовищу, калібрування в сертифікованій лабораторії є дороге, ризик пошкодити датчик під час транспортування, потрібна велика кількість кваліфікованого персоналу для обслуговування системи.

Ряд дослідників поділяють думку [6] про те, що слід відмовитися від традиційних методів періодичних калібрувань. Було розроблено багато спроб дистанційного калібрування, але жодна з них не була успішною на ринку. В ідеалі датчики повинні мати можливість автоматичного калібрування або, принаймні, мати певну резервну систему, таку як алгоритм для зміни даних у випадку, якщо показання виявляться неточними.

У [7] акцент зроблено на управлінні отриманими даними та на забезпеченні надійності дистанційного калібрування та шифруванні отриманих даних. Вони також пропонують систему для автоматичного калібрування, засновану на комбінованій архітектурі для дистанційного калібрування.

У [8] акцент зроблено на концепції усунення впливу невідомих середовищ, які використовуються для поширення одиниці вимірювання. Хоча концепція представлена лише в

моделюванні, подальше вдосконалення цієї концепції допомогло б знизити приписувану невизначеність і, отже, допомогти спільноті дистанційного калібрування в цілому. Також вказують на незручність транспортування вимірювальних приладів між лабораторіями та майданчиками, де проводиться калібрування.

У [9] автори представляють концепцію самокалібрування приладів, що використовуються у небезпечних середовищах, де уникають взаємодії людини. Їхній підхід повністю базується на апаратному забезпеченні та легко застосовний. Подальші оновлення можуть бути можливими шляхом впровадження допоміжного алгоритму для самодіагностики або алгоритму, який би вимірював справність і стан системи.

Таким чином можна підвищити надійність усієї системи, оскільки буде реалізовано перевірку на помилки. Ця концепція може виявитися хорошим способом дистанційного калібрування датчиків температури без необхідності видаляти їх із системи.

Новий підхід до повторного калібрування – це вирішення проблеми на основі алгоритму. По суті, він фільтрує та відхиляє результати вимірювань, які виявилися помилковими. Хоча він простий і ефективний за своєю природою, він все ж позбавляє безперервності отриманих даних, що робить його не ідеальним для деяких цілей, які покладаються на швидку реакцію у разі дефектів. З іншого боку, алгоритмічний підхід має одну величезну перевагу, яка потенційно може співіснувати з нашою пропозицією. Його можна використовувати для аналізу поточного стану датчика та прогнозування його стану в найближчому або далекому майбутньому залежно від типу та стану датчика. Таким чином датчик можна відкалібрувати повторно в той момент, коли це запропонує алгоритм.

Ще одна проблема з підходом алгоритму полягає в тому, що алгоритм повинен мати можливість відрізнити рідкісні екстремальні умови (наприклад, раптову зміну клімату навколишнього середовища) і несправний датчик. У деяких випадках це може виявитися проблематичним, оскільки алгоритм для екстремальних випадків може бути розроблений спеціально для певного середовища. Цей процес не можна виконати, якщо ми не впевнені, що датчик має правильні показання. Якщо ринок обирає алгоритмічний підхід, необхідно враховувати умови вимірювання, які виходять за межі визначених робочих умов. Крім того, алгоритми повинні бути сертифіковані та перевірені лабораторіями або мати певну форму стандартизованого цифрового сертифіката калібрування [6]. Хоча рішення алгоритму дає потрібне передбачення того, коли отримані дані будуть ненадійними, і дає пропозиції щодо того, як впоратися з такими відхиленнями, воно не вирішує основної проблеми: датчик все ще видає неправильні показання та фактично не працює.

Як зазначають, основною складністю вирішення завдань метрологічного самоконтролю, розглянутих у статті, є проблема наявності еталонного значення для порівняння значень відповідних отриманих величин. Одним із варіантів може бути використання значень, отриманих одразу після повторного калібрування, оскільки це значення має бути максимально точним, якщо процес повторного калібрування сертифіковано. В даний час періодичне повторне калібрування пристрою виконується за допомогою певного стандарту вимірюваної одиниці поруч із датчиком і періодичного використання його для перевірки, чи збільшилося відхилення. Проблема цього методу полягає в тому, що іноді чистий зразок є дуже дорогим або дуже великим. Крім того, ці стандарти також схильні до старіння і з часом стають не такими надійними, як раніше. Звичайно, найпростішим способом вирішення цієї проблеми є наявність персоналу для заповнення/зміни/контролю використаного стандарту, і сам по собі цей процес може бути дуже трудомістким, небезпечним і дорожчим, ніж процес повторного калібрування в калібрувальній лабораторії.

Знову ж таки, ключовою проблемою залишається: людський фактор не прибраний з процесу автоматизації. Деякі стандарти, такі як час, частота, напруга та потужність, можна усунути відомим способом. Сьогодні рівномірний час легко доступний, оскільки ми можемо повторно відкалібрувати всі цифрові таймери, які мають підключення до Інтернету або GPS. Ми можемо відкалібрувати напругу за допомогою стандарту напруги Джозефсона, а також використовувати його для «синтезу» потужності. Щоб спростити дистанційне повторне калібрування, необхідно буде розробити нові переносні стандарти.

Однією з проблем дистанційного калібрування є те, що багато виробників виготовляють свої датчики та/або інструменти, тому вони несумісні з інструментами чи датчиками інших

виробників. Це ускладнює процес дистанційного калібрування, оскільки, можливо, виключається модульність датчиків. Для віддаленого калібрування потрібне підключення до мережі/Інтернету. Якщо датчики не мають комунікаційного порту, їх доведеться або відкалібрувати за допомогою звичайного методу, відправивши його в лабораторію, або додати до нього спеціально розроблений комунікаційний пристрій, щоб зробити його «розумним» датчиком.

ВИСНОВОК. Першочерговою проблемою технології IoT є створення ретельної системи правил і стандартизації для можливості глобалізації і розширення сфер впровадження цієї технології. Необхідно вирішити питання уніфікованості датчиків, що будуть входити до складу системи.

Друга проблема створення методології та апаратного забезпечення дистанційного калібрування датчиків і виконавчих механізмів з метою підвищення надійності і якості якості отриманих даних. Частково проблему калібрування можна вирішити збільшенням величини міжкалібровочного інтервала, але це призведе до удорожчання датчиків. Більш перспективним є алгоритмічний підхід для покращення якості даних, але зараз це представляє дослідницьку проблему. Хоча можна очікувати гарних результатів, якщо поєднати алгоритмічний підхід з апаратним, наприклад дистанційним калібруванням.

ЛІТЕРАТУРА

1. Sandrić B, Jurčević M. Metrology and Quality Assurance in Internet of Things. Conference: 2018 First International Colloquium on Smart Grid Metrology (SmaGriMet) DOI: 10.23919/SMAGRIMET.2018.8369849 Available from: https://www.researchgate.net/publication/325493356_Metrology_and_quality_assurance_in_internet_of_things [accessed Oct 24 2022]
2. Z. Sheng, C. Mahapatra, C. Zhu, and V. C. M. Leung, "Recent Advances in Industrial Wireless Sensor Networks Toward Efficient Management in IoT," IEEE Access, vol. 3, no. Omam, pp. 622–637, 2015.
3. L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," IEEE Trans. Ind. Informatics, vol. 10, no. 4, pp. 2233–2243, 2014.
4. J. A. Stankovic, "Research directions for the internet of things," IEEE Internet Things J., vol. 1, no. 1, pp. 3–9, 2014.
5. R. Taymanov, K. Sapozhnikova, and A. Ionov, "Topical Metrology Problems in the Era of Cyber-physical Systems and Internet of Things," 18th Int. Congr. Metrol., vol. 09006, p. 09006, 2017.
6. Joint Committee For Guides In Metrology, "Vocabulaire international de métrologie," VIM3 Int. Vocab. Metrol., vol. 3, no. third, p. 104, 2012.
7. L. Chen, Y. Shao, Y. Fu, and M. Xie, "A Remote Calibration System for Industrial Thermometers," pp. 1574–1577, 2013.
8. Z. Nakutis, M. Saunoris, R. Ramanaukas, and V. Daunoras, "A method for remote wattmeter calibration," I2MTC 2017 -2017 IEEE Int. Instrum. Meas. Technol. Conf. Proc., 2017.
9. A. S. Shakyia, "Temperature Monitoring System with Remote Calibration Capability," pp. 8–13, 2017.

БИКОВА Тетяна Володимирівна – к.т.н., доцент кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: tetiana.bykova@karazin.ua; ORCID: 0000-0003-0484-5388.

Наукові інтереси:

– *інтернет речей*;

– *прогнозування стану систем на основі даних моніторингу*.

БУРЛАЙ І.С., БАКУМЕНКО Н.С.

КОМП'ЮТЕРНА МОДЕЛЬ СИСТЕМИ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ЦІН НА ЕНЕРГОНОСІЇ ЗА ДОПОМОГОЮ ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

Topic: computer model of energy price analysis and forecasting system using convolutional neural networks.**Purpose:** development of a computer model of the energy price forecasting system using time series analysis methods.**Research methods:** methods of time series analysis and forecasting, computer modeling were used to solve the task**Result:** a computer model was obtained with a forecast of energy prices, the model error was calculated.**Conclusions:** The methods of analysis and forecasting of time series were considered in this work. One of the methods most suitable for solving the problem was chosen - convolutional neural networks. At the end, the forecast was obtained in graphical form and the error of this forecast**Ключові слова:** методи прогнозування, часові ряди, штучні нейронні мережі, згорткові нейронні мережі.

Сьогодні однією з провідних галузей науки є Data science. Ця дисципліна вивчає методи роботи з великими об'ємами даних, у тому числі і задачу аналізу та прогнозування часових рядів. Часові ряди, які відслідковують динаміку змін різних економічних показників, у тому числі і цін на енергоносії, є важливим інструментом для формування економічної та інвестиційної політики. Саме через постійні коливання показників перманентно існує потреба в їх аналізі та прогнозуванні.

У цій роботі було розглянуто найбільш поширені методи вирішення даної задачі і проведена комп'ютерна реалізація одного з них – згорткових нейронних мереж.

Регресійний аналіз

Регресійний аналіз – один з найперших методів аналізу та прогнозування часових рядів. Модель можна описати як функцію залежної і незалежних змінних. В рамках цього методу необхідно знайти функціональну залежність:

$$y = f(x_2, x_3, \dots, x_m), \quad (1)$$

де y – залежна змінна;

$x_1, x_2, \dots, x_m$ – незалежні змінні.

Регресійний метод має безліч різновидів, таких як лінійна, гіперболічна, параболічна, логарифмічна тощо.

Даний метод, разом з усіма своїми різновидами має свої певні переваги і недоліки, а саме такі:

- Простота – усі операції, що виконуються дуже прості і не потребують додаткових роз'яснень;
- Швидкість – оскільки алгоритм має небагато кроків виконання і прості операції, що і дозволяє отримувати результат так швидко.
- Прозорість – через свою простоту, можна легко узяти будь-який проміжний результат і перевірити його коректність;
- Вибір функціональної залежності – завдяки великій кількості різновидів, важко підібрати потрібний вид;
- Погана адаптивність – цей метод дозволяє працювати з обмеженою кількістю часових рядів.

Метод Хольта-Вінтерса

Метод Хольта-Вінтерса є покращеною версією методу Хольта. На відміну від вищезгаданого регресійного аналізу, даний метод складніший, він базується вже на трьох компонентах і представити це можна у вигляді:

$$\hat{y}_{t+1} = a_1 + b_1 + c_{t+1-m}, \quad (2)$$

де a_t – базова компонента;
 b_t – лінія тренду;
 c_t – сезонний ефект;
 m – період сезону.

Метод Хольта-Вінтерса має свої характерні риси, а саме:

- Будь-який об'єм вхідних даних – можна працювати з будь-яким набором даних, незалежно від його розміру;
- Простота алгоритму – дозволяє легко проаналізувати розрахунки і виявити помилку, якщо вона є;
- Негнучкий алгоритм – не дозволяє працювати з деякими видами часових рядів;
- Тонке налаштування – важко налагодити високу точність згладжувальних функцій.

ARIMA

Метод ARIMA – покращена версія методу ARMA, що виправляє його головний недолік – обмеження на роботу зі стаціонарними часовими рядами. На відміну від попередніх двох згаданих методів, модель ARIMA значно складніша, її можна описати як:

$$x_t^k = \alpha_1 x_{t-1}^k + \alpha_2 x_{t-2}^k + \dots + \alpha_p x_{t-p}^k + \delta - \theta_1 \delta_{t-1} - \dots - \theta_q \delta_{t-q}, \quad (3)$$

де α – деякий числовий коефіцієнт, що не більше в абсолютній величині за 1;

$\theta_1, \dots, \theta_q$ – кінцевий набір параметрів β ;

$\delta_1, \delta_2, \dots$ – це послідовність випадкових величин або «білий шум».

При цьому:

$$x_t^k = \Delta^k x_t = x_t - C_k^1 x_{t-1} + C_k^2 x_{t-2} - \dots + (-1)^k x_{t-2k}, \quad (4)$$

де C_{t-k}^k – число поєднань;

Δ – різницевий оператор.

Головні особливості цього методу:

- Гнучкий алгоритм – дозволяє працювати з великою кількістю різновидів часових рядів;
- Високоточний результат – серед подібних методів один із найкращих;
- Мінімальне обмеження об'єму вхідних даних – без 40 елементів часового ряду модель не може коректно працювати.;
- Вибір параметрів моделі – дуже важко точно підібрати найкращі параметри моделі, алгоритму для цього не існує;
- Перерахунок моделі – треба здійснювати при появі нових спостережень.

Радіально-базисні нейронні мережі

Радіально-базисні нейронні мережі – це нейронні мережі, в котрих навчання протікає без зворотного розповсюдження помилок.

Радіально-базисні нейронні мережі мають такі риси:

- Масштабованість – при збільшенні структури нейронної мережі якість результату не змінюється;
- Ітеративність – з першої ітерації вже буде приблизний результат і чим більше ітерацій буде виконано, тим більш точним буде прогноз;
- Великий алгоритм – необхідно виконати велику кількість операцій;
- Повільний – відносно інших нейромережевих методів.

Згорткові нейронні мережі

Згорткові нейронні мережі – це штучні нейронні мережі, що характеризуються головним чином, здатністю працювати з даними структурованими у вигляді сітки. Даний метод дозволяє працювати з одно-, дво- і тривимірними часовими рядами. Базисом згорткових нейронних мереж є математична операція згортки. Будь-який елемент у вихідній карті ознак.

Згорткові нейронні мережі планувалися в першу чергу для роботи із зображеннями, але з'ясувалося, що вони ефективні при аналізі та прогнозуванні найпростіших одновимірних часових рядів. Перелік застосування цих мереж постійно розширюється, що дозволяє вирішувати нові і нові задачі.

Оскільки тема доповіді передбачає методи аналізу та прогнозування часових рядів, то і модель даного методу приведемо відповідну:

$$f(i) = (xw)(i) = \sum_{j=0}^{k-1} x(i-j)w(j), \quad (5)$$

де f – карта функцій;

x – вхід;

i, j – номери елементу згортки;

w – ядро;

k – розмір ядра.

Слід також розглянути основні переваги і недоліки даного методу:

- Гарна зв'язність - тісні зв'язки вхідних нейронів з прихованими;
- Спільного доступ – певне ядро можна використовувати для будь-якого розташування;
- Швидкість виконання – обумовлена двома вищезгаданими властивостями;
- Простота – даний метод, відносно інших нейромережових методів, легко реалізувати;
- Ефективний підбір параметрів – вибір параметрів формалізований і може легко бути програмно реалізований;
- Гнучкість навчання – у рамках цього сімейства можна користуватися різними методами навчання;
- Повільний підбір параметрів – виконання алгоритму вибору параметрів потребує багато часу;
- Велика кількість навчальних матеріалів – потрібен набір даних з чималим об'ємом даних.

Дані

При подальшій реалізації згорткових нейронних мереж було обрано набір даних «BrentOilPrices.csv». Даний датасет зберігає динаміку зміни значення ціни на нафту марки «Brent» від 20 травня 1983-го року і закінчуючи 21 квітня 2020-го року.



Рис. 1 – Часовий ряд «BrentOilPrices.csv»

Програмна реалізація моделі

Комп'ютерну модель було створено за допомогою мови програмування Python на базі хмарного інтерактивного середовища – Google Colaboratoty. Програма аналізує часовий ряд, робить висновки про його властивості, розподіляє набір даних на навчальну і тестувальні частини і проводить прогноз. У підсумку ми отримали графічне відображення прогнозу цін на нафту марки «Brent»:

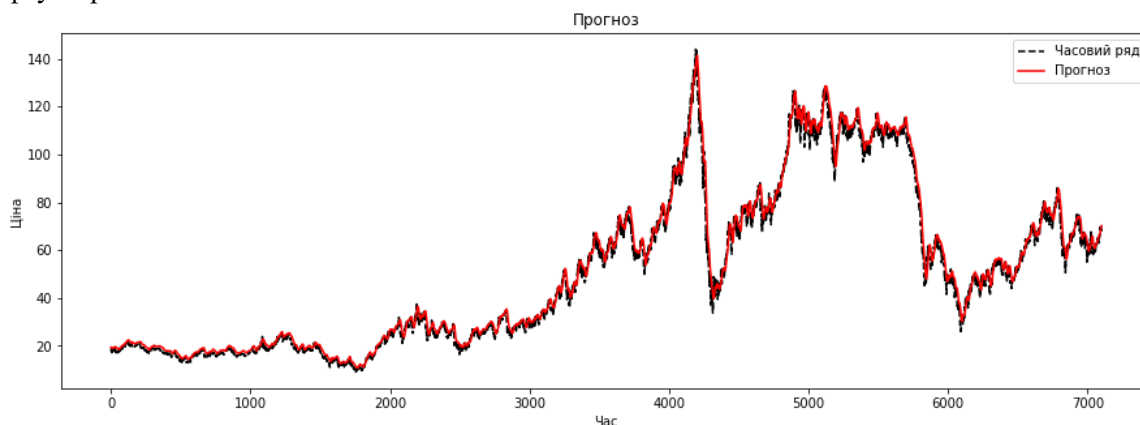


Рис. 3.9 – Прогноз часового ряду

Похибка цього прогнозу склала 7,94 %.

Підсумки

Таким чином, під час виконання даної роботи було здійснено огляд основних методів аналізу та прогнозування часових рядів. Серед розглянутих методів було обрано згорткові нейронні мережі і була створена комп'ютерна модель системи прогнозування цін на енергоносії. Розроблена програма досліджує набір даних і розраховує прогноз. Створене програмне забезпечення може бути використане для широкого кола задач аналізу та прогнозування часових рядів.

Список літератури

1. Макогон Ю.В., Пилипенко В.В. Основи наукових досліджень в економіці: Навч. посібник. – Донецьк: Альфа-прес, 2007. – С. 7-23
2. Пілюшенко В.Л., Шкрабак І.В., Славенко Е.І. Наукове дослідження: організація, методологія, інформаційне забезпечення: Навчальний посібник. – Київ: Лібра, 2004. – С. 52-60.
3. Інтелектуальні системи автоматизації : монографія / Аврунін О. Г., Владов С., І., Петченко М. В., Семенець В. В., Татарінов В. В., Тельнова Г. В., Філатов В. О., Шмельов Ю. М., Шушляпіна Н. О. – Кременчук : Видавництво «НОВАБУК», 2021. – 322 с.
4. Стеченко Д.М., Чмир О.С. Методологія наукових досліджень: Підручник: - К.: Знання, 2005. – С. 91-190.
5. Ткаченко Р. О. Нейромережеві засоби штучного інтелекту : навч. посіб. / Роман Олексійович Ткаченко, Павло Романович Ткаченко, Іван Вікторович Ізонін ; Нац. ун-т "Львівська політехніка". – Львів : Вид-во Львівської політехніки, 2017. – 206 с.

БУРЛАЙ Іскандер Сергійович – магістр, студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: xa11867825@student.karazin.ua; ORCID: 0000-0002-5802-8183.

Наукові інтереси: – *аналіз та прогнозування часових рядів.*

БАКУМЕНКО Ніна Станіславівна – к. т. н., доцент; доцент кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: n.bakumenko@karazin.ua; ORCID: 0000-0003-3496-7167.

Наукові інтереси:

- *інтелектуальний аналіз даних*

ВЕРЕМІЄНКО Б.О., ЧУБ О.І.

ВИКОРИСТАННЯ АФІННИХ ПЕРЕТВОРЕНЬ У ГЕОМЕТРИЧНОМУ МОДЕЛЮВАННІ

In this article the features of studying one of the convenient plane transformations – affine transformation, which is practically not studied in the course of geometry. The authors focus on the properties of the affine transformation. The importance and difficulty of studying the theory of affine transformation is primarily due to the versatility and variability as such. The article presents a classification of affine transformations on various grounds. Particular attention is paid to the general idea of applying the theory to the solution of the problem. In geometric modeling, the affine transformation allows you to move from more complex to simpler, so it is convenient for solving problems. Comprehension and application of this method forms a constructive approach to problem solving and critical thinking in researchers.

Ключові слова: афінні перетворення, геометричне моделювання, поверхнева модель, комп'ютерна графіка, візуалізація об'єктів, теорема Фалеса.

Вступ

Геометричне моделювання використовується для опису об'єктів, що мають геометричні властивості. Розвиток геометричного моделювання пов'язаний з розвитком машинної графіки, основним завданням якої є візуалізація складних об'єктів. Для цього розробляються методи та алгоритми формування довільних ліній та поверхонь, видалення невидимих елементів тощо. Геометрична модель об'єкта є машинним уявленням його форми та розмірів. Розрізняють двовимірне та тривимірне геометричне моделювання. Двовимірні моделі широко застосовуються і зараз для вирішення різноманітних завдань, але вони є не зручними для складних об'єктів. В цьому дослідженні будемо розглядати тривимірні моделі.

Метою дослідження є вивчення процесу афінних перетворень, а також перетворення паралелепіпеду в куб зі збереженням необхідних властивостей геометричних тіл.

Основна частина

Афінна геометрія, що вивчає властивості фігур, інваріантних за будь-яких афінних перетвореннях, почала розвиватися в наукових працях німецьких математиків Мебіуса А. та Плюккера Ю. У 80-х роках ХХ століття Клейн Ф. ввів в класифікацію геометричних перетворень та відповідних гілок геометрії поняття групи.

Тривимірна модель служить для того, щоб створити віртуальне уявлення об'єкта у всіх трьох вимірах. Виділемо каркасні та поверхневі тривимірні моделі. Каркасні моделі використовують лише координати (x, y, z) вершин і ребра, що з'єднують їх. Поверхневі моделі дозволяють описувати складні поверхні. Тому поверхневі моделі часто комбінують з каркасними для опису поверхонь об'єкта, які неможливо автоматично визначити за каркасною моделлю. Реалізація такої комбінації моделей та відповідних методів є дуже вживаною у промисловості.

Афінним перетворенням площини називається таке перетворення площини, яке зберігає колінеарність точок та просте відношення трьох точок. Існує єдине афінне перетворення площини, що перетворює три довільно задані неколінеарні точки відповідно в три довільно задані неколінеарні точки. Тому будь-яке афінне перетворення площини може бути задано вказаними трьома точками, що не лежать на одній прямій, та їх образів.

Властивості афінних перетворень:

- 1) афінне перетворення відображає три точки, що не лежать на одній прямій, в три точки, що не лежать на одній прямій;
- 2) афінне перетворення перетворює пряму в пряму, відрізок у відрізок, промінь у промінь, напівплощину в напівплощину;
- 3) афінне перетворення відображає паралельні прямі в паралельні прямі;
- 4) афінне перетворення зберігає відношення паралельних відрізків;
- 5) афінне перетворення перетворює прямі, що перетинаються, в прямі, що перетинаються, і точку перетину прямих, що перетинаються, в точку перетину їх образів;

- 6) афінні координати точки інваріантні щодо афінних перетворень площини;
- 7) афінне перетворення може зберігати орієнтацію площини;
- 8) афінне перетворення може змінювати орієнтацію площини.

Розглянемо афінне перетворення на прикладі розв'язання математичної задачі. В умові задачі зазначається, що на діагоналях AC і BA' бічних граней паралелепіпеда $ABCD A'B'C'D'$ обрані точки M і N так, що відрізок MN паралельний діагоналі паралелепіпеда DB' . Потрібно знайти співвідношення MN до DB' (рис. 1).

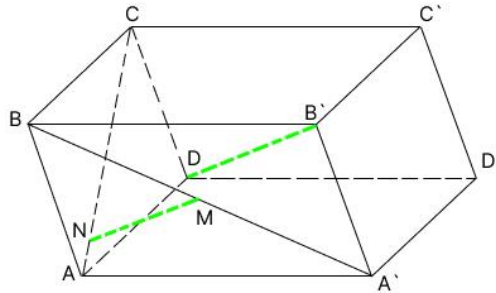


Рис. 1 Паралелепіпед

Для розв'язання цієї задачі нам необхідно знати властивості афінних перетворень, зазначені вище. Застосовуючи деякі властивості афінного перетворення, перетворимо паралелепіпед у куб (рис. 2).

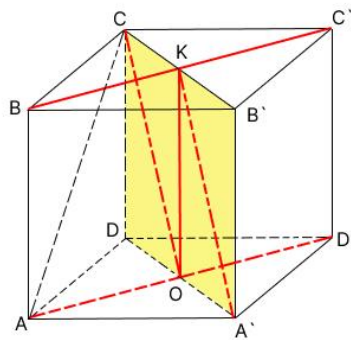


Рис. 2 Куб

Подивимося на цей куб вздовж діагоналі BC' , спроектувавши потрібні нам точки в площину $DCB'A'$ (рис. 3).

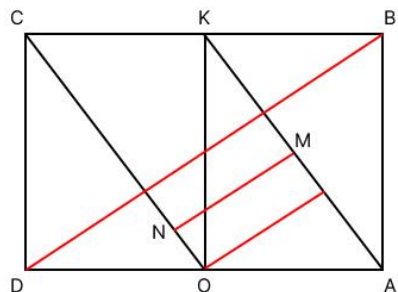


Рис. 3 Площина $DCB'A'$

З теореми Фалеса випливає, якщо паралельні прямі, що перетинають дві задані прямі a і b , відтинають на одній прямій рівні відрізки, то вони відтинають рівні відрізки й на іншій прямій.. Тобто, образ відрізка MN дорівнюватиме одній третині образу діагоналі DB' . Тож відповідь $MN:DB' = 1:3$.

Ідея застосування перетворень площини до вирішення завдань полягає в наступному. Фігуру, задану в задачі, за допомогою деякого перетворення перетворюють на фігуру, для якої ця задача має простіше рішення. Потім отримані результати переносять на цю фігуру, застосувавши перетворення, обернене до розглянутого перетворення. Афінні перетворення мають у порівнянні з рухом і подобою найбільшими можливостями, оскільки дозволяють цей трикутник перетворити на будь-який інший трикутник. Тому довільний паралелограм можна перетворити на квадрат. Для цього треба задати афінне перетворення так, щоб трикутник перетворився на рівнобедрений, прямокутний трикутник. Афінне перетворення дозволяє перетворити будь-яку трапецію на рівнобедрену трапецію. Зрозуміло, що за допомогою афінних перетворень можна вирішувати лише завдання, пов'язані з афінними властивостями фігур.

В наступних дослідженнях буде розглядатися процес афінного перетворення у таких ситуаціях: перегляд тривимірних об'єктів із різних точок зору, проектування тривимірних об'єктів на площину та відображення сцени у вікні, а також складання плоского зображення або тривимірної сцени шляхом компонування з однотипних елементів за допомогою копіювання, перетворення і переміщення в різні місця зображення.

Висновки

Афінні перетворення знаходять застосування в різних галузях науки та техніки. Особливе місце займають комп'ютерні науки. З розвитком комп'ютерної графіки при роботі з тривимірними об'єктами потрібно здійснювати по відношенню до них різні перетворення: рухати, повертати, стискати або розтягувати. Але важливо, щоб після застосування цих змін зберігалися певні властивості. В цій роботі розглянутий процес перетворення паралелепіпеда на куб, а також побудована його проєкція на площину.

ЛІТЕРАТУРА

1. Томашевський В.М. Моделювання систем – Київ, Видавнича група ВНУ, 2019. – 352 с.
2. Рудавський Ю. К., Костробій П. П., Луник Х. П. Лінійна алгебра та аналітична геометрія. – Львів : Видавництво Державного університету «Львівська політехніка», 2020. – 262 с.
3. Боресков А., Шикін Е. Комп'ютерна графіка: від пікселів до програмованого графічного обладнання. Бока Ратон : CRC Press, 2020. – 568 с.
4. Яковець В. П., Боровик В. Н., Ваврикович Л. В. Аналітична геометрія. Суми : Університетська книга, 2018. – 296 с.
5. Руденко Є.А, Семушева О.Ю. Афінні перетворення об'єктів у комп'ютерній графіці на прикладі 3D-моделі, що обертається – Київ : Нацкова думка, 2020. – 102 с.
6. Kokkos A. Expanding Transformation Theory Affinities between Jack Mezirow and Emancipatory Educationalists. – Taylor & Francis Ltd, 2021. – 234 p.

ВЕРЕМІЄНКО Богдан Олександрович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022; e-mail: xa12850433@student.karazin.ua; ORCID: 0000-0001-9699-1168.

Наукові інтереси:

– *інформаційні технології, комп'ютерна графіка, математичне моделювання технічних і технологічних процесів.*

ЧУБ Ольга Ігорівна – к.е.н., доцент кафедри Теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022; e-mail: o.i.chub@karazin.ua; ORCID: 0000-0002-1216-856X.

Наукові інтереси:

– *математичне та комп'ютерне моделювання складних систем, логістичне управління.*

ВОЛИНСЬКИЙ В.В, МІРОШНИК М.А.

РОЗРОБКА МОДЕЛІ НАУКОВО-ДОСЛІДНИЦЬКОЇ ЛАБОРАТОРІЇ

Theme of the research: development of a scientific research laboratory model

The main purpose of the research - the purpose of the work was to consider the main concepts of the Internet of Things. What is the technology based on, where it can be used. The main disadvantages and possible vulnerabilities of the Internet of Things. Its prospects for the future are also discussed. Based on the obtained data, was decided to create a model of a research laboratory using the Internet of Things, namely the smart home and the industrial IoT. Such a model should make independent decisions, autonomously process data, be managed remotely, be scalable and easily adopt changes.

The methods of the research - To build this model, it was decided to use the Cisco Packet Tracer. Python and JavaScript programming languages were used to create the necessary devices that are not in the application. Tests for devices are written in the same programming languages.

Results - the model of a scientific research laboratory was created. It consists of chemical and biological laboratories created based on the knowledge gained from the Internet of Things and with the help of the Cisco Packet Tracer application. The application provides a visual flow of all processes, remote control and shows how these systems can be integrated into real laboratories.

Conclusions - we can say that the Internet of Things technology has great potential. The study of applications that can create models of the use of the Internet of Things takes place. This will make it possible to clearly show all processes at the modeling stage, find errors or vulnerabilities and save time. The model can speed up the work process, ensure the safety of employees and adequately respond to unusual situations by itself.

Key words: DNS, DHCP, IoT, Internet of Things, IP, WAN, Cisco Packet Tracer, ABCDE.

Огляд сучасного стану Інтернет речей та “розумні” пристрої

Інтернет речей об'єднує пристрої в комп'ютерну мережу і дозволяє збирати аналізувати, обробляти та передавати дані іншим об'єктам через програмне забезпечення, додати та технічні девайси. Сама концепція була сформована ще в 1999 році. Але поштовх в широкомасштабному використанні отримала тільки в 2010. З тих пір в Брюсселі кожен рік проводиться конференція, на якій демонструють свої напрацювання чиновники з усього ЄС, представники передових корпорацій, таких як SAP, SAS[7]. А в США заявили, що вважають технологію Інтернету речей однією з шести підричних технологій[8].

Інтернет речей функціонує самостійно, хоч люди іноді можуть давати їм доступ до вже існуючих даних. Ці системи працюють в режимі реального часу, зазвичай складаються з мережі розумних пристроїв в хмарній платформі, до якої вони підімкнуті (під'єднані) за допомогою WiFi, Bluetooth чи інших видів зв'язку.

Виникає питання: з чого складається інтернет речей? Експерти ввели коротку та зрозумілу аббревіатуру: ABCDE. А саме:

- Analytics - головна частина Інтернету речей, яка об'єднує пристрої, дані з них та оптимізує бізнес-процеси;
- BigData - інформація з пристроїв зберігається в хмарі. Вона дозволяє автоматизувати існуючі процеси чи додавати нові;
- Connection - це канали, по яким пристрої отримують і передають інформацію;
- Devices - пристрої, які підімкнуті до мережі, які для коректної роботи в залежності від задач повинні мати свою частоту повідомлень
- Experience - робота з уже отриманим досвідом від інших пристроїв IoT.

Але де ж використовується інтернет речей? Він дозволяє компаніям автоматизувати процеси та знизити труда затрати. Це зменшує об'єм відходів, покращує якість послуг та робить логістику і дешевше. Тому майже всі передові компанії світу уже активно використовують Інтернет речей, а компанії поменше активно інтегрують в усі процеси управління чи розробки своєї продукції. Україна не є винятком, тому що активно розвиваються так названі “розумні міста”. За даними IoT Analytics[2] розподіл ринку має наступний вид:

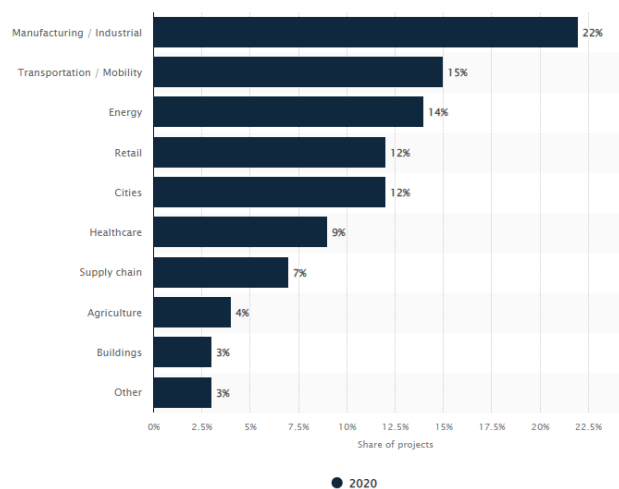


Рис. 1 Розподіл ринку Інтернету речей.

Наприклад в логістиці, Інтернет речей зменшує витрати на грузоперевозку, мінімізує вплив людського фактору. Системи IoT також можуть оптимізувати розходи, виходячи із існуючих чи новоздобутих даних. В електроенергії інтернет речей дозволяє покращити контроль підстанцій чи ліній електропередач. Або керувати електроенергією в приміщенні чи поверсі.

Навіть в Інтернету речей є свої недоліки. По-перше, основні технології, які використовує IoT: сенсорні технології, 5G ще самі по собі “молоді” і їм є куди розвиватися. Такий розвиток створює можливості для створення нового хакерського ПО, яке може негативно впливати на саму мережу IoT, так і на процес отримання та обробки даних. Тому і зараз хакерські атаки на IoT дуже поширені - це і DDoS атаки, і фішинг або навіть диверсії на підприємствах. Також важливим недоліком є те, що декілька пристроїв, які становлять одну систему можуть мати несумісність ПО. Наприклад, якщо загрузити оновлене програмне забезпечення на один з пристроїв, то інші можуть конфліктувати з ним, тому що мають “старі” версії програм. І напевно найбільшим недоліком є те, що пристрої IoT з часом замінять людей на робочих місцях. Та незважаючи на це все, технологія Інтернету речей має шалені перспективи. На рис.2 можна побачити передбачення тенденцій розвитку[1].

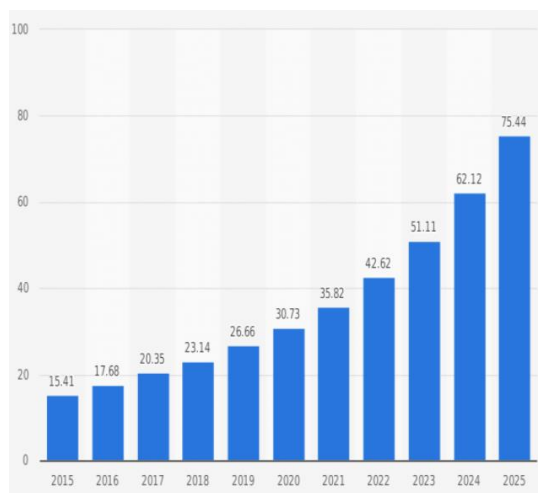


Рис. 2 Перспективи розвитку Інтернету речей

Тобто, вже через 2 роки пристроїв, які будуть підключені до мережі Інтернет стане на 50% більше, ніж зараз. Це високий рівень росту, що показує всі перспективи розвитку технологія і доцільність її вивчення і використання як в навчальних процесах, так і в науково-дослідницьких.

Постановка задачі дослідження

Після дослідження технологія Інтернету речей, можна сказати, що за допомогою спеціальних додатків легко можна створити модель, яка має всі характеристики для проведення наукових праць і дослідів. Такий підхід має ряд переваг, а саме: автоматичне прийняття рішень; забезпечення комфорту і безпеки; виявлення несправностей, проблем лабораторії ще на етапі моделювання, що дозволить зекономити кошти та людино-години; візуалізація всіх процесів в залежності від впливу зовнішнього середовища. Для отримання цих результатів, потрібно дослідити наступні питання: 1) аналіз сучасного Інтернету речей; 2) аналіз “розумного” дому та індустріального Інтернету речей; 3) вивчити як можна самостійно створювати пристрої IoT, якщо вони відсутні в додатку; 4) розробити методику тестування моделі і новостворених пристроїв вчасності; 5) створити модель науково-дослідницької лабораторії на основі вибраного додатку; 6) налаштувати взаємодію компонентів моделі і управління нею через мережу Інтернет; 7) налаштувати чи створити власні параметри зовнішнього середовища і наглядно показати протікання всіх процесів та самостійність прийняття рішень моделі.

Існуючі програмні засоби для створення моделей з використанням технології Internet of Things

Вибрати додаток для моделювання науково-дослідницької лабораторії надважливий етап. Адже варто розрізняти емулятори і симулятори. На сьогоднішній день найпопулярнішими є наступні програмні продукти: Cisco Packet Tracer, NetSim, Dynamips. У кожного свої плюси, мінуси та особливості. Наприклад в NetSim дуже мало пристроїв Інтернету речей і відсутня можливість створювати самостійно, так як і в Dynamips, котрий до того ж, сильно навантажує операційну систему, тому що встановлює на комутатори реальні образи ISO.

Тому для створення такої масштабної моделі краще всього взяти симулятор Cisco Packet Tracer. На це є ряд причин. По-перше: цей симулятор включає можливість розподілення моделі на фізичні та логічні рівні, що дозволяє легко масштабувати, налаштовувати зв'язки між рівнями чи навпаки розділити елементи моделі на незалежні компоненти. По-друге: в додатку присутня повна можливість використовувати Інтернет речей. Присутні як пристрої “розумного дому” так і індустріального Інтернету речей, що значно покращує можливості та сфери використання тих чи інших пристроїв[3]. По-третє, наявна можливість створювати практично будь-які пристрої IoT самостійно, використовуючи мови програмування Python та JavaScript. Важливо правильно написати тести для таких пристроїв, щоб вони реагували тільки на задані для них параметри. По-четверте, надзвичайно широкі можливості задання параметрів зовнішнього середовища. починаючи від рівня кисню в кімнаті і закінчуючи інтенсивністю ультрафіолетового випромінювання [4]. По-п'яте, широка доступність мережевих пристроїв, таких як комутатори, хаби, VoIP -пристрої, DHCP-сервери, DNS-сервери. Все це дозволяє з легкістю організувати передачу даних між елементам як одного логічного рівня, так між різними[5]. Також за допомогою індустріальних пристроїв, є можливість контролю моделлю за допомогою Інтернету. На рис. 3 зображено всі типи пристроїв, які можна додавати в симуляторі Cisco Packet Tracer.



Рис. 3 Прилади “розумного дому” Cisco Packet Tracer

Серед приладів для створення моделей є:

- 1) мережеві пристрої, які діляться на маршрутизатори, комутатори, хаби, бездротові девайси, WAN-емюлятори;
- 2) кінцеві пристрої, які діляться на: пристрої “розумного дому” та “розумного міста”, індустріальний IoT та енергосистему;

- 3) плати, за допомогою яких можна створювати власні пристрої та підключення, які діляться на: звичайні підключення та структуровану кабельну проводку.

Тобто можна сказати, що використання всіх можливостей цього емулятора дозволить створити таку модель науково-дослідницької лабораторії, яка буде візуально відображати всі особливості тієї чи іншої лабораторії, завчасно і самостійно реагувати на екстрені ситуації, самостійно приймати рішення. Буде легко масштабованою, керованою будь-звідки, завдяки налаштуванню дротових зв'язків разом з бездротовим. Це все збільшує безпеку персоналу чи студентів, які працюють в лабораторії, пришвидшує процес наукових досліджень і дає впевненість в тому, що в випадку надзвичайних ситуацій система зробить все можливе для забезпечення безпеки працівників і самої себе, щоб не втратити наукові напрацювання.

Фізична топологія моделі

Загалом, модель науково-дослідницької лабораторії складається з трьох частин: моделі хімічної лабораторії; моделі біологічної лабораторії; серверної; Фізичну топологію моделі можна побачити на рис.4.

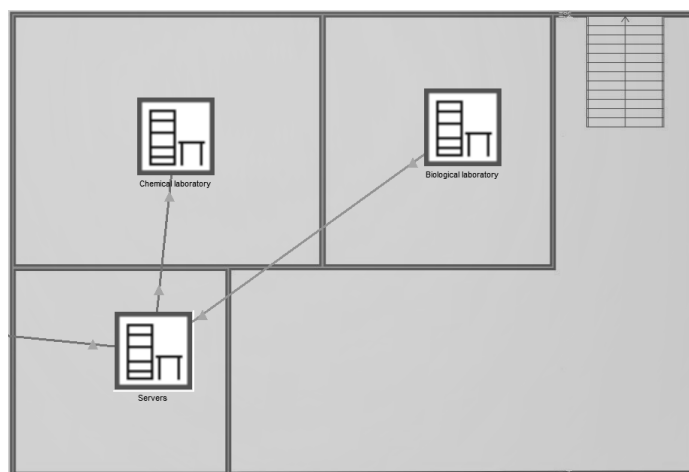


Рис. 4 Фізична топологія науково-дослідницької лабораторії

Для коректної роботи моделі та для віддаленого доступу та керування підсистемами, потрібно налаштувати доступ, задати IP-адреси всім пристроям, створити сервер для IoT пристроїв, де буде зберігатися вся логіка їх функціонування. Для цього потрібно створити DHCP та DNS- сервера, створити та дозволити обмінюватися даними через WAN. Створення DHCP та WAN можна побачити на рис.5.

```

IOS Command Line Interface
%LINK-5-CHANGED: Interface GigabitEthernet0/0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet0/0, changed state to up
Router(config-if)#int g0/2
Router(config-if)#ip address 209.165.201.225 255.255.255.224
Router(config-if)#no shutdown

Router(config-if)#
%LINK-5-CHANGED: Interface GigabitEthernet0/2, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet0/2, changed state to up
Router(config-if)#int g0/1
Router(config-if)#ip address 209.165.200.225 255.255.255.224
Router(config-if)#no shutdown

Router(config-if)#
%LINK-5-CHANGED: Interface GigabitEthernet0/1, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface GigabitEthernet0/1, changed state to up
Router(config-if)#exit
Router(config)#ip dhcp excluded-address 209.165.201.225 209.165.201.229
Router(config)#ip dhcp pool CELL
Router(dhcp-config)#network 209.165.201.224 255.255.255.224
Router(dhcp-config)#default-router 209.165.201.225
Router(dhcp-config)#dns-server 10.0.0.254
Router(dhcp-config)#exit
Router(config)#ip dhcp excluded-address 209.165.200.225 209.165.200.229
Router(config)#ip dhcp pool WAN
Router(dhcp-config)#network 209.165.200.224 255.255.255.224
Router(dhcp-config)#default-router 209.165.200.225
Router(dhcp-config)#

```

Рис. 5 Створення WAN та DHCP

Завдяки таким налаштуванням кожен пристрій моделі науково-дослідницької лабораторії отримує свій адрес в мережі і може бути ідентифікований іншими пристроями та приймати команди з будь-якого пристрою, який має доступ до системи керування. Можна організовувати підсистеми, адже кожен пристрій може обмінюватися інформацією з будь-яким іншим.

Тобто, створення такої моделі, яка в залежності від впливу зовнішнього середовища, дій персоналу чи студентів, дає змогу автоматично приймати оптимальні рішення, швидко обробити великі об'єми даних та на основі них приймати рішення уже в майбутньому, своєчасно сповіщати персонал та інженерів. Такий підхід на етапі моделювання покаже всі плюси і мінуси інтеграції Інтернету речей в реальні лабораторії, можливі технічні проблеми. Це дозволить завчасно уникнути проблемних ситуацій. Організація доступу з будь-якої точки за допомогою Інтернету має свої переваги, адже навіть віддалено можна слідкувати та коригувати роботи підсистем. Використання Інтернету речей допомагає пришвидшити всі процеси, зробити робочий процес менш утомливим, а слідкування за безпекою та своєчасне реагування на нестандартні ситуації довірити "розумним" пристроям.

ЛІТЕРАТУРА

- 1) Hassan, Q.F.; Khan, A. ur R.; Madani, S.A.. Internet of Things: Challenges, Advances, and Applications. – 2017. P. 39-41
- 2) Bastos, D.; Shackleton, M.; El-Moussa, F. (2018). "Internet of Things: A Survey of Technologies and Security Risks in Smart Home and City Environments". Living in the Internet of Things: Cybersecurity of the IoT - 2018. P. 30
- 3) Adrian McEwan: Designing the Internet of Things – 2016. P. 165-180
- 4) Andrea Finardi. IoT simulations with Cisco Packet Tracer. [Електронний ресурс]. Режим доступу до ресурсу <https://www.theseus.fi/bitstream/handle/10024/150158/Andrea%20Finardi%20%20Master%20of%20Engineering%20%20Information%20technology.pdf?sequence=1&isAllowed=y>. – 2017.
- 5) Zhang, Yongbin; Liang, Ronghua; Ma, Huiling (2012). "Teaching Innovation in Computer Network Course for Undergraduate Students with Packet Tracer" - 2018. P. 504–510
- 6) Cisco Packet Tracer. Cisco Networking Academy. [Електронний ресурс]. Режим доступу до ресурсу: <https://www.netacad.com/portal/learning>. – 2021
- 7) Want, Roy; Schilit, Bill N.; Jenson, Scott. "Enabling the Internet of Things" – 2015. P. 28–35.
- 8) Nordrum, Amy "Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated" – 2018. P. 65-71.

ВОЛИНСЬКИЙ Валентин Володимирович – кафедра прикладної і теоретичної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: xa11867787@student.karazin.ua; ORCID 0000-0002-2692-2279

Наукові інтереси:

– *технологія Internet of Things;*

МІРОШНИК Марина Анатоліївна – д.т.н., професор; професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна; м.Харків, майдан Свободи, 6, 61022; e-mail: m.miroshnyk@karazin.ua, ORCID 0000-0002-2231-2529

Наукові інтереси:

– *комп'ютерні системи та компоненти, розподілені обчислення, Інтернет речей;*

ГАМЗАЄВ Р.О., ШКОДА Б.В.

МОДЕЛІ ТА ПРИНЦИПИ ПОБУДОВИ ЛІНІЙКИ ПРОГРАМНИХ ПРОДУКТІВ НА ПРИКЛАДІ СИСТЕМ ОБМІНУ ПОВІДОМЛЕННЯМИ

Вступ

Системи обміну повідомлень (СОП) використовуються як незалежні програмні продукти, а також як додаткова функціональність у великій кількості програмних рішень. Крім того в залежності від потреб користувачів та їх ролей СОП можуть мати різний набір функціональних елементів, а також доступні функції можуть змінюватись динамічно в залежності від таких характеристик як права, версія, тощо. Розробка окремих рішень для кожного можливого користувача, а також їх динамічна адаптація може бути дуже затратною і буде вимагати багато часу, тому доречно скористатися одним з можливих підходів для розробки так званих лінійок програмних продуктів (software product line - SPL) та динамічних лінійок програмних продуктів (DSPL) [1]. Одна з можливих моделей та принципи побудови СОП на основі концепції DSPL розглянуті у цієї статті.

Актуальність та мета дослідження

З появою всесвітньої мережі Інтернет, у людей з'явилась потреба у комунікації, яка і надалі, з кожним роком зростає. На ринку існує багато різних систем, що вирішують дану проблему в різних контекстах. Але такі системи не є простими у підтримці, розробці, тестуванні і, тому, не кожна компанія здатна з нуля розробити схожий продукт під свої потреби, в більшості моментів, не через нестачу інженерного персоналу, а через високі витрати на сам процес розробки, починаючи з архітектурних концепцій, закінчуючи процесом повного (end-to-end) тестування готового програмного рішення. Крім того, час, що потрібен на розробку таких додатків є теж проблемою. Тому дана тема є актуальною та робота покриває вирішення описаних проблем за допомогою підходів до розробки DSPL [2]. Метою дослідження є не просто розробка продукту з використанням DSPL на основі СОП, а визначення загального процесу до побудови СОП з використанням підходів DSPL, аналіз доменної моделі СОП для подальшого проектування архітектури, розробки та тестування системи. Крім того, програмне рішення повинно вміти адаптуватись під сучасні проблеми, мати можливість повного конфігурування функціоналу, чи як ще його називають - feature. Кожна компанія чи користувач, що будуть використовувати даний генератор повинні на виході отримати готовий продукт під їх вимоги, тому глибокий аналіз усіх можливих функцій є невід'ємною частиною даної роботи, що повністю відрізняється від традиційних підходів до розробки програмного забезпечення (ПЗ).

Основні принципи

SPL та DSPL включають в себе багато практик, інструментів, підходів задля розробки адаптивного ПР. Загальні процеси, що дозволяють побудувати необхідну SPL описав К. Поль (K. Pohl et al.) [3]. На рисунку зображено 2 основні рівня до розробки SPL - рівень доменної інженерії (Domain Engineering) та рівень інженерії додатків (Application Engineering). На рівні доменної інженерії, створюються різні варіабельні рішення для багаторазового використання як моделі вимог, архітектурні моделі, компоненти домену, моделі тестування та ін. На рівні інженерії додатків, на основі отриманих проєктних рішень з доменного рівня, реалізують поставлені вимоги, програмні компоненти, архітектуру, тести тощо. Для кожного рівня існує велика кількість методів розробки ПЗ та CASE-інструментів, що стисло описані в [4].

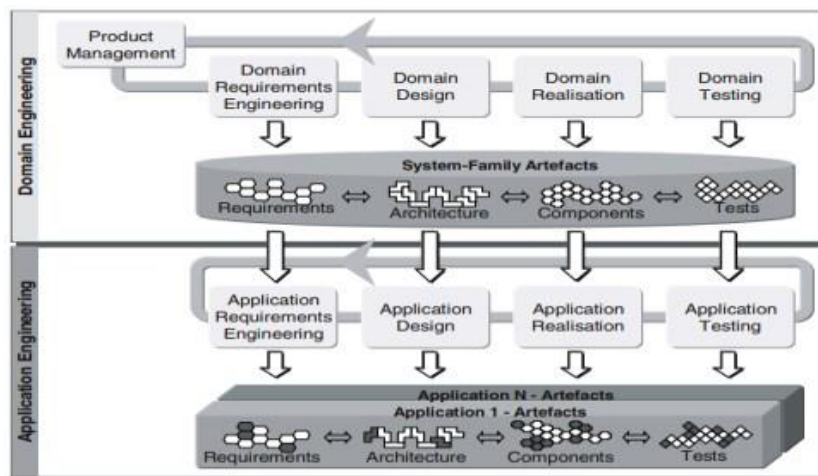


Рис. 1 Процеси розробки програмного рішення на основі концепції SPL

Під час доменної інженерії, використовується підхід FODA, що допоможе описати модель варіабельності, провести детальний аналіз контексту з усіма ключовими сутностями, зв'язками. Цей підхід також включає проектування функціональної моделі обраної предметної області (ПрО), тобто месенджери.

Для інженерії додатків, використовуються основні принципи FOSD, що дозволяє побудувати адаптивне ПР, що є не просто розширюваним, а також легким у підтримці (maintainability), розгортанні (deployment), оптимізаціях (optimizations).

Незважаючи на те, що моделі та методи розробки лінійок програмного забезпечення (SPLE) використовуються для керування мінливістю в парадигмах DSPL, інженерні парадигми SPLE та DSPL переслідують різні цілі, кожна з яких призводить до різних продуктів. Хальстайзен (Hallsteinsen et al.) [5] характеризує інженерію DSPL за допомогою набору властивостей, які відрізняють її від інженерії SPL з точки зору мети та властивостей їхніх продуктів. Наприклад, метою вибору іншого варіанту системи в SPL є задоволення конкретних потреб ринку; в той час як мета в парадигмі DSPL полягає в задоволенні вимог різних робочих контекстів [6]. Хальстайзен [5] також визначає два типи властивостей для інженерних продуктів DSPL, які відрізняють їх від продуктів SPLE: реконструкція під час виконання та самокерування. Властивості реконструкції під час виконання представляють здатність продукту «конфігурувати та прив'язувати змінність під час виконання», «прив'язувати зміни кілька разів протягом життя», «змінювати точки варіації» та «впоратися з неочікуваними змінами» [5]. Властивості самоуправління представляють здатність продукту «зважати на контекст», «самоадаптивно» та функціонувати як «автоматичний засіб прийняття рішень» [5]. Властивості реконструкції під час виконання вважаються обов'язковими для продукту, але властивості самокерування вважаються необов'язковими в контексті DSPL [2, 7, 8].

Особливості моделювання

Основний функціонал месенджерів - обмін миттєвими повідомленнями задля комунікації між людьми. Загалом, повідомлення містять текстову інформацію, але, в залежності від реалізації, можуть мати можливість відправляти ще й графічну інформацію, документи, аудіо файли тощо. Додатково, в месенджерах, крім комунікації, загальними функціями є управління даними користувачів (налаштування профілю), пошук користувачів, додавання користувачів до друзів, створення каналів для групової комунікації та ін.

Розробка системи на основі DSPL орієнтована на функціонал СОП, тобто одним з підходів що використовується є функціонально-орієнтований підхід до розробки ПЗ (FOSD), тому, перед розробкою архітектури та програмного рішення необхідно провести детальний аналіз функцій, що є також використанням підходу доменного аналізу орієнтованого на FODA та побудувати функціональну модель, яка надалі буде використовуватись для конфігурування програмного рішення [3]. Функціональна та компонентна моделі зображені на рисунку 2.

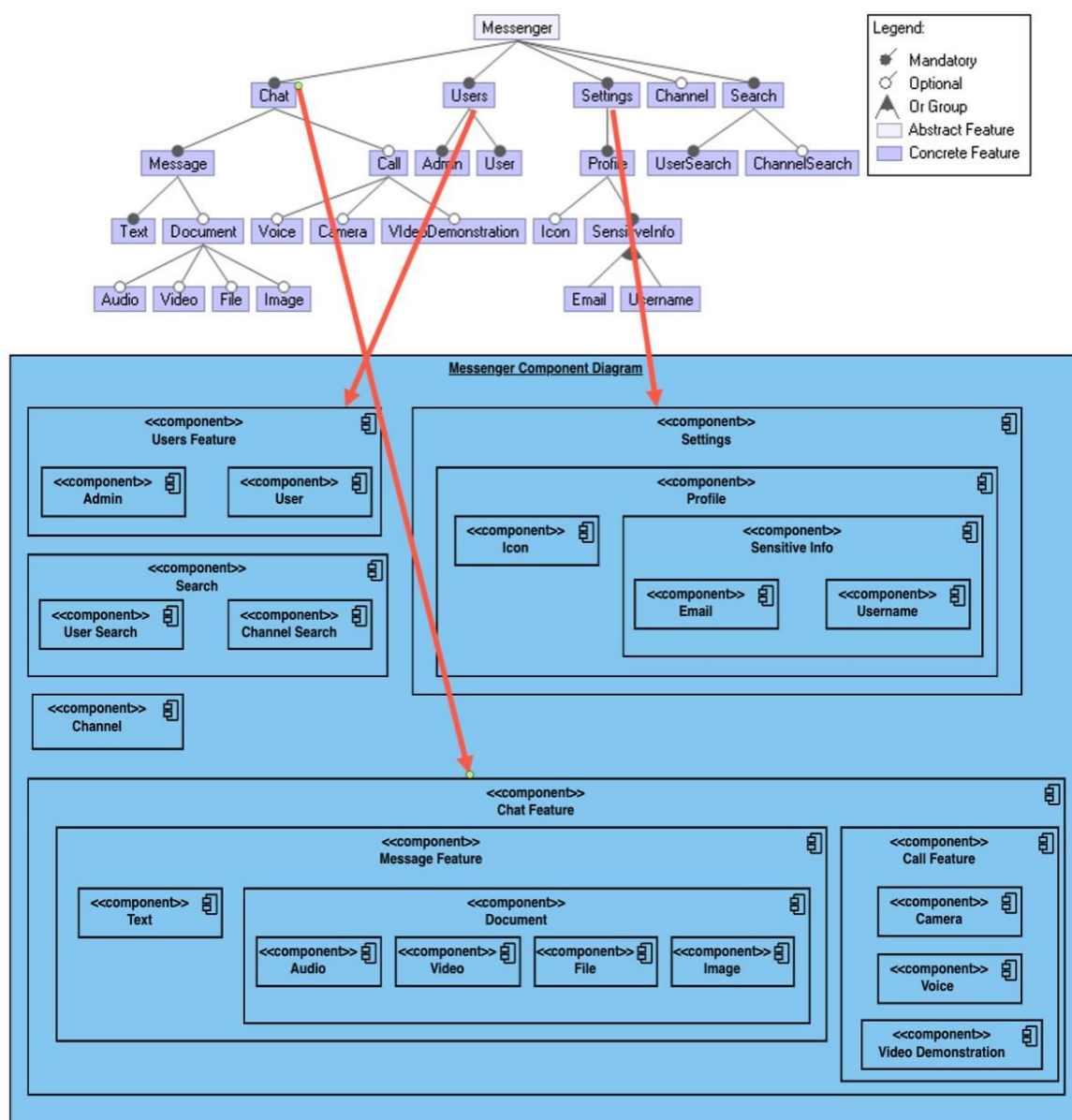


Рис. 2 Функціональна та компонентна моделі системи обміну повідомлень

FODA модель описує набір функцій у домені та їхні зв'язки. Вона описує, які функції надає DSPL (тобто її обсяг), які не обов'язковими та в які комбінації функцій можна вибрати для отримання прототипів. Кількість потенційних варіантів може експоненціально зростати зі збільшенням кількості функцій. В моделі зазначено чорними точками ті функції, що не можуть бути увімкненими ще до генерації програмного рішення, а також білими, що є опціональними. Також, модель описує операції OR що дозволяють обрати певну функцію зі списку можливих.

Висновки

Запропоновані підходи та моделі дозволять розробити та отримати очікуваний результат у вигляді системи, здатної конфігурувати ПР на прикладі СОП під будь які умови та вимоги. Розроблені ПР дозволять уникнути описані раніше проблеми компаній, швидко та ефективно запуснути згенеровані рішення для різних доменних контекстів, потребуючих реалізації у вигляді СОП.

ЛІТЕРАТУРА

1. Eleutério J., “A Comparative Study of Dynamic Software Product Line Solutions for Building Self-Adaptive Systems”, Technical Report in Eniversidade estadual de Campinas, 30 p. 2017.
2. Bashari, Mahdi & Bagheri, Ebrahim & Du, Weichang. (2017). Dynamic Software Product Line Engineering: A Reference Framework. International Journal of Software Engineering and Knowledge Engineering. 27. 191-234. 10.1142/S0218194017500085.
3. Pohl K., et al: Software Product Line Engineering: Foundations, Principles, and Techniques., Springer, 2005. P. 467.
4. Ткачук М.В., Гамзаєв Р.О., Товстокоренко О.Т., Мартінкус І.О., Нагорний К.А. Аналіз та експериментальне дослідження CASE - засобів для моделювання варіабельності в процесах розробки лінійок програмних продуктів // Системи управління, навігації та зв'язку: зб. наук. пр. ЦНДІ НУ. – К., 2020. – Вип. 1 (45).– С. 105-110.
5. S. Hallsteinsen, E. Stav, A. Solberg and J. Floch, Using product line techniques to build adaptive systems, in Software Product Line Conference, 2006, pp. 10–150.
6. S. Hallsteinsen, M. Hinchey, S. Park and K. Schmid, Dynamic software product lines, Computer 41(4) (2008) 93–95.
7. L. Baresi, S. Guinea and L. Pasquale, Service-oriented dynamic software product lines, Computer 45(10) (2012).
8. N. Bencomo, P. Sawyer, G. Blair and P. Grace, Dynamically adaptive systems are product lines too: Using model-driven techniques to capture dynamic variability of adaptive systems, in 2nd Int. Workshop on Dynamic Software Product Lines, 2008.
- 9.

ГАМЗАЄВ Рустам Олександрович – доцент кафедри моделювання систем і технологій; Харківський національний університет імені В.Н. Каразіна, м. Харків, Майдан Свободи 6, 61022; e-mail rustam.gamzayev@karazin.ua; ORCID: 0000-0002-2713-5664

Наукові інтереси:

– аналіз та моделювання розподілених програмних систем, моделі та технології розробки лінійок програмних продуктів, знання-орієнтовані методи розробки та інженерії якості програмного забезпечення.

ШКОДА Богдан Владиславович - студент кафедри моделювання систем і технологій; Харківський національний університет імені В.Н. Каразіна, м. Харків, Майдан Свободи 6, 61022; e-mail xa11867771@student.karazin.ua; ORCID: 0000-0001-6312-8384

Наукові інтереси:

– об'єктно орієнтоване програмування, програмування баз даних, проектування високо-навантаженої архітектури, аналіз та проектування, що орієнтовані на доменах (domain-driven design)

ГЛУХОТА А.В, ШЕВЧЕНКО Д.О, СТРИЛЕЦЬ В.Є.

НЕЙРОМЕРЕЖЕВІ ТЕХНОЛОГІЇ ДЛЯ ОБРОБКИ ШУМУ В ЗОБРАЖЕННЯХ

Актуальність: існує багато алгоритмів усунення шуму зображення, але проблема підвищення якості знешумлення зображення залишається відкритою. Тому виникає потреба у вдосконаленні методів знешумлення зображення за рахунок використання автокодера.

Мета: метою роботи є підвищення рівня стійкості до шуму в робототехнічних системах шляхом використання автокодера для зменшення шуму (denoising autoencoder).

Методи дослідження: метод для зменшення шуму на основі згорткових нейронних мереж, загальний автокодер, автокодер для зменшення шуму (denoising autoencoder).

Результати: проведено аналіз математичних моделей та методів для зменшення шуму та зроблений опис загального автокодера та детальний опис автокодера для зменшення шуму (denoising autoencoder).

Висновки: головним результатом роботи був огляд автокодера для зменшення шуму в зображенні та його відновлення в робототехнічних системах. В наш час використання згорткових нейронних мереж є прогресивним підходом. Головним напрямком використання автокодера для зменшення шуму є використання в робототехнічних системах моніторингу.

На основі отриманих результатів далі буде розроблена програмна реалізація автокодера для зменшення шуму в зображеннях та проведення його тестування при різних об'ємах вхідних даних.

Ключові слова: згорткові мережі, зменшення шуму, автокодер, машинне навчання, робототехнічні системи.

Постанова задачі та її актуальність

На даний момент у робототехнічних системах активно використовується робота із зображеннями (наприклад, обробка знімків з камер), але часто при отриманні зображення виникають шуми.

Зазвичай виникнення шумів викликано різними внутрішніми та зовнішніми умовами, які часто неможливо уникнути в практичних ситуаціях. Таким чином, зменшення впливу шуму на зображення відіграє важливу роль у широкому спектрі додатків, таких як робототехнічні системи, відновлення зображень, реєстрація зображень, сегментація зображень та класифікація зображень, де отримання вихідного вмісту зображення має вирішальне значення для високої продуктивності.

На сьогодні існує чимало підходів видалення шумів, окремо можна виділити використання згорткових нейронної мережі для загальних зображень: ADNet(attention-guided CNN)[1], NERNet(Noise estimation removal network - NERNet)[2], MP-DCNN[3], SW-CNN(side window CNN)[4], PCLDCNet(patch complexity local divide and deep conquer network)[5], прямий метод CNN[6], SSDRN(spectral-spatial denoising residual network)[7].

Хоча дослідниками було запропоновано багато алгоритмів усунення шуму зображення, проблема знешумлення зображення залишається відкритою. Тому *метою* роботи є підвищення рівня стійкості до шуму в робототехнічних системах шляхом використання автокодера для зменшення шуму (denoising autoencoder)

Нейромережева технологія для обробки шуму в зображеннях

Проаналізувавши різні методи зниження шуму, найкращим методом для робототехнічних систем був визначений метод з використанням автокодерів. Визначимо поняття автокодера. Автокодери були вперше представлені в книзі «Parallel distributed processing: Explorations in the microstructure of cognition»[8] як нейронні мережі, які навчаються реконструювати свої вхідні дані.

Автокодер – це специфічний тип нейронної мережі, яка в основному призначена для кодування вхідних даних у стиснене та змістове уявлення, і потім для декодування його назад так, щоб реконструйований вхід був максимально схожий на оригінал. Зазвичай автокодер складається з трьох основних частин: кодер(encoder), декодер(decoder), і код(code) як показано на рис. 1. Основна мета кодера – стиснути дані таким чином, щоб їх основні характеристики були збережені, таким чином допомагаючи декодеру відновити дані зі стисненого представлення.

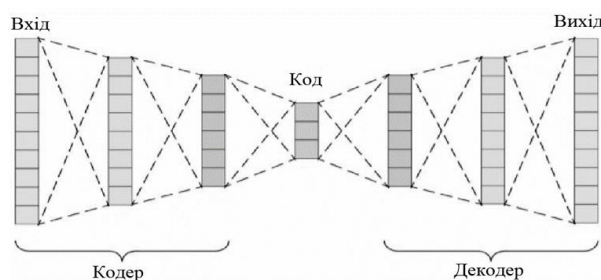


Рис. 1. Архітектура автокодера

Автокодери навчаються за допомогою алгоритму зворотного поширення таким чином, щоб різниця між вихідними і оригінальними даними була мінімізована.

Проблема автокодера, як формально визначено в статті «Autoencoders, unsupervised learning, and deep architectures»[9], полягає в тому, щоб вивчати функції A : (кодер) і B : (декодер) що задовольняє:

$$\arg \min_{A,B} E[\Delta(x, B \circ A(x))] \quad (1)$$

де $\square$ - очікування щодо розподілу $\square$, а Δ - втрата на реконструкцію функція, яка вимірює відстань між виходом декодера і введенням.

У найпопулярнішій формі автокодерів $\square$ і $\square$ є нейронними мережами. У окремому випадку, коли $\square$ і $\square$ є лінійними операціями, отримується лінійний автокодер. У випадку лінійного автокодера, також відкидаються нелінійні операції, тому автокодер досягне такого ж прихованого представлення, що й аналіз основних компонентів (Principal Component Analysis). Таким чином, автокодер фактично є узагальненням аналізу основних компонентів, де замість пошуку низьковимірної гіперплощини, в якій лежать дані, вона може вивчити нелінійне різноманіття.

Для зменшення впливу шуму запропоновано використовувати автокодер для зменшення шуму (denoising autoencoder). Він відноситься до виду регуляризованих автокодерів. Знешумлювані автокодери є розширенням базової архітектури автокодерів. Нейронна мережа автокодера намагається відновити зображення з прихованого простору коду. При знятті шумів автокодерів вносяться деякі шуми в зображення. Мережа автокодера, що знімає шум, також намагатиметься відновити зображення. Але перед цим він повинен буде приглушити шум від вхідних даних зображення. При цьому автокодер навчиться фіксувати всі важливі характеристики даних. Приблизна архітектура знешумлюваного автокодера зображена на рис. 2.

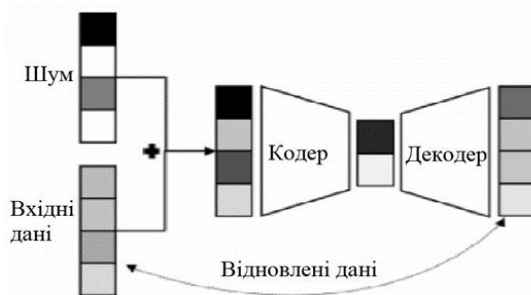


Рис. 2. Архітектура знешумлюваного автокодера

Зменшення шумів рекомендується для навчання моделі, а DAE (denoising autoencoder) надає моделі два важливі аспекти: DAE зберігають вхідну інформацію (кодування введення) і DAE намагаються видалити шум, доданий до входу автоматичного кодера. Знешумлювані автокодери працюють краще порівняно з традиційними фільтрами для зменшення шуму, оскільки їх можна модифікувати на основі вхідних даних, на відміну від традиційних фільтрів, які не залежать від даних.

При зменшенні шуму дані певним чином пошкоджуються через додавання випадкового шуму, і модель навчається передбачати вихідні пошкоджені дані. Інший варіант цього полягає в пропуску частин вхідних даних на відміну від додавання шуму до входу, щоб модель могла навчитися передбачати вихідне зображення. У цьому випадку ідея полягає в тому, щоб зберігати

вихідні дані, згенеровані кодером, у вигляді вектора ознак, який можна використовувати в підході прогнозування поїздів із контрольованою моделлю.

Результати та висновки

У роботі був проведений аналіз математичних моделей та методів для зменшення шуму; описаний автокодер для зменшення шуму.

Головним результатом наукової роботи є автокодер для зменшення шуму в зображеннях та його відновлення при використанні в робототехнічних системах, тож на основі отриманих результатів далі буде розроблена програмна реалізація автокодера для зменшення шуму в зображеннях та проведення його тестування при різних об'ємах вхідних даних.

Автокодер для зменшення шуму слід використовувати в робототехнічних системах моніторингу.

ЛІТЕРАТУРА

1. Schwenker F, Kestler HA, Palm G (2001) Three learning phases for radial-basis-function networks. *Neural Netw* 14(4):439–458.
2. Guo B, Song K, Dong H, Yan Y, Tu Z, Zhu L (2020) NERNet: noise estimation and removal network for image denoising. *J Vis Commun Image R.*
3. Gai S, Bao Z (2019) New image denoising algorithm via improved deep convolutional neural network with perceptive loss. *Expert Syst Appl.*
4. Yin H, Gong Y, Qiu G (2020) Fast and efficient implementation of image filtering using a side window convolutional neural network. *Signal Process.*
5. Hong I, Hwang Y, Kim D (2019) Efficient deep learning of image denoising using patch complexity local divide and deep conquer. *Pattern Recogn.*
6. Islam MT, Rahman SMM, Ahmad MO, Swamy MNS (2018) Mixed Gaussian-impulse noise reduction from images using convolutional neural network. *Signal Process Image Commun.*
7. Xi W, Li Y, Jia X (2018) Deep convolutional networks with residual learning for accurate spectral-spatial denoising. *Neurocomputing.*
8. Rumelhart, D. E., Hinton, G. E., Williams, R. J.: Parallel distributed processing: Explorations in the microstructure of cognition, vol. 1. chap. Learning Internal Representations by Error Propagation, pp. 318–362. MIT Press, Cambridge, MA, USA (1986). URL <http://dl.acm.org/citation.cfm?id=104279.104293>.
9. Baldi, P.: Autoencoders, unsupervised learning, and deep architectures. In: I. Guyon, G. Dror, V. Lemaire, G. Taylor, D. Silver (eds.) *Proceedings of ICML Workshop on Unsupervised and Transfer Learning, Proceedings of Machine Learning Research*, vol. 27, pp. 37–49. PMLR, Washington, USA (2012).

ГЛУХОТА Андрій Валерійович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: andreygluhota03@gmail.com; ORCID 0000-0002-7623-9187.

Наукові інтереси: – *методи обчислювального інтелекту;*
– *прогнозування стану систем на основі даних моніторингу.*

ШЕВЧЕНКО Дмитро Олександрович – аспірант; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: dimyich24@gmail.com; ORCID: 0000-0002-7897-250X.

Наукові інтереси: – *методи обчислювального інтелекту;*
– *прогнозування стану систем на основі даних моніторингу.*

СТРИЛЕЦЬ Вікторія Євгенівна – кандидат технічних наук, доцент кафедри теоретичної та прикладної системотехніки, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: viktoria.strilets@karazin.ua, ORCID 0000-0002-2475-1496.

Наукові інтереси: – *математичне моделювання складних процесів;*
– *методи штучного інтелекту;*
– *методи машинного навчання.*

ГОЛОВЧЕНКО В.Б., СТРИЛЕЦЬ В.Є.

РЕГРЕСІЙНІ МОДЕЛІ ПРОГНОЗУВАННЯ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ

Дослідження присвячене проблемі прогнозування трафіку комп'ютерних мереж. Його **метою** є підвищення точності прогнозування трафіку комп'ютерних мереж на основі алгоритму випадкових лісів та алгоритмів лінійної регресії. Під час виконання даної роботи були розглянуті класичні методи прогнозування стану систем на основі машинного навчання, проведено аналіз методів лінійної регресії, методу випадкових лісів. Для досягнення поставленої мети був знайдений відповідний набір даних, обрано пакет програмного забезпечення, а також метрики для оцінки її ефективності.

Результатом даного дослідження є декілька натренованих моделей, які здатні спрогнозувати стан комп'ютерної мережі, що базуються на методі випадкових лісів та методах лінійної регресії. Областю використання запропонованої моделі та методу є розробка моделей, обчислювальних методів та засобів, направлених на спрощення прогнозування трафіку та стану комп'ютерних мереж.

Ключові слова: прогнозування трафіку, лінійна регресія, випадковий ліс, штучний інтелект, машинне навчання, навчання з вчителем.

Прогнозування мережного трафіку представляє значний інтерес у таких областях як відстеження перевантажень у мережі, контроль потоків даних та мережеве управління. Явище та класифікація аномалій мережевого трафіку з багатьох подій мережі стає корисним при створенні моделі роботи мережі, оптимізації налаштувань обладнання та підтримці заданого QoS. Ретельно підібрана модель трафіку здатна виявити та передбачити найважливіші характеристики мережевого трафіку, такі як короткочасно та довготривалі залежні процеси, самоподібність на великих тимчасових масштабах. В цей час велика увага приділяється методам прогнозування роботи комп'ютерних мереж у різних умовах. У роботі пропонується розроблення методу прогнозування трафіку комп'ютерних мереж на основі алгоритмів машинного навчання. А саме з використанням алгоритму випадкових лісів та лінійного алгоритму регресії.

Регресія – один з розділів машинного навчання, присвячений прогнозуванню цільовій змінній місця на числовій прямій. Завдання регресії – прогноз з урахуванням вибірки об'єктів з різними ознаками. На виході має вийти число (6, 25, 236.58 тощо.), наприклад вартість квартири, ціна цінного паперу після пів року, очікуваний дохід магазину наступного місяця, якість вина при сліпому тестуванні.

Задача регресії може бути розв'язана за допомогою різних методів, кожен з яких має як свої переваги, так і недоліки. До них відносяться дерева регресії і випадкові ліси, метод найменших квадратів (зокрема лінійна регресія), метод опорних векторів, нейронні мережі та інші.

Лінійна регресія — це метод машинного навчання, задача якого моделювання зв'язку між скалярною змінною у та векторною (зазвичай) змінною X . Якщо змінна X також є скаляром, це називається простою регресією. Лінійна регресія – це широко використовувана модель, що використовується для оцінки реальних величин.

У лінійній регресії зв'язок між незалежною та залежною змінними розглядається через лінію, яка зазвичай представляє зв'язок між двома змінними.

Випадковий ліс є одним із універсальних алгоритмів машинного навчання, оскільки може бути застосований для розв'язання різноманітних задач, таких як, класифікація, регресія, кластеризація, відбір ознак тощо. Метод заснований на побудові великої кількості (ансамблю) дерев ухвалення рішень (один із параметрів методу), кожне з яких будується за вибіркою, що отримується з початкової навчальної вибірки за допомогою бутстрепа (тобто вибірки з поверненням) [2].

Метод випадкових лісів має такі *переваги*:

- гарантія захисту від перенавчання навіть в тому випадку, коли кількість ознак перевищує кількість спостережень.

- для побудови випадкового лісу по навчальній вибірці потрібно задати лише два параметри: розмір ансамблю та кількість випадково підібраних ознак, за якими виконується розщеплення в вершинах;
- оцінка ймовірності помилкової класифікації випадкових лісів на основі спостережень, що не входять до навчального бутстреп вибірки, але використані для побудови дерев, забезпечується методом Out-Of-Bag (OOB), запропонований Брейманом [3].
- випадкові ліси можуть використовуватися не тільки для задач класифікації і регресії, а й для задач виявлення найбільш інформативних ознак, кластеризації, виділення аномальних спостережень і визначення прототипів класів;
- навчальна вибірка для побудови випадкового лісу може містити ознаки, виміряні за різними шкалами: числовими, порядковими і номінальними, що неприпустимо для багатьох інших класифікаторів;
- метод допускає легку паралелізацію (тобто програмну реалізацію, придатну для паралельних обчислень), що має велике значення при великому розмірі навчальної вибірки.

Гradientний бустинг дерев регресії – це техніка машинного навчання для завдань класифікації та регресії, яка будує модель передбачення у формі ансамблю слабких моделей, що передбачають, зазвичай дерев рішень. Навчання ансамблю проводиться послідовно на відміну, наприклад, від беггинга. На кожній ітерації обчислюються відхилення передбачання уже навченого ансамблю на навчальній вибірці. Наступна модель, яка буде додана в ансамбль, буде передбачати ці відхилення. Таким чином, додавши передбачення нового дерева до передбачень навченого ансамблю, ми можемо зменшити середнє відхилення моделі, яке є ціллю оптимізаційної задачі. Нові дерева додаються в ансамбль доти, доки помилка зменшується, або доки виконується одне з правил "ранньої зупинки".

Основними перевагами бустингу є:

1) простота реалізації. Бустинг має прості для розуміння та інтерпретації алгоритми, здатні вчитися на своїх помилках. Ці алгоритми вимагають попередньої обробки даних, і навіть мають вбудовані процедури обробки відсутніх значень. Крім того, більшість мов вбудовані бібліотеки для реалізації алгоритмів бустингу з безліччю параметрів, що дозволяють точно задавати продуктивність.

2) зменшення зміщення. Зміщення – це наявність невизначеності чи неточності у результатах машинного навчання. Алгоритми бустингу поєднують кілька слабких моделей у послідовний метод, що ітеративно покращує спостереження. Такий підхід допомагає зменшити високе усунення, яке поширене в моделях машинного навчання.

3) ефективність алгоритмів. Алгоритми бустингу фокусуються на елементах, що підвищують точність прогнозування під час навчання. Вони здатні зменшувати кількість атрибутів даних та ефективно обробляти великі набори.

У роботі була розглянута задача прогнозування характеристик комп'ютерної мережі. Для побудови моделі був обраний набір даних Optical Interconnection Network Data Set. Усі симуляції для формування набору даних виконані за допомогою програмного забезпечення під назвою OPNET Modeler. Набір даних Optical Interconnection Network Data Set містить 640 записів про використання мережі. Фрагмент набору даних показаних на рис. 1.

Node_Number	Thread_Number	T_R	Processor_Utilization	Channel_Waiting_Time	Input_Waiting_Time	Network_Response_Time	Channel_Utilization
640.000000	640.000000	640.000000	640.000000	640.000000	640.000000	640.000000	640.000000
40.000000	7.000000	0.550000	0.649013	377.459157	333.247102	1504.247529	26.347886
24.018772	2.237817	0.287453	0.194737	381.974899	233.721860	1202.606968	223.782214
16.000000	4.000000	0.100000	0.202377	0.950721	33.036130	0.529210	0.136979
16.000000	5.500000	0.300000	0.492530	29.247560	137.730986	580.676198	0.587539
40.000000	7.000000	0.550000	0.624787	265.614624	261.855556	1232.150369	0.773611
64.000000	8.500000	0.800000	0.833106	664.965408	485.943680	2115.326618	0.905573
64.000000	10.000000	1.000000	0.986516	1627.330246	892.852416	6065.736672	2895.323131

Рис. 1. Фрагмент використаного набору даних

Набір даних був розбитий на дві частини: тренувальний (512) та тестовий (128). Тренувальний набір використовувався для знаходження оптимальних параметрів моделі, а тестовий набір даних був створений для оцінки кінцевої моделі.

Після тренування моделей на тестовому наборі даних були отримані результати, які наведені у табл. 1.

Таблиця 1. Результати роботи моделей прогнозування

Метод	Score для навч.	Score для тест.
Випадковий ліс	1.000	0.645
Градiєнтний бустинг	1.000	0.619
Лiнійна Регресiя	0.764	0.012
Випадковий лiс CV	1.000	0.859
Градiєнтний бустинг CV	1.000	0.895

Значення метрики Score для моделей на основі випадкових лісів та градієнтного бустингу знаходиться на рівні 0.859 та 0.895 відповідно, що позитивно оцінює роботу моделей. Висока оцінка крос-валідації виправдовує використання цієї процедури. Таким чином, модель повністю задовольнила поставлені вимоги.

На основі проведеного аналізу було виявлено, що найкращі результати показала модель побудована на алгоритмах випадкового лісу.

ЛІТЕРАТУРА

1. Ho, Tin Kam. Random decision forests // Proceedings of 3rd International Conference on Document Analysis and Recognition. – 1995. – vol.1. – P. 278–282.
2. Breiman, Leo. Random Forests // Machine Learning: journal. – 2001. – Vol. 45, no. 1. – P. 5 – 32.
3. Breiman L. Out-of-bag estimation // Technical report, Statistics Department University of California, Berkeley. 1996. P. 1– 13.
4. Wolf B. J., Slate E. H., Hill E. G. Logic Forest: An ensemble classifier for discovering logical combinations of binary markers
5. Ruczinski I., Kooperberg C., LeBlanc M. Logic Regression methods and software / Eds: D. Denison et al. // Proceedings of the MSRI workshop on Nonlinear Estimation and Classification. New York: Springer, 2002. P. 333– 344.
6. Prinzie A., Poel D. Random Multiclass Classification: Generalizing Random Forests to Random MNL and Random NB // Working paper, Department of Marketing, Ghent University, 2007

ГОЛОВЧЕНКО Віталій Борисович – студент кафедри прикладної і теоретичної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, Україна, 61022; e-mail: xa11867788@student.karazin.ua, ORCID 0000-0002-0008-3999

Наукові інтереси:

- методи машинного навчання;
- методи штучного інтелекту.

СТРИЛЕЦЬ Вікторія Євгенівна – к.т.н.; доцент кафедри прикладної та теоретичної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, Україна, 61022; e-mail: viktorija.strilets@karazin.ua; ORCID: 0000-0002-2475-1496.

Наукові інтереси:

- методи машинного навчання;
- методи штучного інтелекту.

ДААС Т.І., ТКАЧУК М.В.

ПРО ОДИН ПІДХІД ДО РОЗРОБКИ БАНКІВСЬКИХ ІНФОРМАЦІЙНИХ СИСТЕМ: ЗНАННЯ-ОРІЄНТОВАНІ МОДЕЛІ ТА МІКРОСЕРВІСНІ АРХІТЕКТУРИ

Актуальність теми та мета дослідження

У сучасному світі зростає необхідність впровадження все більшої кількості електронних послуг у банківській сфері. Перелік таких послуг та їх складність зростає щоденно. Бізнес-вимоги користувачів вимагають від розробників забезпечення швидкого впровадження нового функціоналу банківські інформаційні системи (БІС), забезпечення мінімізації, а краще, відсутності впливу нових модулів на якість та логіку роботи вже впровадженого функціоналу, а також забезпечення відмовостійкості (robustness) при пікових навантаженнях на систему. Банківські установи, як власники продукту, також висувають свої вимоги до програмного забезпечення (ПЗ) ІС, серед яких можна виділити: можливість усунення помилок в роботі певного функціоналу у мінімальний строк з забезпеченням "недоторканності" інших функцій системи, мінімізація ресурсів при горизонтальному масштабуванні, впровадження вимог регулятора (для України – це НБУ) у відведені строки, мінімізація складнощів супроводу ПЗ ІС, можливість моніторингу робочого стану ІС, швидке усунення вразливостей (vulnerability), знайдених у системі чи сторонніх компонентах.

Більшість БІС побудовані досі з використанням монолітних архітектур (МА) через певні організаційно-технічні фактори. Але МА за багатьма показниками вже не встигають за динамікою розвитку бізнес-потреб користувачів та власників програмного продукту, тому це призвело до необхідності пошуку нових методів розробки та архітектур, які б як і МА забезпечували розподілену обробку даних, але б давали кращі можливості для виконання вимог, які висуваються до ПЗ БІС. Саме тому в ІТ-галузі широкої популярності набирають мікросервісні архітектури (МСА), які забезпечують гнучкість ІТ-інфраструктури, незалежність компонентів (сервісів) один від одного, можливість адаптивного масштабування та здатність до швидкої модернізації ПЗ. Саме тому метою цього дослідження є стислий аналіз проблем використання МСА для БІС та визначення можливих перспективних підходів до їх вирішення[1].

Модельні підходи до побудови банківських інформаційних систем на основі мікросервісних архітектур

БІС, які побудовані з використанням МА, як правило, починали розроблятися у минулому десятилітті, що відобразилося на підходах, які були використані при проектуванні ІС. Більшість таких систем розроблялися як інструмент для спрощення та автоматизації певних загально-використовуваних задач, таких як міжбанківська взаємодія (пересилання платежів), залишки на рахунку, тощо. Такі системи не мали різноманіття зовнішніх користувачів і зазвичай виконували лише «підрахунки». З точки зору модельного підходу вони були зовсім простими, а тому не потребували використання модельних методів при їх проектуванні та розробці. З розвитком електронних технологій розвивалися і БІС, які починали охоплювати все більш різноманітний функціонал, зокрема і той, що потребував взаємодію з великою кількістю кінцевих користувачів. Але оскільки майже усі існуючі БІС були побудовані з використанням МА, то кожен новий функціонал додавався у вигляді нового «шару» до вже існуючого, а отже не передбачав широке використання моделей, особливо доменних і знання-орієнтованих. Використовувалися зазвичай лише UML-діаграми, і лише для зображення особливостей функціоналу, який додається, але без урахування взаємодії компонентів системи в цілому. Так відбувалося, бо доменні моделі передбачали необхідність відносно частого рефакторінгу як архітектури ІС, так і програмного коду, що потребувало додатковий ресурс; а знання-орієнтовані моделі, для

впровадження якісних результатів в кінцевий продукт, потребували широкої аудиторії експертів, яких через новизну предметної області не вистачало. Ці фактори, а також залежність компонентів між собою у МА, призвели до необхідності пошуку такого типу архітектури, який би дозволяв, для підвищення якості ПЗ БІС та їх подальшого стрімкого розвитку, використовувати при проектуванні та розробці ще й науковий підхід: знання-орієнтовані та доменні моделі, експертні методи пошуку рішень (зокрема, метод аналізу ієрархій та метод Дельфі), методи об'єктно-орієнтованого аналізу та синтезу ПЗ з використанням мови UML (Unified Modeling Language), графічних нотацій сімейства IDEF та мови BPMML (Business Process Modeling Language).

МСА прийшла на заміну МА, яка була основним архітектурним патерном розробки усіх клієнт-серверних систем. МА характеризується чітко визначеною структурою застосунку, повною залежністю компонентів, єдиним стеком технологій та фреймворків та залежністю від платформ розгортання. У свою чергу, МСА є набором невеликих незалежно розміщених програмних сервісів, кожен з яких реалізує певний кінцевий функціонал, має свій власний процес виконання та взаємодія з яким виконується з використанням чітко визначеного протоколу обміну даними, але без залежності від мови програмування та технології, за якими він є реалізованим, платформи, на якій він розгорнутий та апаратних засобів, які він використовує [2-3].

МСА також є можливість застосування різних механізмів контейнеризації та оркестрації, що забезпечує можливість масштабування з мінімальними ресурсами та моніторингу робочого стану ІС. На сьогодні найбільш популярним засобом контейнеризації є Docker, а оркестрації – Kubernetes. Це дає можливість мінімізувати ресурси, які витрачаються на розгортання додаткових екземплярів мікросервісу, оскільки розгортається лише мікросервіс, на який підвищилося навантаження через зовнішні чинники, на відміну від МА, для якої в разі необхідності масштабування потрібно розгорнути сервери, які реалізують увесь функціонал, в тому числі і той, використання якого не збільшилося. Для БІС це особливо важливо, оскільки вони характеризуються досить нерівномірним використанням з боку клієнтів різного функціоналу ІС, а також є закономірність часу, у який відбуваються пікові навантаження на систему. Технології оркестрації дозволяють досить гнучко конфігурувати таке масштабування, а також, що досить важливо, утилізацію ресурсів при зниженні навантажень[4]. Використання МСА дозволяє значно ширше та якісніше використовувати модельні методи як при побудові нових систем, так і при додаванні нового функціоналу у вже існуючі. Оскільки МСА передбачає наявність купи невеликих мікросервісів, які відповідають кожен за свій функціонал, то при необхідності додавання нового мікросервісу з новим функціоналом, немає потреби перебудовувати усі раніше створені моделі та визначені онтології. В той же час, розробка та використання ПЗ, яке побудовано з використанням МСА, створює ряд проблем, до яких можна віднести: додаткові витрати на обмін даними між мікросервісами; необхідність забезпечувати коректну роботу мікросервісу у разі збою в іншому компоненті ІС; підтримка спадкоємності інтерфейсів, у разі розробки нового функціоналу; необхідність мати кваліфікований персонал для супроводу [2-3].

У теперішній час значна увага приділяється питанням удосконалення підходів розробки і оцінки якості ПЗ БІС, які проектуються та будуються з використанням МСА. Так, наприклад, в роботі [5], на основі аналізу останніх закордонних та вітчизняних публікацій, розглядаються такі показники якості розробки та впровадження МСА як: середній обсяг коду мікросервісу, ступінь зв'язаності як окремих мікросервісів, так і програмного коду всередині мікросервісу. Вони дозволяють кількісно оцінити якість архітектурного рішення, але залишає поза увагою зовнішні чинники, такі як інфраструктура, стек технологій, тощо. У [6] описується різниця між ПЗ, які побудовані на МА та МСА, а також приклад реалізації конкретної системи з визначенням основних показників якості. Але при цьому також практично не розглядається більш складні питання вдосконалення методологічних підходів до адаптивної розробки БІС.

Деякі недоліки та обмеження існуючих підходів

При побудові ПЗ БІС, з використанням МСА, особливу увагу потрібно звертати на

процес поділу функціоналу на мікросервіси, оскільки в таких ІС присутній службовий функціонал, який може використовуватися як один з етапів бізнес-логіки для великої частини кінцевого функціоналу, який виділений кожний в окремий мікросервіс. Це створює ситуацію, коли неможливо повністю реалізувати одну з вимог МСА – «незалежність»[3]. Прикладом такого функціоналу є робота з електронним підписом (ЕП) чи ОTR-кодами, які використовуються як підтвердження вводу платіжних документів чи неплатіжних операцій. Реалізувати його можна двома способами: в кожному мікросервісі, який обробляє платежі/операції або в окремому мікросервісі, до якого за потреби звертатимуться інші мікросервіси. Перший варіант призводить до дублювання логіки, і таке дублювання зростає зі зростанням кількості мікросервісів, а другий варіант підвищує залежність мікросервісів між собою. Тому вибір варіанту для реалізації повинен залежити від функціональних особливостей ІС, та виконується лише після розрахування кількісних та якісних показників, до яких можна віднести: кількість строк коду, незмінність функціоналу, швидкодія, тощо [7].

Перспективні напрямки дослідження проблем БІС на основі МСА

Враховуючи наведені вище обмеження та особливості банківських систем та недоліки у підходах до побудови цих систем з використанням МСА, можливо запропонувати наступні напрямки досліджень:

1. Розробка технології побудови БІС на основі МСА, яка відрізняється від існуючих послідовним застосуванням при проектуванні та розробці ПЗ:
 - a. знання-орієнтованих моделей та методів: NLP (Natural Language Processing) – методів, методів побудови та обробки онтологій, лінгвістичних моделей та ін.;
 - b. експертних методів пошуку рішень: метод аналізу ієрархій, метод Делфі, методи логічного висновку на основі аналізу прецедентів CASE-based Reasoning) та ін.;
 - c. методів об'єктно-орієнтованого аналізу та синтезу ПЗ використанням уніфікованої мови моделювання UML, нотацій сімейства IDEFX та мови BPMML для формалізації процедур проектування ПЗ БІС
 - d. методів статистичного аналізу експериментальних даних для кількісної оцінки ефективності запропонованого підходу
2. Удосконалення проектних патернів для створення ПЗ, які мають забезпечити кращі показники якості за рахунок використання мікросервісної архітектури у різноманітних БІС порівняно з монолітною архітектурою;
3. Подальший розвиток методологій адаптації життєвого циклу процесу розробки (SDLC) для БІС за умов постійних змін у середовищі їх функціонування.

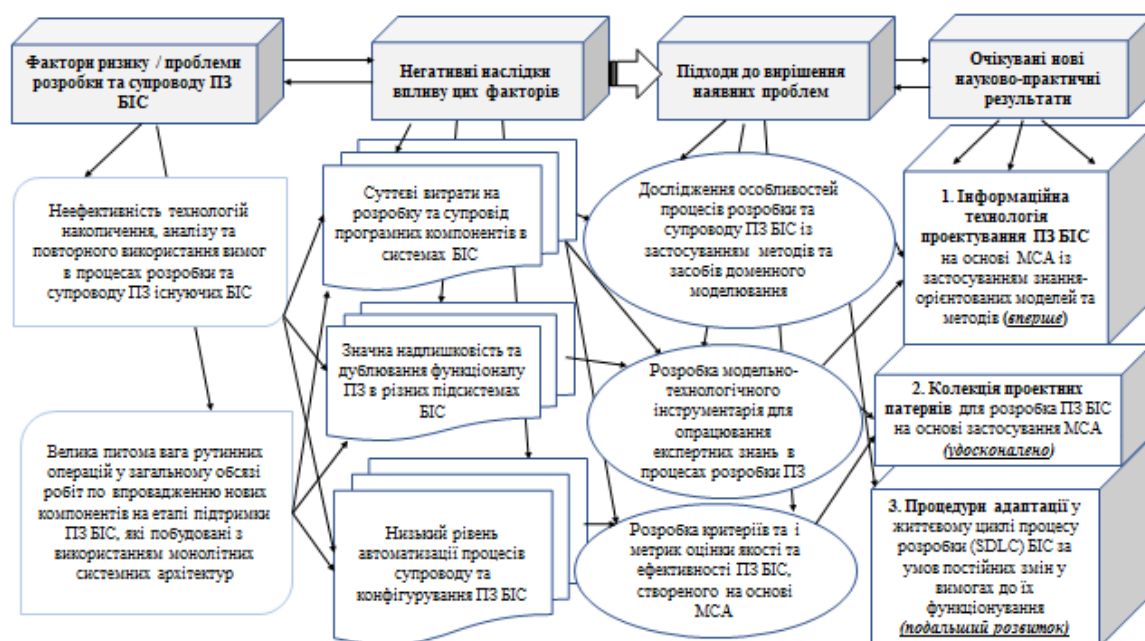


Рисунок 1 - Концептуальна схема дисертаційного дослідження

Концептуальна схема виконання запланованого дисертаційного дослідження за п.п. (1)-(3) представлена на рис.1. Практична реалізація запропонованого підходу також передбачає розробку інтелектуального інструментального CASE-засобу для проведення обчислювальних експериментів та аналізу їх результатів.

Висновки

У представлений роботі обґрунтована актуальність вирішення проблеми побудови, розробки та супроводу ПЗ банківських інформаційних систем (БІС) з використанням мікросервісної архітектури (МСА). На основі огляду сучасних досліджень щодо побудови БІС та використання МСА, зроблено мотивований висновок щодо необхідності пошуку нових підходів до комплексного вирішення цих проблем, зокрема, з використанням знання-орієнтованих моделей і технологій. В результаті запропонована концептуальна схема основних етапів і завдань дисертаційного дослідження, виконання якого має на меті отримати нові науково-практичні результати у цій галузі.

ЛІТЕРАТУРА

1. Avornicului Mihai-Constantin, Bresfelean Vasile Paul, MODEL DRIVEN DEVELOPMENT OF ONLINE BANKING SYSTEMS, 07.2011, URL: [OF_ONLINE_BANKING_SYSTEMS/fulltext/024ea98d0cf257003d2bc17c/MODEL-DRIVEN-DEVELOPMENT-OF-ONLINE-BANKING-SYSTEMS.pdf](https://vps.ua/blog/docker-and-linux-containers/) (Останній перегляд 11.11.2022)
2. Фаулер М. Архитектура корпоративных программных приложений – М., Изд. дом Вильямс, 2006. – 544 с.
3. Ньюмен С. Создание микросервисов. – СПб.: Питер, 2016. – 304 с.
4. Docker і технологія контейнерів Linux [Electronic resource]. URL: <https://vps.ua/blog/docker-and-linux-containers/> (Останній перегляд: 12.11.2022).
5. Зінов'єв Д.В., Ткачук М.В., Тріщенко І.В. Моделі та технології забезпечення якості сервіс-орієнтованих програмних систем: сучасний стан та перспективні напрямки досліджень // Матеріали міжн. науков-техн. конференції КМНТ-2021, (м. Харків, 23-25 квітня 2021 року) – Х.: ХНУ імені В.Н. Каразіна, 2021. – С. 166-169.
6. Michel Cojocaru;Alexandru Uta;Ana-Maria Oprescu. MicroValid: A Validation Framework for Automatically Decomposed Microservices. IEEE International Conference on Cloud Computing Technology and Science (CloudCom - 2019), 2019. - pp. 78 – 86
7. Julia Korsun, 5 Innovations in Banking Application Development: From Challenges to Opportunities, 28.06.2019, URL: <https://djangostars.com/blog/banking-application-development> (Останній перегляд: 12.11.2022)

ДААС Тимур Імад Ахмад – аспірант кафедри моделювання систем і технологій факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, м. Харків, Майдан Свободи 6, 61022; e-mail: timurkadaas@gmail.com, ORCID: <https://orcid.org/0000-0003-2796-8145>

Наукові інтереси:

– об'єктно-орієнтоване програмування, архітектура програмного забезпечення, розробка високонавантажених систем, розробка Web-застосунків, захист Web-застосунків, криптографія

ТКАЧУК Микола Вячеславович – д.т.н., професор, завідувач кафедри моделювання систем і технологій факультету комп'ютерних наук, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022; mykola.tkachuk@karazin.ua, ORCID: <https://orcid.org/0000-0003-0852-1081>

Наукові інтереси:

– концептуальне моделювання інформаційних систем і технологій, проєктування проблемно-орієнтованих програмних систем, знання-орієнтовані моделі і методи розробки та інженерії якості програмного забезпечення.

ДЕЙНЕГА Т.С., СВАТОВСЬКИЙ І.І.

ДОСЛІДЖЕННЯ ЗАСТОСУВАННЯ АЛГОРИТМІВ ШТУЧНОГО ІНТЕЛЕКТУ В СИСТЕМАХ ВИЯВЛЕННЯ/ЗАПОБІГАННЯ ВТОРГНЕНЬ

Вступ

Сучасні кібератаки спрямовані на несанкціонований доступ, шантаж, підміну та інші види зловмисних дій над інформацією, які зосереджені на порушенні певних механізмів безпечної обробки інформації. Для виявлення кібератак зазвичай використовуються статистичні методи, методи, які застосовують сигнатури, та методи знаходження аномалій. Методи знаходження аномалій, у порівнянні зі статистичними та сигнатурними методами, дають можливість виявляти нові загрози та атаки «нульового дня», що дозволяє ефективніше протистояти сучасним кібератакам.

На практиці вказані методи реалізуються в системах виявлення та запобігання вторгненням (Intrusion Detection and Prevention Systems – IDS/IPS) [1,4]. Ці методи застосовуються для ефективного дослідження та типізації широкого спектру даних, які передаються в мережах.

Сьогодні швидко розвиваються та застосовуються різноманітні технології штучного інтелекту (ШІ). Їх основним завданням є виявлення та захист від складних мережових атак. Однією з найдинамічніших технологій ШІ вважають технології штучних нейронних мереж [1,3]. Вони дозволяють вирішувати велику кількість практичних технічних завдань.

Дослідження технологій нейронних мереж

Нейронна мережа імітує людський мозок, створюючи систему обробки інформації, яка складається з великої кількості взаємопов'язаних вузлів (нейронів), які працюють один з одним, щоб вирішити конкретне завдання [1]. Рішення про вихід кожного вузла зважується і обробляється для подачі на вхід для всіх інших вузлів в наступному шарі. Процес навчання, генерує архітектурою нейронної мережі дозволяє покривати дуже складні і нелінійні відносини між даними [5]. На рис. 1 показана архітектура нейронної мережі.

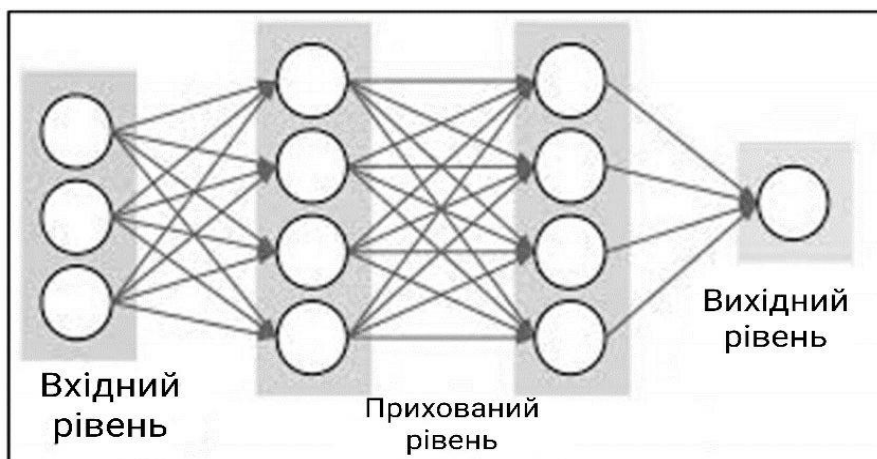


Рис. 1 Архітектура нейронної мережі

Нейронні мережі можуть отримувати закономірності, виявляти відносини в даних і вчитися на зібраних знаннях. Для отримання поставлених завдань нейронна мережа повинна бути правильно спроектована. Головна перевага нейронної мережі - здатність робити висновки із зібраних даних без будь-яких попередніх знань про закономірності цих даних, а також її здатність узагальнювати вивчені дані, які дозволяють нейронній мережі виявляти і

класифікувати невідомі атаки і різні типи відомих атак. Крім того, як тільки атака розпізнана, нейронна мережа не може бути схильна до цієї атаки у майбутньому. Висока швидкість обробки даних також є однією з переваг нейронної мережі. Щоб успішно впровадити цей метод для системи виявлення та запобігання вторгнень, дані, що представляють нормальну і ненормальну поведінку мережі, повинні бути введені в неї, щоб вона могла автоматично налаштовувати мережеві коефіцієнти на етапі навчання [2, 3].

За останні кілька років було проведено багато досліджень застосування нейронних мереж для виявлення вторгнень [4]. Нейронна мережа зворотного поширення використовується для виявлення вторгнення в систему через її здатність швидкого розпізнавання шаблонних пакетів, отриманих з мережі. Особливості багатьох атак були витягнуті й проаналізовані з стандартних і незвичайних пакетів. Проаналізовані результати цих пакетів були використані для навчання нейронної мережі за зразком пакетів обох типів, використовуючи стандартний алгоритм зворотного поширення помилки.

Щоб подолати низький рівень виявлення, високу кількість помилкових спрацьовувань і інших дефектів, був розроблений новий алгоритм виявлення вторгнень на основі нечіткої нейронної мережі. Цей алгоритм класифікує об'єкти і розпізнає їх нормальну і ненормальну поведінку. В ході великої кількості експериментів продемонстровано, що запропонована модель ефективна і має краще узагальнення. Крім того, швидкість правильного виявлення вторгнень збільшується, а ймовірність помилкового виявлення знижується [1].

Практична реалізація технології нейронної мережі на мові програмування Java

Головною метою цієї роботи була реалізація алгоритму нейронної мережі на мові програмування Java для класифікації мережевого трафіку у дослідній локальній мережі як загрозового, безпечного чи трафіку, що потребує додаткової перевірки людиною (середній рівень загрози між безпечним та загрозовим). Дослідження пропонує створення алгоритму підготовки тестової вибірки даних для нейронної мережі, алгоритму навчання мережі та застосування навченої нейронної мережі на реальних даних з розрахунком відсотку правильних відповідей та кількості помилок у залежності від ітерації на тестовій вибірці даних.

Алгоритм підготовки тестової вибірки для нейронної мережі, заснований на алгоритмах нечіткої логіки (Fuzzy Logic), приймає на вхід текстовий файл з пакетною інформацією, яка була перехоплена у локальній мережі за допомогою sniffера пакетів WireShark та видає на виході тестову вибірку для нейронної мережі у вигляді пар. Перше значення пари – це масив вхідних даних про мережевий пакет, його розмір = 3. До цих даних відноситься: розмір пакету в кілобайтах, наявність у пакеті скриптів та наявність використання мережевих утиліт у пакеті. Друге значення пари – правильна відповідь щодо загрози пакету.

Після того, як навчальна вибірка буде зібрана, запускається алгоритм навчання нейронної мережі. Даний алгоритм заснований на трьох частинах: пряме розповсюдження помилки (forward propagation), зворотного поширення помилки (backward propagation) та оновлення вагових коефіцієнтів зв'язку між нейронами [5]. Шляхом практичних експериментів було знайдено наступні гіпер-параметри для нейронної мережі (Табл. 1):

Табл. 1 Гіпер-параметри дослідної нейронної мережі

Назва параметру	Значення
Кількість шарів нейронної мережі	3
Кількість нейронів на другому шарі	10
Кількість ітерацій по навчальній вибірці	1000

Схематично отриману нейронну мережу можна зобразити наступним чином (рис. 2):

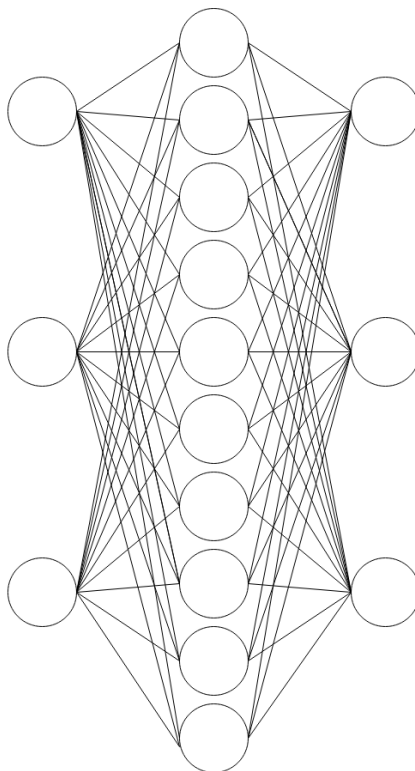


Рис. 2 Схема нейронної мережі

Після завершення етапу навчання у консолі середовища розробки IntelliJ IDEA Community Edition 2021.3 можна бачити відсоток правильних відповідей нейронної мережі на останньому циклі проходи по навчальній вибірці (рис. 3):

```
[0.000000, 0.000000, 0.041016]
0
[0.000000, 0.000000, 0.041016]
0
[0.000000, 0.000000, 0.044922]
0
[0.000000, 0.000000, 0.044922]
0
[0.000000, 0.000000, 0.041016]
0
[0.000000, 0.000000, 0.041016]
0
[0.000000, 0.000000, 0.041016]
0
[0.000000, 0.000000, 0.041016]
0
[0.000000, 0.000000, 0.072266]
0
Accuracy: 0.9844357976653697
```

Рис. 3 Відсоток правильних відповідей нейронної мережі на навчальній вибірці

Як можна бачити з рисунку 3 - відсоток правильних відповідей складає 98.44%, що оцінює мережу як дуже ефективну. Також після виконання алгоритму навчання можна побачити графік залежності помилки відповіді від ітерації:

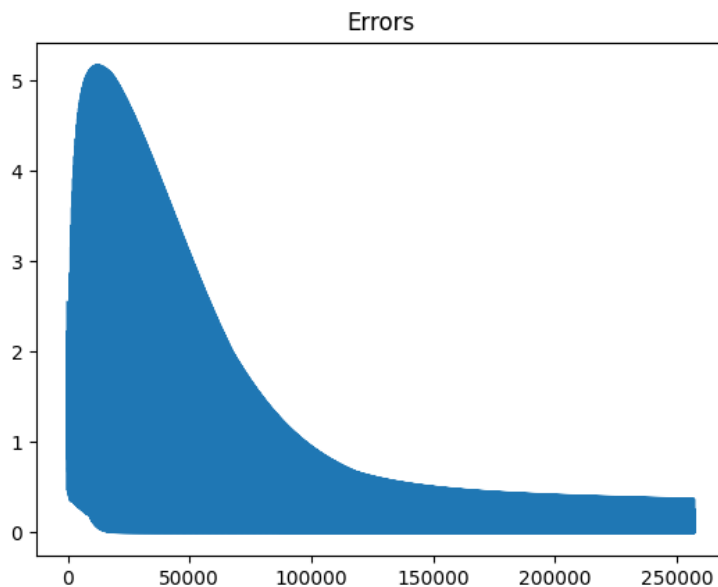


Рис. 4 Графік залежності помилки відповіді від ітерації

Як можна бачити з графіку на рисунку 4 - кількість помилок зменшується зі збільшенням числа пройдених ітерацій. Стрімке зниження кількості помилок спостерігається після 20000 ітерацій алгоритму.

Після того, як нейронна мережа була успішно навчена у консоль виводяться матриці вагових коефіцієнтів W1 та W2. Їх значення наступні:

- W1 = {0.155353, 0.848620, -0.058468, 0.649779, 0.413327, 0.594318, 0.085725, 0.792108, -0.043930, 0.496328; -0.154272, 0.449024, 0.231025, -0.216459, 1.486742, 0.779725, 1.256136, 0.576876, 0.852188, 0.200115; 0.577778, 0.420876, 0.438282, 0.021567, 1.871525, 0.422121, 1.624109, 0.890067, 0.570674, 0.683717}
- W2 = {0.653276, 0.313120, 0.187101; 0.773230, 0.899240, 0.504340; 0.777735, 0.305611, 0.144749; 1.484853, 0.283755, 0.228076; -0.820227, 0.304082, 1.926313; 0.207484, 0.740289, 0.990918; -0.120909, 0.094986, 1.811141; 0.019589, 0.410993, 1.024896; 0.193575, -0.076539, 1.092565; 0.977203, 0.157549, 0.870955}

Коли система навчена та має правильні вагові коефіцієнти можна протестувати її на реальних даних (мережових пакетах). У якості реальних пакетів було зібрано інформацію з 20 пакетів мережевого трафіку, у яких містяться і загрозові пакети даних, і безпечні, і ті, які потребують додаткової перевірки адміністратором мережі.

```
Predication: Need clarify - Correct
Predication: Need clarify - Correct
Predication: Need clarify - Correct
Predication: Not secure - Incorrect
Predication: Secure - Incorrect
Predication: Secure - Correct
Predication: Secure - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Predication: Secure - Correct
Predication: Secure - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Predication: Need clarify - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Predication: Not secure - Correct
Real accuracy: 0.9
```

Рис. 5 Відсоток правильних відповідей нейронної мережі на реальній вибірці

Нейронна мережа надала правильної відповіді для 90% мережевих пакетів (рис. 5). Даний показник трохи відрізняється від того, що був отриманий на тестовій вибірці через те, що у якості реальних даних були обрані також пакети даних, що містять у собі більше одного параметра, який вказує на потенційну загрозу у пакеті на відміну від тестової вибірки, яка містить у собі переважно пакети з одним таким параметром.

Висновки

Таким чином, досліджену та практично реалізовану технологію нейронних мереж доцільно використовувати для класифікації мережевого трафіку в сучасних мережах зі складними потенційними загрозами. Можна з упевненістю сказати, що запропонована модель нейронної мережі є достатньо ефективною у виявленні та запобіганні мережевих атак. Дана модель має переваги у швидкодії та надійності одержання правильних висновків детектування загроз над традиційними експертними системами та системами, заснованими на інших технологіях штучного інтелекту. На реальних мережевих даних нейронна мережа показала результат близько 90% правильних відповідей та 98.44% - на тестових вибірках. В подальшому розглянута модель системи виявлення/запобігання вторгнень може буде вдосконалена шляхом додавання до навчальної вибірки більш складних випадків, наприклад, пакетів з декількома параметрами, які вказують на можливі загрози, що може забезпечити кращий результат на реальних даних.

ЛІТЕРАТУРА

1. Ganesh Kumar P., Devaraj D. Intrusion detection using artificial neural network with reduced input features. URL: http://ictactjournals.in/paper/IJSC_V1_I1_PAPER_5_30_36.pdf. Coimbatore, 2010. 36 p.
2. Kumar G., Kumar K., Sachdeva M. The use of artificial intelligence-based techniques for intrusion detection. Punjab: Springer Science + Business, 2010. 387 p.
3. Gupta B., Singh B., Jain V. Artificial Intrusion Detection Techniques. URL: <https://arxiv.org/ftp/arxiv/papers/1701/1701.02145.pdf>. New Delhi, 2014. 43 p.
4. Sikos L. F. AI in Cybersecurity. New York: Springer, 2018. 205 p.
5. Shahriar Usman Khan, Fariha Eusufzai, Md. Azharuddin Redwan, Mohiuddin Ahmed, Saifur Rahman Sabuj. Artificial Intelligence for Cyber Security: Performance Analysis of Network Intrusion Detection. Cham: Springer, 2022. 140 p.

ДЕЙНЕГА Тарас Сергійович – студент кафедри безпеки інформаційних систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: deynega.taras@ukr.net; ORCID: 0000-0002-7659-9721.

Наукові інтереси:

– *штучний інтелект, об'єктно-орієнтоване програмування, алгоритмізація.*

СВАТОВСЬКИЙ Ігор Іванович – доцент кафедри безпеки інформаційних систем і технологій, кандидат технічних наук, старший науковий співробітник; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: i.svatowsky@karazin.ua; ORCID: 0000-0002-1836-5599.

Наукові інтереси:

– *завадостійке кодування, завадостійкі інформаційно-телекомунікаційні системи, методи підвищення захищеності передачі цифрової інформації по радіоканалах.*

ДЕРЕВ'ЯНКО Я.А., ГОРБЕНКО І.Д.

ГЕНЕРАЦІЯ НЕЛІНІЙНИХ ВУЗЛІВ ЗАМІН З ВИКОРИСТАННЯМ ФАКТОРІАЛЬНОЇ СИСТЕМИ ЧИСЛЕННЯ

Вступ

Перестановки є одним із важливих елементів комбінаторики, який повсюдно використовується у різних сферах і реальних задачах. Одним із варіантів використання перестановок у криптографії є шифрування повідомлень або використання перестановок в різних криптографічних перетвореннях. В деяких випадках, необхідно мати змогу швидко отримати номер перестановки за її значеннями, або навпаки значення перестановки за її номером у загальному просторі. З такою метою доцільним буде використання факторіальної системи числення, оскільки завдяки її використанню можна швидко відтворити перестановку, маючи тільки її номер, а також виконати зворотнє перетворення в разі необхідності.

В даній роботі описано загальні принципи використання факторіальної системи числення для роботи з перестановками і продемонстровано роботу алгоритмів для застосування ФСЧ. Зазначені алгоритми реалізовано на мові програмування С. Ми наводимо чисельні приклади для демонстрації та перевірки отриманих результатів. Також в роботі досліджено зміну криптографічних показників простору блоків, що може допомогти в подальшому краще розуміти простір блоків та оптимізувати пошукові алгоритми.

Ключові слова: підстанова, факторіальна система числення, S-box, криптографічні показники, залишки.

Факторіальна систем числення

У факторіальній системі числення основами є послідовність факторіалів $b_k = k!$. Кожне натуральне число x представляється у вигляді:

$$x = \sum_{k=1}^n a_k k!, \quad (1)$$

де $0 \leq a_k \leq k$

Факторіальна система числення використовується при декодуванні перестановок списками інверсій, тобто маючи номер перестановки, можна відтворити її саму наступним чином: номер перестановки (починаючи з 0) записується в факторіальній системі числення, при цьому коефіцієнт при числі $i!$ буде позначати число інверсій для елементу $i+1$ в тій множині, в якій виконуються перестановки (тобто число елементів, менших за $i+1$, але які знаходяться правіше нього в шуканій перестановці).

Алгоритми представлення цілого числа в факторіальній системі числення і у вигляді перестановки

Для представлення цілого числа без знаку (натуральні числа) x у факторіальній системі числення потрібно знайти набір цифр a_k , $k=1, \dots, n$ таких, що:

$$x = \sum_{k=1}^n a_k k! = a_n n! + a_{n-1} (n-1)! + \dots + a_2 2! + a_1 1! \quad (2)$$

де $0 \leq a_k \leq k$

Для швидшого та більш ефективного розрахунку можна скористатися схемою Горенра:

$$x = a_n n! + a_{n-1} (n-1)! + \dots + a_2 2! + a_1 1! = ((a_n n + a_{n-1})(n-1) + \dots + a_2) 2 + a_1 \quad (3)$$

При такому записі числа x для знаходження всіх a_k , $k=1, \dots, n-1$ достатньо виконати послідовність таких кроків:

$$\begin{aligned} x &= q_1 2 + a_1; \\ q_1 &= q_2 3 + a_2; \\ q_2 &= q_3 4 + a_3; \\ &\dots \\ q_{n-3} &= q_{n-2} (n-1) + a_{n-2}; \\ q_{n-2} &= q_{n-1} n + a_{n-1}; \\ q_{n-1} &= a_n. \end{aligned} \quad (4)$$

Набір залишків $a_n, a_{n-1}, \dots, a_2, a_1$ представляє ціле число без знаку в факторіальній системі числення.

Приклад 2. Представимо число 77 в факторіальній системі числення (з використанням алгоритму 2):

$$\begin{aligned} 77 &= 38 \cdot 2 + 1; \quad a_1 = 1; \\ 38 &= 12 \cdot 3 + 2; \quad a_2 = 2; \\ 12 &= 3 \cdot 4 + 0; \quad a_3 = 0; \\ 3 &= 3; \quad a_4 = 3. \end{aligned} \quad (5)$$

Вихід: число у факторіальній системі має вигляд $\{3,0,2,1,0\}$.

Число у факторіальній системі числення буде означати наступне для перестановки довжини 5:

- першим залишком факторіального числа, буде число елементів, які є меншими за 5, але стоять правіше нього – 3;
- другим залишком буде число елементів менших за 4, але які стоять правіше нього – 0;
- третім залишком є число елементів менших за 3, які стоять правіше нього – 2;
- четвертим залишком буде число елементів, менших за 2, які стоять правіше нього – 1.
- останній елемент перестановки завжди ставиться на вільне місце, тому значення останнього залишку завжди буде нульовим.

Виходячи з перерахованих вище правил переведення факторіального числа у перестановку, отримуємо наступну перестановку для числа 77 ($\{3,0,2,1,0\}$ у ФСЧ) – (3,5,2,1,4).

Детальний опис алгоритмів, а також експериментів можна знайти у [1].

Виконання перетворень показано на рисунку 1:

```

START NUMBER
77

FACTORIAL NUMBER
3, 0, 2, 1, 0,

SUBSTITUTION
3, 5, 2, 1, 4,

FACTORIAL NUMBER
3, 0, 2, 1, 0,

FINAL NUMBER
77

```

Рис. 1 Переведення числа 77 у перестановку і назад

Експерименти

Для демонстрації та перевірки правильності застосування розроблених алгоритмів проведено низку експериментів із представлення нелінійних підстановок відомих стандартів симетричного криптоперетворення. Детальні експерименти надаються у [1]. Ми наводимо приклади переходу від перестановок до факторіальних чисел і від них, до чисел у десятковій системі числення. Також ми наводимо заміри швидкості виконання операцій такого переходу.

Нами проведено заміри швидкодії операцій перетворення перестановки у номер, а також зворотного перетворення. Для цього було проведено тестування, шляхом виконання операцій над одним і тим же блоком 1000000 разів, щоб порахувати середню швидкість виконання операції в мілісекундах, оскільки ця швидкість має дуже маленьке значення і не відображається коректно шляхом меншої кількості тестів. Код усіх методів перетворення і заміру швидкостей наведено у [2-3]. Результати надано в таблиці 1:

Табл.1 Результати вимірювання швидкодії операції в мілісекундах

Перехід від перестановки до номеру	2.86940 мс
Перехід від номеру до перестановки	0,60700 мс

Використання пам'яті при виконанні алгоритмів в залежності від кількості блоків, задіяних для пошуку відрізнявся від 10,5 Мб для 10000 блоків до 101 Мб для 100000 блоків. Отримані результати свідчать про те, що виконання всіх необхідних операцій є доволі швидким і при цьому не сильно навантажує оперативну пам'ять комп'ютера.

Далі показано роботу з нелінійними вузлами заміни на прикладі AES блоку:

Табл.2 Представлення нелінійного вузла заміни алгоритму AES у вигляді номеру у одновимірному просторі

AES-SBox	43880208007746150693148788658785927962956643589594057609124938966311869 06385111763357572778332571141682146209499454052643751923036183936623352 77010380153387264179445016571705252302920569441228045303327567040823942 74612393094935659378418967270259726890339492473242137991953517032463783 60648066895595033278711020844867589240650723513074058254311572860596324 97461568460660658015651849289360349395308511707577545803238296080881647 26585567607716607901471646633359943117599023937932382874636793320251875 7906752833
-----------------	---

Для блоку AES ми отримали номерне представлення у одновимірній множині усіх блоків і тепер можемо робити будь-які маніпуляції з цими номерами для пошуку, наприклад, нових блоків з хорошими показниками, але відмінними від тих, що використовуються в стандарті. Далі буде показано можливе практичне застосування цього.

Дослідження простору нелінійних вузлів заміни поряд з AES-блоком

Далі на рисунках будуть надані графіки параметрів блоків, що знаходяться біля блоку AES: нелінійність, лінійну збитковість, оскільки алгебраїчний імунітет на дельта-рівномірність таких блоків будуть майже однаковими.

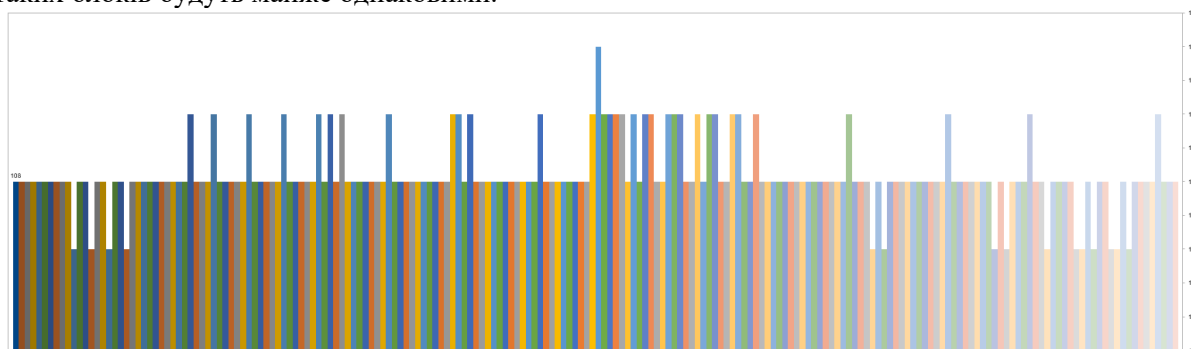


Рис. 2 Нелінійність блоків навколо AES

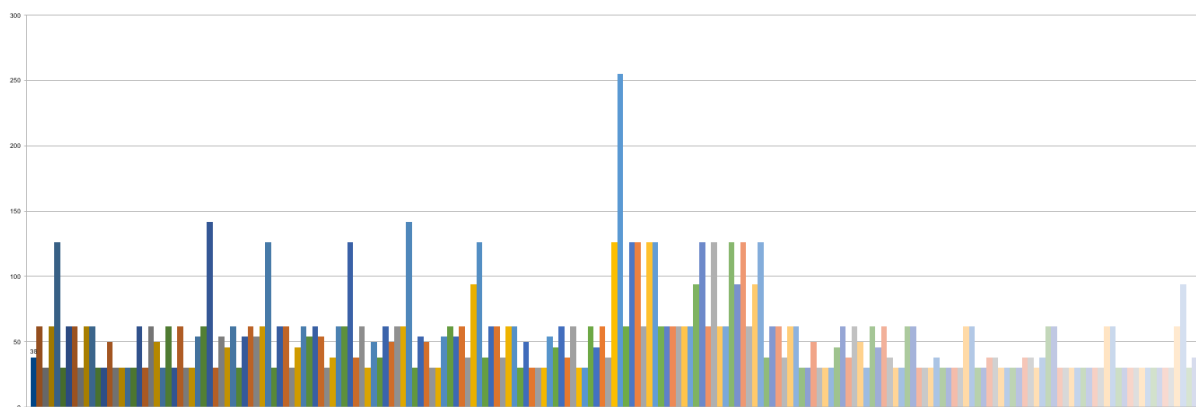


Рис. 3 Лінійна збитковість блоків навколо AES

Я видно з рисунку, більшість блоків по сусідству з AES S-Box мають нелінійність 108, менша кількість 106 або 110 і тільки сам AES S-Box має нелінійність 112, а лінійна збитковість більшості блоків не перевищує значення 100, інших 150, і тільки сам AES S-Box має найбільшу лінійну збитковість 254 – що є дуже поганим показником.

Висновки

Оскільки робота з перестановками є важливою задачею в багатьох сферах сучасної діяльності, застосування факторіальної системи числення з метою доступу до певної перестановки за її номером або навпаки отримання номеру за її значеннями є доволі важливою задачею. В даній було розглянуто та пояснено загальні принципи використання факторіальної системи числення для роботи з перестановками. Також у цьому дослідженні було проведено дослідження простору блоків поруч з AES-Sbox для виявлення показників блоків, що знаходяться поруч із ним. Це може допомогти в подальшому краще розуміти простір блоків та оптимізувати пошуки як цим, так і іншими алгоритмами пошуку.

ЛІТЕРАТУРА

1. Я.А. ДЕРЕВ'ЯНКО, Ю.І. ГОРБЕНКО, О.О.КУЗНЕЦОВ. ФАКТОРІАЛЬНА СИСТЕМА ЧИСЛЕННЯ ДЛЯ ГЕНЕРАЦІЇ НЕЛІНІЙНИХ ПІДСТАНОВОК. 2022. URL: <https://doi.org/10.30837/rt.2022.2.209.04>
2. Програмна реалізація Факторіальної системи числення від Дерев'янка Я. А. URL: <https://github.com/DereviankoYaroslav/SBoxDereviankoFactorial>.
3. Програмна реалізація алгоритму PSO з використанням факторіальної системи числення від Дерев'янка Я. А. URL: <https://github.com/DereviankoYaroslav/FactorialPSO-NL>.

ДЕРЕВ'ЯНКО Ярослав Андрійович – студент кафедри безпеки інформаційних систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: yarik0009258@gmail.com; ORCID: <https://orcid.org/0000-0002-3290-3373>

Наукові інтереси:

– електронні підписи, нелінійні підстановки, постквантова криптографія, симетрична криптографія, постквантові алгоритми.

ГОРБЕНКО Іван Дмитрович – доктор технічних наук, професор, головний конструктор АТ «Інститут інформаційних технологій», професор кафедри безпеки інформаційних систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: GorbenkoI@iit.kharkov.ua; ORCID: <https://orcid.org/0000-0003-4616-3449>

Наукові інтереси:

– електронні підписи, постквантова криптографія, нелінійні підстановки, симетрична криптографія, асиметрична криптографія.

ДОЛЯ Г.М., БОНДАРЕНКО К. О.

SIMULATION MODEL OF THE MULTIMODE LASER RADIATION TRANSFORMATION ON A GLASS BEADS RETROREFLECTIVE SURFACE

The article considers a model of transformation of multimode laser radiation on a retroreflective surface consisting of a collection of miniature glass beads light-reflectors in the context of laser fire alarm systems based on the analysis of the turbulence state, which are triggered when changing the parameters of detected laser emission. Descriptions of the physical processes that occur during transformation and the mathematical simulation model are given, along with the intermediate results of the simulation. The interrelation of multimode, dispersion and contrast of the resulting speckle pattern is considered. An analysis is made of the influence of the multimode nature of a laser radiation source on the output amplitude-phase distribution. Application of the core idea of the research to other developed models and schemas of the various laser-based detection systems is discussed. The prospects for further development of the study are presented.

Keywords: *laser, retroreflective surface, glass beads, simulation model, specklometry.*

Introduction

Recently, a new class of light retroreflecting surfaces (RRS), consisting of a collection of miniature light-reflectors, has appeared for practical applications. Their essential difference from rough surfaces or scattering particles in the flow is the patterns of spatial distribution of scattered radiation [1]. Taking into account these differences, a number of new effective devices can be built [2, 3], such as, for example, sensors of vibrations or acoustic oscillations, as well as meters of linear or angular speed of movement. Among others, there are also laser fire detectors based on the analysis of the state of turbulence, which are triggered when certain parameters of laser radiation are changed. Previously, in work [4], a study was conducted that shows the possibility of creating such sensors using monochromatic radiation and corner reflectors. Of interest is to study the features of such detection systems consisting of a multimode laser emitter and a light-reflecting surface made of glass beads. As part of the research of this type of device, it is beneficial to develop a simulation model of the system and apply it to analyze the influence of the variation of the system parameters on the quality of detection. In the context of work on this model, it is necessary to develop its part that describes the transformation of multimode laser radiation on RSS, taking into account parameters that are relevant for the subject area, which is the focus of this work.

Model's scope definition

In the developed model, the RRS is assumed to be two parallelepipeds with a common side located in space. They represent the covering layer and the substrate. The covering layer is assumed to be a dielectric, and the substrate can be either a dielectric or a metal. Beads are enclosed between the parallelepipeds. In this model they are considered ideal spheres made of dielectric, with a given radius, as well as a center for each of them. The emitter is presented as a Gaussian beam of a multimode polarized laser, the location of which is set arbitrarily in space. The reflector and emitter are assumed to be placed in some dielectric isotropic medium. Model parameters: spatial arrangement of the RRS, emitter, radiation wavelength, mode order, polarization of the laser beam, its half-width, refractive indices (functions of the radiation frequency) for the environment, cover layer, substrate, and beads, thickness of the cover layer, beads' immersion depth, their radius, type and step of the grid along which beads are located, their offset relative to the grid nodes, sampling frequency. The result of the simulation is the intensity distribution of the radiation reflected from the RSS in the far zone. Of particular interest is the analysis of the influence of the multimodality of the emitter on the form of the detected distribution.

Description of the simulation model

Modeling takes place in three-dimensional space, iteratively for each of the discrete points on the surface of the RRS and is divided into three main stages. At the first stage, a discrete point is selected on the RRS plane. It is converted into a point in space, for which the parameters of the emitter field to be calculated (for more details refer to [5]). Earlier in the works [4, 5], systems were

considered in which a single-mode monochromatic laser was assumed as a radiation source. In practice, lasers in the form-factor required for the sensor are multimode. To describe the field at a point in this case, one should use multimode Gaussian beams [6]

$$E_{l,m}(x, y, z) = E_0 \frac{w_0}{w(z)} H_l \left(\frac{\sqrt{2}x}{w(z)} \right) H_m \left(\frac{\sqrt{2}y}{w(z)} \right) \times \exp \left(-\frac{x^2 + y^2}{w^2(z)} \right) \exp \left(-i \frac{k(x^2 + y^2)}{2R(z)} \right) \exp(i\psi(z)) \exp(-ikz) \quad (1)$$

where l, m - longitudinal mode and transverse mode respectively,

H - Hermite polynomial.

The cross section of the intensity distribution across the beam propagation direction then looks as shown in the figure 1 (a), with TEM LM designating modes: L designating the longitudinal mode and M the transverse mode.

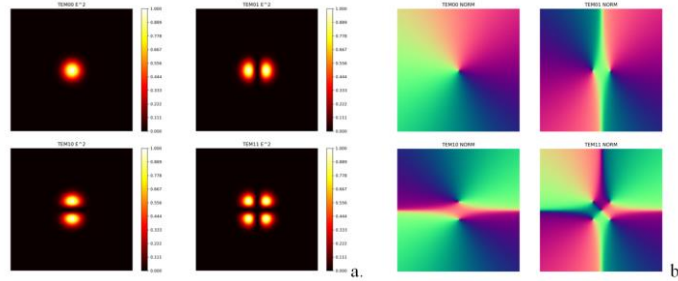


Fig. 1 Intensity distribution (a) and normal map (b) of the multimode Gaussian beam cross section

For the further transition to the geometric optics, which is used in the model at the stage of calculating the field transformation inside of the RSS, it is necessary to restore the field's directional vector at a point, as well as the polarization vector perpendicular to it and lying in a given polarization plane. To do this, it is enough to find the value of the gradient of the function at the desired point. To do this, one can use the numerical methods for finding the derivative of the formula 1. The figure 1 (b) shows an example of normal maps for the cross sections of the fields of Gaussian beams of different combinations of the orders of longitudinal and transverse modes. It seems obvious that for a multimode emitter it is possible to carry out an experiment for each of the chosen order of modes and then combine the obtained intensities in the far zone.

Knowing the field's amplitude, phase, directing vector and electric field vector in space at the point of interest to us, we can proceed to the second stage - the transformation of the field inside the RRS. To construct the ray trajectory, the directing vector $\vec{k}_1$ at a point P_1 , selected on the RRS plane, turns into $\vec{k}_2$ according to the Snell's law (figure 2 (a)). Next, the bead with which the intersection occurred is determined, and the point P_2 - the intersection of the ray $\{P_1, \vec{k}_2\}$ with the sphere representing the bead. At the found point, a normal to the sphere is constructed, then similar actions are carried out.

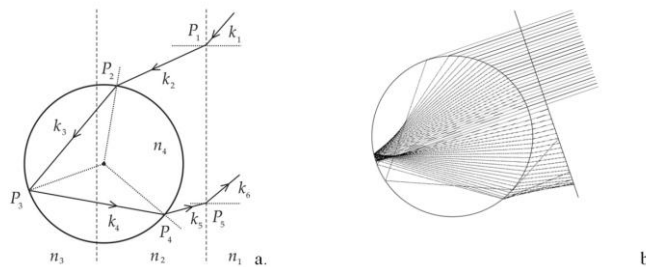


Fig. 2 Ray trajectory inside glass beads RRS: calculations (a) and simulation example (b)

According to the model, the points P_2 and P_4 should lie on the verge between the bead and the covering layer, and the point P_3 should lie on the verge between the bead and the substrate; re-reflections inside the bead or between neighboring beads are not considered. The point P_5 is sampled and is the point of the output amplitude-phase distribution (APD). In general case, the vectors $\vec{k}_1$ and $\vec{k}_2$ lie in one plane, and the vectors $\vec{k}_2, \vec{k}_3, \vec{k}_4, \vec{k}_5, \vec{k}_6$ - lie in another. Given each of the segments $l_i = P_{i-1}P_i$ that make up the trajectory of the light beam, the total phase shift φ is calculated

$$\varphi = \frac{2\pi}{\lambda} \sum_{i=1}^6 n_i l_i, \quad (2)$$

where λ - wavelength in vacuum,

n_i - absolute index of refraction of the medium for λ .

The phase shift when reflected at the point P_3 is considered separately, since the shift in the general case for the parallel and perpendicular components is different. Also, for each transition from one medium to another, the amplitude transmission and reflection coefficients for both field components are calculated using the Fresnel formulas [7].

The third stage occurs after the tracing is completed for each of the points, which means that the APD on the surface of the RRS is known - the problem of transforming the light field on the reflector is solved. Knowing this APD, applying the Fourier transform to it, we obtain the APD in the far zone [8], which means that it is possible to calculate the light intensity, which is an input parameter for quadratic detectors.

Analysis of the specifics of using a multimode emitter

Estimating the relative contribution of higher-order modes is essential for modeling, since this allows one to choose the order of the mode, which can be limited in a computer experiment while maintaining a satisfactory correspondence with the actual physical processes. Therefore, it is important to understand how the total power of the emitter is distributed over the modes of different orders [6].

It should be noted that multimode means the difference between the frequencies in the emitter. A dielectric, as a medium for propagation of radiation, respectively, has a dispersion that in practice, for example, can be described by the Sellmeier equation [7]

$$n^2 = 1 + \sum_i \frac{B_i \lambda^2}{\lambda^2 - C_i}, \quad (3)$$

where B_i and C_i - are experimentally determined Sellmeier coefficients.

The presence of dispersion means that the refractive index of the medium varies for different modes, which entails changes in the spatial shape of the signal reaching the RRS and also in the ray trajectory inside the RRS. As can be seen from formula 2, dispersion has an effect on the phase shift as well. All three of these consequences of multimode usage mean that the picture of the output APD is different for each of the modes.

In the formulation of the problem of studying the relationship of parameters in the detection system in the context of detecting a reflected signal, of interest is the contrast ratio of the speckle pattern - the clarity of the boundaries of each speckle, as well as the half-width of the first-order maximum outline dependance on the order of the mode.

It follows from the described influence of dispersion that the pattern contrast should decrease with an increase in the number of modes, the ratio of fluctuations of the maximum amplitude of the useful signal, which carries information about signal fluctuations due to turbulence, to the average should decrease. It is of interest to evaluate this influence not only from the point of view of studying the nature of the processes occurring in the system, but also from a practical point of view - such an

assessment allows you to choose from the modules of the laser radiation source on the market the most suitable for implementing the sensor in terms of such a parameter as cost / multimode, What you have in terms of the target parameters of the system translates into the cost/accuracy of detection (the fewer modes, the better the signal, but the laser is more expensive). To solve such a problem, it was necessary to create a transformation model in order to implement it as part of a simulation model and, after conducting a computer experiment, further give a specific quantitative assessment of the influence of multimode on the quality of detection.

Conclusion

A model for the transformation of multimode laser radiation on a retroreflective surface is presented. The physical processes that took place during the transformation were analyzed, and the intermediate results of the modeling were provided. The analysis of the influence of the multimode nature of the laser radiation source on the output amplitude-phase distribution shows that, due to dispersion, the contrast of the detected speckle pattern should decrease with an increase in the number of modes, and the ratio of fluctuations of the maximum amplitude of the useful signal carrying information about the acoustic signal to the average should decrease. The presented model is supposed to be used in the context of simulation modeling of laser fire alarm systems based on turbulence analysis. Upon this model the flexible computer simulation program is being developed, which allows for fine control over parameters of the system. Once such simulation is verified via comparison results of the physical and computer experiments, it can be used to determine optimal parameters for different configurations of such systems given target specification or constrains for the applied use case, thus assisting in creation and tuning such class of systems without spending much resources and time on the early stages of prototyping. Furthermore, the core idea of the research can be extended by applying developed models and schemas to the registration of the acoustic oscillations, vibration control, velocimetry and measuring turbulence parameters of both gas and liquid flows, besides further improving upon complex early fire ignition detection systems themselves.

REFERENCES

1. M. Francon Laser Speckle and Applications in Optics. Academic Press, 2012. 174 p.
2. Measures R. M. Laser Remote Sensing: Fundamentals and Applications. Malabar, FL : Krieger Publishing, 1992. 524 p.
3. Tetsuo Fukuchi, Tatsuo Shiina. Industrial Applications of Laser Remote Sensing. Bentham Science Publishers, 2012. 194 p.
4. Dolya G. M., Lytvynova O. S. Modeling of fluctuations of laser radiation scattered on the reflector array in a turbulent atmosphere // Laser and Fiber-Optical Networks Modeling (LFNM), 2011. 108-111 p.
5. Dolya G., Bondarenko K., Laser detector of acoustic oscillations utilizing a retroreflective surface with glass beads // EESA, Vol. 3, No. 11(63), 2020. 7-16 p.
6. Siegman, Anthony E. Lasers. Revised ed. edition. University Science Books, 1986. 1283 p.
7. Борн М., Вольф Э. Основы оптики. Москва : Наука, 1970. 856 с.
8. Гудмен Дж. Введение в Фурье-оптику. Москва : Мир, 1970. 363 с.

DOLYA Gregory – Doctor of Technical Sciences, professor; professor of the Department of Theoretical and Applied Systems Engineering, Faculty of Computer Science, V.N.Karazin Kharkiv National University, Svobody Sq., 4, 61022, Kharkiv, Ukraine; e-mail: gdolya@karazin.ua; ORCID: 0000-0001-8888-5189.

Scientific interests:

– *optoinformatics, laser location, specklometry.*

BONDARENKO Kostiantyn – postgraduate (Ph.D. student); Ph.D. student of the Department of Theoretical and Applied Systems Engineering, Faculty of Computer Science, V.N.Karazin Kharkiv National University, Svobody Sq., 4, 61022, Kharkiv, Ukraine; e-mail: konstantinb2016@gmail.com; ORCID: 0000-0002-2430-0716.

Scientific interests:

– *specklometry, computer simulation of optical processes.*

DONETS V., UGRYUMOV M., STRILETS V.

A MEASURE OF COMPACTNESS FOR FUZZY CLUSTERING BASED ON ENTROPY

The traditional method of fuzzy clustering and its modifications the Mixture Models, Gaussian Mixture Models does not aim to determine the optimal clusters count and requires this number as the input parameter of the method. To solve this problem with the method creation, it was suggested to evaluate the clusters received on some Partition Coefficient, which did not really solve the problem. Therefore, the use of a new metric based on Kullback-Leibler Entropy is considered.

To determine the metrics quality, it was proposed to compare the metric with earlier and existing ones: Partition Entropy Coefficient, Fukuyama-Sugeno Index, Partition Coefficient and Exponential Separation. For metrics evaluation experiments were done on clustering 4 classic datasets with different target classes, parameters and spatial resolution of clusters: Fisher Irises Dataset, Glass Identification Dataset, Wine Dataset, Breast Cancer Wisconsin Dataset. The experiments purpose was to get metrics indicating the target clusters count.

The experiments results are shown on several graphics for each of the test dataset and selected metrics. The results showed some better opportunities for the detection of the optimal clusters count had a metric based on Kullback-Leibler Entropy and the previously proposed metric based on Mahalanobis distance in combination with the membership value. This means that in this form, the proposed metric can be used to determine the optimal clusters count, but it is possible to consider the spatial resolution of clusters in the future works.

Key words: fuzzy c-means, data mining, clustering, compactness rate, intraclass distance, Mahalanobis distance, Kullback-Leibler divergence, determining clusters count..

Introduction

Data Mining is critical area for finding patterns in an unprocessed flow of information that appears in many areas such as automatic categorization of textual information [1], time series analysis, information flows processing in various technical and biological systems [2] and many others. One of the methods of Data Mining is clustering, the main task of which is to break the input data into clusters with maximum plausibility. Clustering methods usually have a problem with determining the required number of clusters, so this work considers and a few metrics for determining (compactness) of the optimal clusters count.

Related works

1. The method of fuzzy clustering [3, 4], which aims to separate a set of data into J clusters belonging to which is determined by some fuzzy function. The fuzzy clustering algorithm for determining the position of cluster centers is aimed at reducing the value of the cost function:

$$LOSS_m = \sum_{j=1}^J \sum_{i=1}^N \omega_{ij}^m d^2(x_i, c_j), \quad (1)$$

where $m \in (1, \infty)$ – the fuzzy parameter, N – samples count in the training set, J – clusters count, ω_{ij} – the membership value the i-th sample to the j-th cluster, x_i – the i-th sample, c_j – the j-th cluster, $d(x_i, c_j)$ – Euclidian distance between the i-th sample and the j-th cluster.

Accordingly, membership is determined by the following expression:

$$\omega_{ij} = 1 / \sum_{k=1}^J (d(x_i, c_j) / d(x_i, c_k))^{2/(m-1)}. \quad (2)$$

2. Partition Coefficient proposed Bezdek [3] and Dunn [4] which shows how a fuzzy solution corresponds to a rigid solution. Here, the rigid solution is formed by the classification of the object to the cluster because the greatest importance and the formula proposed by Bezdek and Dunn ($m = 2$):

$$PC = F(U) = 1/N \sum_{j=1}^J \sum_{i=1}^N \omega_{ij}^m. \quad (3)$$

The value of $F(U)$ measures the volume of overlapping between fuzzy clusters by membership mean and greater the value means better result. Among the disadvantages of the coefficient is the lack of taking into account the geometric sample space and its monotonous tendency to decrease with increasing the clusters count [5].

3. Partition coefficient Xie & Beni. To eliminate the disadvantages of Partition Coefficient Xie and Beni proposed a metric to determine compactness along clusters using the index [6]:

$$S = [\sum_{j=1}^J \sum_{i=1}^N \omega_{ij}^m \|c_j - x_i\|^2] / [N \cdot \min \|c_i - c_j\|^2]. \quad (4)$$

4. Conventional within-cluster sum of square (WCSS) is a metric based on the Euclidian distance, which is a conditional sum of squares within cluster [5]:

$$WCSS(k) = \sum_{j=1}^J \sum_{i \in P_j} \|x_i - c_j\|^2. \quad (5)$$

The S index and the WCSS metrics use the Euclidean distance, so they do not consider the spatial properties of clusters [6]. The article [6] shows that the WCSS measure has significant variations, even the proposed solution to the introduction of the punitive function has not solved the problem.

5. Partition entropy coefficient [7] determines the quality degree of fuzzy clustering, the smaller this value, the better the separation result. The disadvantage should be noted is the tendency to increase its value with the increase in the clusters count:

$$PEC = -\frac{1}{N} \sum_{j=1}^J \sum_{i=1}^N \omega_{ij} \cdot \log(\omega_{ij}). \quad (6)$$

6. Fukuyama-Sugeno Index. Fukuyama and Sugeno [8] offered an index, taking into account the compactness and cluster resolution. For this index, the smallest value is an indicator of the best clustering result. The index is determined by the expression:

$$FSI = LOSS_m - \sum_{i=1}^N \sum_{j=1}^J d(c_j, \bar{c}), \text{ where } \bar{c} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (7)$$

7. Partition Coefficient and Exponential Separation (PCES). Wu and Yang [9] proposed an index that combines the normalized PC with the exponential resolution of each cluster. The greater the value of this index, the better the separation results. The coefficient is determined by the expression:

$$PCES = \sum_{j=1}^J \sum_{i=1}^N (\frac{\omega_{ij}^m}{\omega_M}) - \sum_{k=1}^j \exp(-\min\{d^2(c_j, c_k)\} / \beta_T), \quad (8)$$

where $\omega_M = \min_j \{\sum_{i=1}^N \omega_{ij}^m\}$, $\beta_T = \frac{1}{J} \sum_{j=1}^J d^2(c_j, \bar{c})$.

Offered metrics

Eliminating the disadvantages of the existing c-means method and its modifications Mixture Models, Gaussian Mixture Models in the article [10] are considered the fuzzy clustering method modifications, with the developed new determining methods the optimal cluster count. There were proposed to use a number of modifications like using the Kullback-Leibler informative distance (11), the Mahalanobis distance based on the data distribution of a particular cluster (9, 10). In this work, the compactness rate (CR) used to simulate the process of agent evolution should be highlighted as functions for evaluating the validity of clusters, the expressions:

$$CR(c_j, P_j) = 1 / P_j \sum_{i=1}^{P_j} [MD^2(x_i, c_j)], \quad (9)$$

$$CR(c_j, P_j) = 1 / P_j \sum_{i=1}^{P_j} [\omega_{ij}^{-1} \cdot MD^2(x_i, c_j)], \quad (10)$$

$$CR(c_j, P_j) = 1 / P_j \sum_{i=1}^{P_j} [-D_{KL}(x_i, c_j)]. \quad (11)$$

Using the proposed expressions (9, 10) to evaluate the cluster validity, we get new compactness measures and, taking into account the expression (11), we get an Kullback-Leinler entropy based measure of compactness. The overall expression of compactness rate then takes the following look:

$$CR = 1 / J \sum_{j=1}^J CR(c_j, P_j). \quad (12)$$

Formulation of the problem

One of the main problems of the classical method of fuzzy clustering is uncertainty with the optimal number of clusters. Different metrics were used to solve this problem, especially in the context of direct use of the method without a teacher, but they cannot clearly indicate what number of clusters are optimal. Therefore, it will be advisable to compare the proposed metrics with the existing ones and draw preliminary conclusions about the feasibility of using them.

So, after the analysis of the literature, it was decided to test the following metrics: Partition coefficient (PC), the highest value is the best (3); Partition entropy coefficient (PEC), the smallest value is the best (6); Fukuyama-Sugeno Index (FSI), the smallest value is the best (7); Partition coefficient and exponential separation (PCES), the biggest value is the best (8).

And compare them with three metrics were offered: metrics by Mahalanobis distance (Mah), the smallest value is the best (9, 12); metric by combining Mahalanobis distance and membership values (Mah_Inv), the smallest value is the best (10, 12); metric accordingly to Kullback-Leibler entropy (KLC), the smallest value is the best (11, 12).

Used datasets for the metrics evaluation

To evaluate the selected metrics, it was decided to use classic datasets for testing machine learning methods:

- 1) Fisher's Iris. The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant. One class is linearly separable from the other 2; the latter are NOT linearly separable from each other. Contains 4 physical attributes.
- 2) Glass Identification Dataset contains 214 instances, 10 attributes are physical values, the dataset is divided into 7 classes – glass types.
- 3) Wine Data Set contains 178 instances, 13 attributes among which are integer and physical values, the dataset is divided into 3 classes.
- 4) Breast Cancer Wisconsin (Diagnostic) Dataset contains 569 records, 30 parameters, including integer and physical values, the dataset is divided into 2 classes.

Conducting research and results

Among the chosen datasets the number of optimal clusters for them are specified, so using the selected metrics will be a good sign of the optimal clusters count indication will be a reaching a minimum or maximum into the same classes. In order to evaluate not the clustering algorithm, namely metrics, it was decided to use the fuzzy clustering algorithm for all cases and expand it, adding the ability to calculate the selected metrics.

For the experiments results, we concluded that none of the existing metrics showed good results into the clear selection of the required number of clusters. FSI worked for Fisher's iris (Fig. 1.1) and for the classification of Wine (Fig. 1.3), although it had a large oscillation of results, for the classification. PC, PC, PEC and MAh_inv metrics, KLC behave exponentially by increasing or falling, reaching almost minimums or maxima near the required clusters count. The using of Mah_Inv and KLC metrics, but this was due to the using of the modified fuzzy c-means method, perhaps through these metrics, they have produced more accurate results for all dates. It should be noted that the Mah metric had quite poor results, reaching a minimum for the classification of Glass (Fig. 1.3), for others it had a great oscillation. The PCES metric, which have to reach the maximum, reached a maximum only on Fisher irises (Fig. 1.1), on others it simply increased linearly or oscillating.

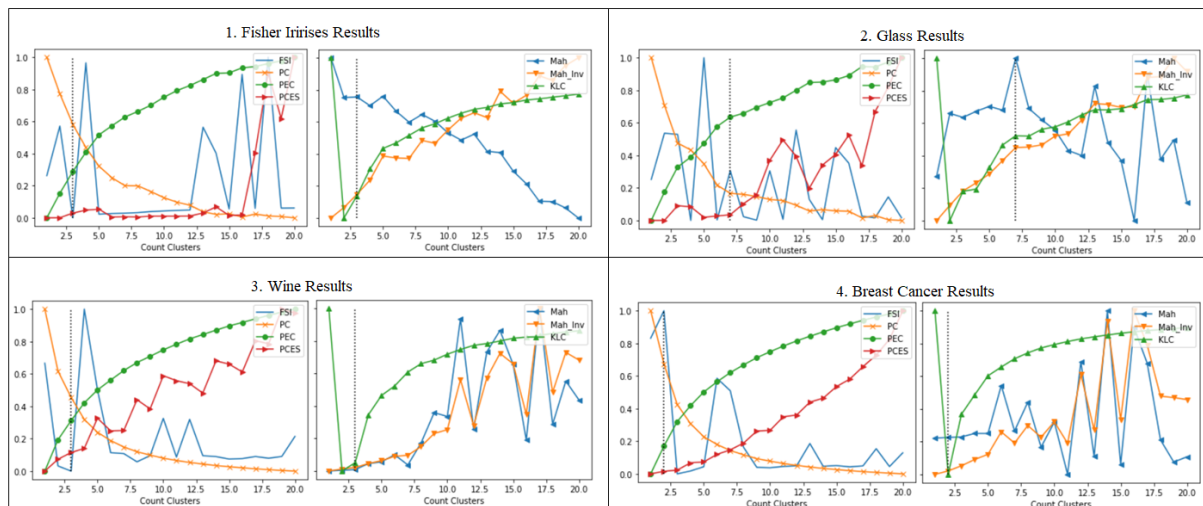


Fig. 1 Metric evaluation for each dataset and selected metric

Conclusions

According to the experimental results, we conclude that existing and proposed metrics are able to determine the optimal number of clusters in some cases. PC, PEC, Mah metrics required, in most cases, human intervention to determine the optimal cluster count, better results had FSI and PCES and the most persistent results were obtained from Mah_inv and KLC. Therefore, a logical extension of the work can be the development a metric that combines the advantages of the proposed and offsets their disadvantages, one of the possible solutions to the problem can be calculating the average density of clusters and taking into account their separation.

REFERENCES

1. L. M. Abualigah, A. T. Khader, M. A. Al-Betar, and M. A. Awadallah. A krill herd algorithm for efficient text documents clustering. In *Proceedings of the IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, USA, 2016, pp. 67–72.
2. S. R. B. Prabhu, R. Mahalakshmi, S. Nithya et al.. A Review of Energy Efficient Clustering Algorithm for Connecting Wireless Sensor. *Network Fields, Social Science Electronic Publishing*, New York, USA, 2017.
3. Bezdek JC. Pattern recognition with Fuzzy Objective Function Algorithms. New York: Plenum Press, 1981. 198 p.
4. E.Trauwaert. On the meaning of Dunn's partition coefficient for fuzzy clusters. *Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium*. 2003. DOI: [10.1016/0165-0114\(88\)90189-3](https://doi.org/10.1016/0165-0114(88)90189-3) (Last accessed 25.10.2022).
5. Mingrui Zhang, Wei Zhang, Hugues Sicotte, and Ping Yang. A New Validity Measure for a Correlation-Based Fuzzy C-means Clustering Algorithm. *Conf Proc IEEE Eng Med Biol Soc*. 2010. DOI: 10.1109/IEMBS.2009.5332582 (Last accessed 25.10.2022).
6. Xie X. L., Beni G. A Validity Measure for Fuzzy Clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1991. vol. 13, pp. 841–847. DOI: [10.1109/34.85677](https://doi.org/10.1109/34.85677) (Last accessed 25.10.2022).
7. J. C. Bezdek. Cluster validity with fuzzy sets. *Journal of Cybernetics*. 1974. vol. 3, pp. 58–73.
8. Y. Fukuyama, M. Sugeno. A new method of choosing the number of clusters for the fuzzy c-means method. in *Proceedings of the Fifth Fuzzy System Symposium*, Kobe, Japan, 1989.
9. K.-L. Wu, M.-S. Yang. A cluster validity index for fuzzy clustering. *Pattern Recognition Letters*. 2005. vol. 26, no. 9, pp. 1275–1291.
10. Strilets V., Donets V., Ugryumov M., Zelenskyi R., Goncharova T. Agent-Oriented Data Clustering for Medical Monitoring. *Journal of Radioelectronic and Computer Systems*. Kharkov “KhAI”, 2022. ISSN:1814-4225E-ISSN:2663-2012.

DONETS Volodymyr – PhD student of the Theoretical and Applied system engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine, e-mail: vol.donets@gmail.com, ORCID: 0000-0002-5963-9998.

Scientific interests:

– machine learning methods in multicriteria problems of robust design and intellectual diagnosis of systems (ROD & IDS) in the face of uncertainty.

UGRYUMOV Mykhaylo – Doctor of Technical Sciences, Professor, Professor of the Theoretical and Applied system engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine, e-mail: m.ugryumov@karazin.ua, ORCID: 0000-0003-0902-2735.

Scientific interests:

– machine learning methods in multicriteria problems of robust design and intellectual diagnosis of systems (ROD & IDS) in the face of uncertainty.

STRILETS Viktoriia – Candidate of Technical Sciences, Associate Professor of the Theoretical and Applied system engineering Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine, e-mail: viktorija.strilets@karazin.ua, ORCID: 0000-0002-2475-1496.

Scientific interests:

– machine learning methods in multicriteria problems of robust design and intellectual diagnosis of systems (ROD & IDS) in the face of uncertainty.

МЕТОД АНАЛІЗУ ПАРАМЕТРІВ ТРЕНДУ У ТРАФІКУ КОМП'ЮТЕРНОЇ МЕРЕЖІ

The purpose of the work is to develop a computer model for detecting computer network traffic trends and their further evaluation to improve the quality of traffic forecasting.

Research methods

Collection and pre-processing of input data, their formalization and subsequent quantitative and qualitative analysis were used in the work. Statistical analysis methods were used for data analysis. An assessment of existing mathematical models and computational methods was used to predict the traffic of computer networks.

The results

An analysis of the existing methods of detecting trends in time series was carried out. Based on the results of the work, a model for detecting computer network traffic trends was developed using the analysis of a unique data set formed by parsing logs (system files) from monitoring the state of computer network traffic. Improved autoregressive trend data forecasting model using gradient boosting.

Conclusions

Based on the results of the research, a model for detecting computer network traffic trends was proposed. The work created a data set that characterizes the computer network for its further analysis. With the help of methods of researching time series trends, as well as linear regression algorithms, it was possible to increase the accuracy and quality of forecasting the traffic state of computer networks. The results of the work can be used in various types of tasks, such as the tasks of computer network traffic analysis, forecasting of the state of technical systems, development of computer network models, evaluation of the functioning of computer networks. The use of modern methods of traffic analysis and estimation play a big role in the development of computer networks, so it is necessary to constantly improve existing methods and introduce new ones.

Ключові слова: *ARIMA, тренд, часові ряди, комп'ютерні мережі, трафік комп'ютерних мереж, прогнозування, градієнтний бустінг, машинне навчання, регресія, статистичні методи.*

Вступ

Сучасний час можна охарактеризувати стрімким зростанням обсягів трафіку комп'ютерних мереж. Основним завданням при модернізації комп'ютерної мережі є аналіз мережевого трафіку, а саме його структури та обсягу. Особливої уваги потребує виділення тенденцій зміни трафіку мереж за певний період часу, що дасть змогу більш точно спрогнозувати навантаження на канали зв'язку. Саме для вирішення цієї задачі пропонується застосувати методи виявлення трендів трафіку комп'ютерних мереж для розробки комп'ютерної моделі прогнозування показників трафіку комп'ютерної мережі.

Методи виявлення тренду у часових рядах

Виділення тренду мережевого трафіку є дуже важким і водночас дуже важливим завданням, оскільки його вирішення дозволяє здійснювати прогноз, у якому випадкова складова часового ряду використовується з метою оцінки точності прогнозу стану трафіку. В даний час найчастіше використовуються два методи виділення тренду. Перший метод полягає в тому, що за емпіричними даними часового ряду підбирається крива (математична модель), яка з найбільшою точністю описує часовий ряд [1]. При цьому в якості математичних моделей використовуються різні функції: рівняння прямої та експоненти, парабола (квадратична, кубічна і більш високих ступенів), логістична крива, крива Гомперца та ін. Другий метод виділення тренду полягає в згладжуванні ряду за методом ковзного середнього [2]. При цьому зазвичай знаходять середнє значення трьох (або п'яти) перших членів, далі беруться наступні три члени зі зміщенням на одиницю і середнє. Таким чином, вдається зменшити випадкову складову.

При аналізі часового ряду є доцільним виявити трендову складову [3]. Для цього існують критерії виявлення трендів. До таких відносять наступні:

- критерій Аббе-Лінника;
- критерій Кокса-Стюарта;
- критерій Фостера-Стюарта;
- критерій Манн-Кендала.

Кожен з цих критеріїв використовується індивідуально для розглядуваного часового ряду, в залежності від властивостей даних, на яких цей критерій буде застосовуватись.

Набір даних дослідження

За основу для набору даних було обрано логи комерційної компанії, яка спеціалізується на аналізі мережевого трафіку для крупних клієнтів [4]. З отриманих логів за допомогою розробленого програмного забезпечення було сформовано вибірку (з використанням парсингу текстових файлів), яка налічує більше 300 000 рядків. Об'єкти набору даних мають 9 параметрів:

1. Max packet size (bytes) - відповідає за показник максимального розміру пакету, що передається по мережі;
2. Send latency (usec) - показник затримки відправника у мілісекундах;
3. Recv latency (usec) - відповідає за показник затримки отримувача у мілісекундах;
4. Min packet spacing (usec) - відповідає за часову різницю між пересиланням пакетів (інтервал передачі пакетів) у мілісекундах;
5. Max rate (Mbps) - відповідає за показник максимальної швидкості даних в мережі на момент заміру;
6. ADR (Mbps) - показник адаптивного бітрейту мережі;
7. Grey bandwidth resolution - показник сірого трафіку мережі;
8. Measurement date - дата, коли було сформовано дані вимірювання показників трафіку мережі;
9. Available bandwidth (Mbps) - наявна пропускна здатність мережі.

Розробка моделі оцінки тренду та прогнозування даних на його основі

В якості вхідних даних для моделі було обрано два параметри з найбільшим коефіцієнтом кореляції – Grey bandwidth resolution і Available bandwidth (Mbps). Усі дані було згруповано помісячно. Попередня обробка даних проводилась з використанням статистичних методів (усі N/A комірки були замінені на середнє значення всього набору), це дозволило отримати валідні дані для побудови моделі [5].

Наступним кроком був аналіз ряду на стаціонарність за допомогою теста Діка-Фулера. Цей тест показав, що відібрані 2 набори даних є стаціонарними.

В роботі використовується модель ковзного середнього з використанням елементів авторегресії – ARIMA (Autoregressive Integrated Moving Average). Така модель зазвичай використовується для прогнозування даних стаціонарних часових рядів та використовує трендову компоненту ряду, завдяки чому має високі показники точності серед усіх моделей даного типу [6]. Модель була обрана через те, що досить непогано себе поводить саме з даними, в яких немає сезонної компоненти.

Для оцінки наявності трендової компоненти в роботі було використано критерій Манн-Кендала [7]. При цьому для того, щоб перевірити наявність тренду було запропоновано метод, при якому загальна оцінка наявності тренду складається на основі як мінімум трьох тестів: тест Манн-Кендала, модифікований тест Манн-Кендалла (Ю Ванг) та модифікований тест Манн-Кендала (Х. Рао). Результат методу інтерпретується як ціле число в інтервалі $[-1;1]$. Якщо результат є від'ємне число – то тренд існує, та є низхідним. В випадку, коли результат 0 – це свідчить про те, що тренду не існує. Якщо результат дорівнює 1 – це говорить про те, що в даних є тренд, який має висхідний напрямок. Таким чином можна зробити комплексну оцінку наявності тренду, та отримати об'єктивні результати аналізу.

При виборі параметрів моделі ARIMA p, d, q було застосовано ітеративний метод підбору цих параметрів на основі критерію Акаїке. Для обох наборів вхідних даних кращій результат показала модель ARIMA(2,0,0)(0,0,0) (показник критерію Акаїке – 264.068).

Після побудови моделі було отримані результати, графічне представлення яких можна побачити на Рис. 1.

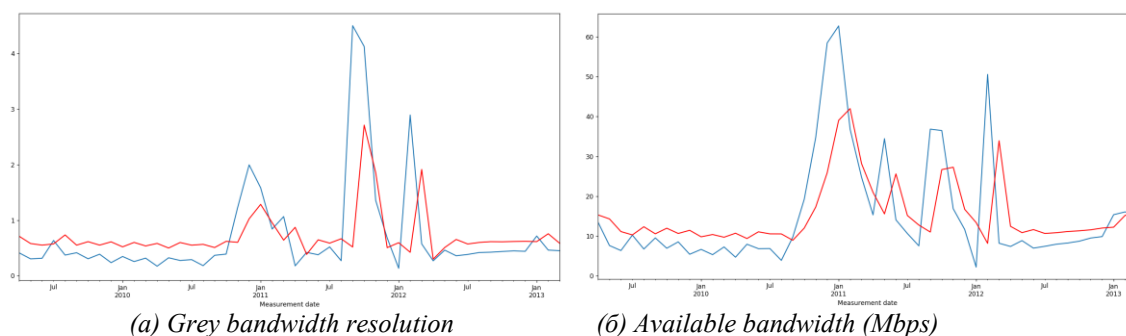


Рис. 1 Порівняльні графіки результатів моделі ARIMA

Синім кольором на графіку відокремлені актуальні дані, червоним – результати моделі прогнозування. Як можна побачити за попередньою оцінкою, точність прогнозування не є високою.

Покращення моделі прогнозування за допомогою методів машинного навчання

Запропоновано покращити результати роботи моделі прогнозування параметрів трафіку з використанням методів машинного навчання, а саме методу градієнтного бустінгу. Даний метод зазвичай використовується для зменшення відхилень та помилок. Модель авторегресії ковзного середнього може включати в себе екзогенні змінні, завдяки яким можна використовувати екзогенні регресори (Exogenous Regressors) для покращення результатів, що і було зроблено в даній роботі [8].

На рис 2. можна побачити результати роботи нової моделі. Золотим кольором позначено результати прогнозування значень, чорним - актуальні дані.

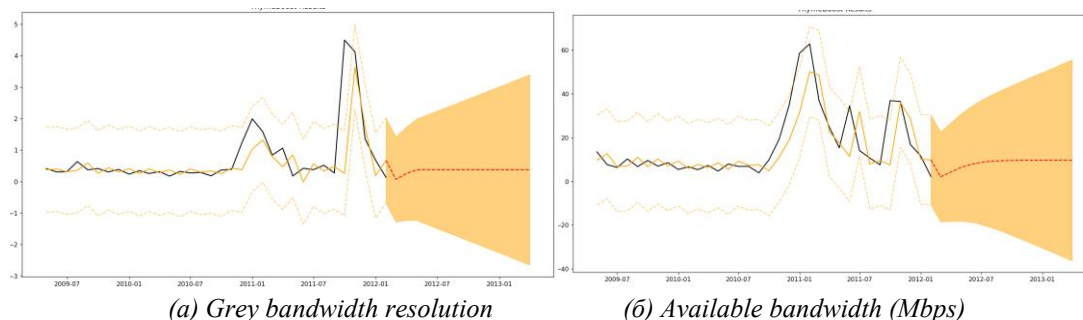


Рис. 2 Порівняльні графіки результатів покращеної моделі ARIMA

Для того, щоб упевнитися в тому, що результати моделі дійсно було покращено, було проведено порівняння актуальних метрик моделей. В якості набору характеристик точності моделі було використано показники MAE (Mean absolute error), MSE (Mean squared error) та RMSE (Root mean squared error). Результати порівняння наведено в Таблиці 1.

Таблиця 1. Порівняння показників метрик точності моделі

	MAE	RMSE	MAPE
ARIMA	7.973	13.626	0.618
Покращена ARIMA	5.712	13.337	0.443

Інтерпретуючи результати порівняння можна сказати, що дійсно, після застосування методу градієнтного бістінгу якість роботи моделі вдалося покращити.

Висновки

На сьогоднішній день існують різні методи виявлення та оцінки трендів у трафіку комп'ютерних мереж. Побудова моделей аналізу трафіку комп'ютерних мереж за допомогою

дослідження трендів часових рядів, а також лінійних алгоритмів регресії дозволять підвищити точність та якість аналізу стану комп'ютерних мереж. У роботі був сформований набір даних, який характеризує комп'ютерну мережу, для його подальшого аналізу і виявлення тренду.

Результати можуть бути використані в різних типах задач, таких як задачах аналізу трафіку комп'ютерних мереж, прогнозування станів технічних систем, розробка моделей комп'ютерних мереж, оцінювання функціонування комп'ютерних мереж, а також використані для аналізу та оцінки поточного трафіку комп'ютерних мереж для подальшої оптимізації, пошуку перспектив для подальшого розвитку у структурі мережі, виявлення проблем, що негативно впливають на систему. Використання сучасних методів аналізу та оцінки трафіку відіграють велику роль у розвитку комп'ютерних мереж, тож необхідне постійне вдосконалення існуючих методів та впровадження нових.

ЛІТЕРАТУРА

1. Gray, Katharine Lynn, "Comparison of Trend Detection Methods" (2007). Graduate Student Theses, Dissertations, & Professional Papers. 228.
2. IBM Corporation Documentation, "Trend detection", 2021. URL: <https://www.ibm.com/docs/en/siffs/2.0.3?topic=learning-trend-detection>
3. Часові ряди. [Електронний ресурс] Режим доступа: <https://kstat.pnu.edu.ua/wp-content/uploads/sites/63/2018/04/%D0%A7%D0%B0%D1%81%D0%BE%D0%B2%D1%96-%D1%80%D1%8F%D0%B4%D0%B8.pdf>
4. Real World Networking datasets. [Електронний ресурс] Режим доступа: <https://gist.github.com/stefanbschneider/96602bb3c8b256b90058d59f337a0e59>
5. Ковтун Н. В. Теорія статистики: Курс лекцій, практикум. – К.: ІмексЛТД, 2007. – 276 с
6. Дуброва, Т. А. Статистические методы прогнозирования: Учеб. пособие для вузов. - М.: ЮНИТИ-ДАНА, 2003. – 206 с.
7. Mann, H.B, Non-parametric tests against trend (1945), *Econometrica*, 13:163-171.
8. Robert Nau, "Statistical Forecasting: notes on regression and time-series analysis" (2020). ARIMA Models for time-series forecasting. URL: <https://people.duke.edu/~rmau/arimreg.htm>

ДОРОШЕНКО Максим Ігорович – студент групи KI-51 факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022; e-mail: xa11867769@student.karazin.ua; ORCID: 0000-0003-3645-252X.

Наукові інтереси:

– *Програмування, структури даних, автоматизація програмних систем.*

СТРИЛЕЦЬ Вікторія Євгеніївна – к. т. н., доцент; доцент кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: viktoria.strilets@karazin.ua, ORCID 0000-0002-2475-1496.

ЄГОРОВ Є.С., ШМАТКОВ С.І.

КЛАСТЕРИЗАЦІЯ ДАНИХ НА ОСНОВІ МЕТОДІВ БАГАТОВИМІРНОГО ДИСКРИМІНАНТНОГО АНАЛІЗУ ТА С- СЕРЕДНІХ

Актуальність: інтелектуальний аналіз даних пов'язаний із пошуком закономірностей, що дозволяють отримати нову інформацію про досліджувані дані. Останнім часом інтерес до засобів та методів аналізу даних підвищився у зв'язку з розвитком технологій для зберігання та обробки великих масивів даних. В той час як традиційні статистичні методи аналізу даних займаються здебільшого оцінкою параметрів для вже виведених закономірностей, методи інтелектуального аналізу даних дозволяють виявити нові, скриті закономірності в досліджуваних даних. Таким чином визначення функціонального зв'язку між показниками об'єктів є складовою частиною актуального завдання пошуку оптимальних впливів на об'єкти і не вирішена в повній мірі на сьогоднішній день. При цьому найбільш цікавими є результати, що адекватно представляють розбиття простору на кластери, які відносяться до різних функціональних співвідношень, що зв'язують показники об'єктів, що розглядаються в даній області. Такі кластери логічно називати функціональними, а їх форма в загальному випадку може бути довільною. Для адекватного поділу вихідної сукупності на такі однорідні групи необхідне застосування нових інформаційних технологій.

Мета: розробити алгоритм кластеризації даних на основі методів багатовимірного дискримінантного аналізу та с-середніх, що дозволяє формувати кластери довільної форми та оцінювати динаміку змін характеристик об'єктів.

Методи дослідження: методологічною базою є роботи по методам кластеризації, в тому числі присвячені практичним аспектам їх застосування. Програмне забезпечення розроблено за допомогою мови Java. Також були використані наступні бібліотеки: pmml4s, Vectorz, tableaw, Weka і інші.

Результати: Проведено огляд існуючих підходів до кластеризації даних з пропусками і здійснено порівняльний аналіз декількох підходів у разі застосування методу k-середніх. Для порівняльного аналізу обрано методи заповнення пропусків середнім, медіаною і на основі методу головних компонент, а також методи кластеризації даних з пропусками с-середніх і с-POD (останні базуються на методі с-середніх).

Висновки: Розроблено програмний продукт, який забезпечує проведення кластеризації даних з пропусками та порівняльного аналізу обраних методів. За допомогою розробленого програмного продукту проведено обчислювальні експерименти, результати яких показали, що найкращу якість кластеризації даних з пропусками забезпечує метод кластеризації с-середніх та заповнення пропусків на основі методу головних компонентів.

Ключові слова: кластеризація, метод с-середніх, пропуски, заповнення пропусків, зовнішні метрики якості, обчислювальний експеримент.

Постановка проблеми та її актуальність

Задача кластерного аналізу полягає у розбитті заданого набору об'єктів на підмножини, що називаються кластерами, так, щоб кожен кластер складався зі схожих об'єктів, а об'єкти різних кластерів істотно відрізнялися.

На практиці часто виникає необхідність розв'язувати цю задачу на різних даних методом багатовимірного дискримінантного аналізу. Це становить проблему, оскільки усі поширені й загальновідомі методи кластеризації (с-середніх, ієрархічні, ЕМ тощо) працюють лише на даних, які заздалегідь підготовлені і мають готову форму. Для подолання цієї проблеми запропоновано декілька різних підходів, але актуальним досі залишається питання вибору найкращого з них. Тому в рамках даної роботи було поставлено за мету проаналізувати існуючі підходи до кластеризації даних і провести їх порівняльний аналіз у разі застосування методу с-середніх[1].

Аналіз останніх досліджень і публікацій

Можна виділити три загальні підходи до кластеризації даних:

1. Видалення з набору даних об'єктів або ознак, у яких є пропущені значення, і подальша кластеризація даних без пропусків.
2. Заповнення пропусків і подальша кластеризація повних даних.
3. Використання методів кластеризації, що враховують наявність пропусків у даних.

Недолік першого підходу полягає в тому, що він призводить до скорочення набору даних та неможливості кластеризувати об'єкти з пропущеними значеннями за допомогою методів багатовимірного дискримінантного аналізу. Проте цей підхід простий в реалізації і може бути рекомендований, коли кількість пропусків дуже мала.

Другий підхід є найбільш розповсюдженим на практиці. В його рамках запропоновано значну кількість методів заповнення пропусків.

В рамках третього підходу до кластеризації даних з пропусками, зазвичай, розробляють модифікації відомих методів. Наприклад, на базі методу с-середніх, що є предметом розгляду в даній роботі

Постановка задачі

Необхідно провести порівняльний аналіз різних підходів до кластеризації даних з пропусками типу MCAR методом с-середніх. Припускається, що набір даних містить лише кількісні показники.

Створення програмного продукту

З метою проведення порівняльного аналізу було розроблено програмний продукт. У процесі його розробки використано такі технології:

- Середовище розробки – IntelliJ IDEA.
- Мова програмування – Java.
- Додаткові бібліотеки – pmml4s, Vectorz, tableaw, Weka. Перші дві були використані у процесі реалізації ЕМ алгоритму, а дві останні – для побудови дендрограми в ієрархічних методах.

Програмний продукт має такий функціонал:

- Завантаження або генерація даних. Передбачено генерацію повних даних із заздалегідь відомим розбиттям по кластерах, а також генерацію даних з пропусками на їх основі.
- Заповнення пропусків середнім або медіаною відповідної ознаки, а також на основі МГК.
- Кластеризація даних з пропущеними значеннями методами с-середніх та с-POD.
- Кластеризація даних без пропусків методом с-середніх (алгоритм Ллойда). Також реалізовано кластеризацію ієрархічними методами та ЕМ алгоритмом.
- Оцінювання якості роботи різних підходів до кластеризації даних з пропусками шляхом порівняння результатів кластеризації із заздалегідь відомими на основі таких індексів: Rand, Jaccard, Fowlkes- Mallows, Sorensen-Dice (усі індекси набувають значень від 0 до 1).

Експериментальна кластеризація даних

За допомогою створеного програмного забезпечення було проведено ряд обчислювальних експериментів на змодельованих наборах даних. У ході кожного експерименту моделювалось 100 двовимірних наборів даних згідно суміші нормальних розподілів. Усі 100 наборів моделювалися з однаковими параметрами кластерів. Кожен набір містив 800 об'єктів, розділених на 4 сферичні кластери (по 200 об'єктів у кожному кластері).

Нижче наведено і проаналізовано результати трьох експериментів. Під час першого експерименту моделювались набори даних, в яких кластери були сильно відокремлені один від одного. У ході другого експерименту розглядалися набори з кластерами, що також не перетиналися, але були розташовані більш близько. У третьому експерименті було розглянуто випадок, коли кластери перетиналися. В усіх експериментах для кожного кластеру дисперсійно-коваріаційну матрицю було задано як одиничну матрицю. Центри кластерів для було задано такими:

- Експеримент 1: (5, 22), (18, 23), (18.5, 11), (35, 29).
- Експеримент 2: (0, 10), (4, 10), (4, 4), (7, 6).
- Експеримент 3: (0, 4), (3, 5), (4, 7), (6, 3).

Усередненні значення Rand Index для кожного з трьох експериментів наведено нижче (табл. 1–3). Чим вищі значення індексу, тим якіснішою є кластеризація даних. Одержані результати дозволяють зробити такі висновки:

- Якщо кластери дуже добре відокремлені один від одного, то найкращу якість кластеризації забезпечує метод с-середніх майже при будькому відсотку пропущених значень. Усі методи заповнення пропусків у такому випадку значно спотворюють дані і, як наслідок, забезпечують гіршу якість кластеризації.
- Якщо кластери досить близько розташовані або взагалі перетинаються, то найкращу якість кластеризації поряд з методом с-середніх показує заповнення пропусків на основі МГК.

Таблиця 1. Усереднене значення Rand Index $\pm$ його середньоквадратичне відхилення під час експерименту 1

Відсоток пропусків, %	Заповнення пропусків середнім	Заповнення пропусків медіаною	Заповнення пропусків на основі МГК	Методс-середні х	Метод с-POD
10	0,708 $\pm 0,00884$	0,849 $\pm 0,00236$	0,989 $\pm 0,0013$	0,944 $\pm 0,00118$	0,837 $\pm 0,00255$
20	0,657 $\pm 0,00985$	0,721 $\pm 0,00289$	0,852 $\pm 0,00149$	0,869 $\pm 0,00046$	0,728 $\pm 0,00059$
30	0,638 $\pm 0,00132$	0,665 $\pm 0,0012$	0,722 $\pm 0,00111$	0,820 $\pm 0,0002$	0,683 $\pm 0,0002$
40	0,622 $\pm 0,00027$	0,640 $\pm 0,00147$	0,654 $\pm 0,00292$	0,769 $\pm 0,00227$	0,622 $\pm 3,3E-05$
50	0,615 $\pm 0,00230$	0,622 $\pm 0,00042$	0,646 $\pm 0,00059$	0,706 $\pm 0,00054$	0,560 $\pm 0,00052$
60	0,612 $\pm 0,00128$	0,602 $\pm 0,0023$	0,647 $\pm 0,00221$	0,639 $\pm 0,00188$	0,500 $\pm 0,01261$

Таблиця 2. Усереднене значення Rand Index $\pm$ його середньоквадратичне відхилення під час експерименту 2

Відсоток пропусків, %	Заповнення пропусків середнім	Заповнення пропусків медіаною	Заповнення пропусків на основі МГК	Метод с-середні х	Метод с-POD
10	0,783 $\pm 0,00565$	0,784 $\pm 0,01586$	0,878 $\pm 0,01005$	0,894 $\pm 0,00089$	0,845 $\pm 0,00042$
20	0,701 $\pm 0,00196$	0,689 $\pm 0,00643$	0,841 $\pm 0,00552$	0,809 $\pm 0,00016$	0,743 $\pm 0,00078$
30	0,679 $\pm 0,00048$	0,661 $\pm 0,00333$	0,793 $\pm 0,00098$	0,761 $\pm 0,00135$	0,693 $\pm 0,00075$
40	0,635 $\pm 0,00215$	0,635 $\pm 0,00098$	0,724 $\pm 0,00063$	0,710 $\pm 0,0015$	0,649 $\pm 0,00157$
50	0,606 $\pm 0,00156$	0,603 $\pm 0,00154$	0,674 $\pm 0,0043$	0,652 $\pm 0,0015$	0,596 $\pm 0,00154$
60	0,568 $\pm 0,00568$	0,573 $\pm 0,00518$	0,650 $\pm 0,00038$	0,586 $\pm 0,00058$	0,543 $\pm 0,00386$

Таблиця 3. Усереднене значення Rand Index $\pm$ його середньоквадратичне відхилення під час експерименту 3

Відсоток пропусків, %	Заповнення пропусків середнім	Заповнення пропусків медіаною	Заповнення пропусків на основі МГК	Метод с-середні х	Метод с-POD
10	0,747 $\pm 0,00072$	0,771 $\pm 0,00357$	0,842 $\pm 0,02166$	0,825 $\pm 0,00136$	0,786 $\pm 0,00258$
20	0,685 $\pm 0,00216$	0,677 $\pm 0,00519$	0,778 $\pm 0,00314$	0,761 $\pm 0,00169$	0,695 $\pm 0,00233$
30	0,656 $\pm 0,00035$	0,652 $\pm 0,00131$	0,730 $\pm 0,00197$	0,724 $\pm 0,00056$	0,647 $\pm 0,00046$
40	0,619 $\pm 0,00283$	0,626 $\pm 0,00112$	0,681 $\pm 0,0064$	0,688 $\pm 0,00192$	0,595 $\pm 0,00108$
50	0,581 $\pm 0,0012$	0,604 $\pm 0,00208$	0,634 $\pm 0,00418$	0,645 $\pm 0,00093$	0,545 $\pm 0,00266$
60	0,546 $\pm 0,00789$	0,578 $\pm 0,00183$	0,620 $\pm 0,01022$	0,592 $\pm 0,00081$	0,494 $\pm 0,00104$

Результати та висновки

Проведено огляд існуючих підходів до кластеризації даних з пропусками і здійснено порівняльний аналіз декількох підходів у разі застосування методу k-середніх. Для порівняльного аналізу обрано методи заповнення пропусків середнім, медіаною і на основі методу головних компонент, а також методи кластеризації даних з пропусками c- середніх і c-POD (останні базуються на методі c-середніх). Розроблено програмний продукт, який забезпечує проведення кластеризації даних з пропусками та порівняльного аналізу обраних методів. За допомогою розробленого програмного продукту проведено обчислювальні експерименти, результати яких показали, що найкращу якість кластеризації даних з пропусками забезпечує метод кластеризації c- середніх та заповнення пропусків на основі методу головних компонентів.

ЛІТЕРАТУРА

1. Steinley D. K-means clustering: A half-century synthesis. *British Journal of Mathematical and Statistical Psychology*. 2006. Vol. 59. P. 1–34.
2. Глушко О. Обработка пропусков в данных – часть 1. URL: <https://basegroup.ru/community/articles/missing> (дата звернення: 14.11.2019).
3. Литтл Р. Дж. А., Рубин Д. Б. Статистический анализ данных с пропусками. Финансы и статистика, 1990. 336 с.
4. Guan N.C., Yusoff M.S.B. Missing values in data analysis: Ignore or Impute? *Education in Medicine Journal*. 2011. Vol. 3 (1). P. e6–e11. DOI:10.5959/eimj.3.1.2011.or1
5. Pigott T.D. A Review of Methods for Missing Data. *Educational Research and Evaluation*. 2001. Vol. 7. No. 4. P. 353–383. DOI:10.1076/edre.7.4.353.8937
6. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. Новосибирск: ИМ СО РАН, 1999. 270 с.
7. Wagstaff K. Clustering with Missing Values: No Imputation Required. *Proceedings of the Meeting of the International Federation of Classification Societies*. 2004. P. 649–658. DOI:10.1007/978-3-642-17103-1_61
8. Chi J.T., Chi E.C., Baraniuk R.G. k-POD. A Method for k-Means Clustering of Missing Data. *The American Statistician*. 2016. Vol. 70. Issue 1. P. 91–99. DOI: 10.1080/00031305.2015.1086685

ЄГОРОВ Єгор Сергійович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: xa12283976@student.karazin.ua; ORCID: 0000-0001-8731-7198.

Наукові інтереси:

- методи обчислюваного інтелекту;
- прогнозування стану систем на основі даних моніторингу.

ШМАТКОВ Сергій Ігорович – завідувач кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: s.shmatkov@karazin.ua; ORCID: 0000-0002-0298-7174.

Наукові інтереси:

- методи обчислюваного інтелекту;
- прогнозування стану систем на основі даних моніторингу.

ЄСІНА М. В., ГОРБЕНКО І. Д., КАЧКО О. Г.

РОЗРОБКА, СТАНДАРТИЗАЦІЯ ТА ВПРОВАДЖЕННЯ СТІЙКИХ ПОСТКВАНТОВИХ КРИПТОПРИМІТИВІВ НА МІЖНАРОДНОМУ ТА НАЦІОНАЛЬНОМУ РІВНЯХ

The object of research is the processes of reducing risks for vulnerable (existing) cryptographic systems, development, standardization and implementation of standardized stable post-quantum crypto primitives of asymmetric encryption (AE), electronic signature (ES) and key encapsulation protocols (KEP) at the international and national levels. The subject of research is the methods of synthesis and methods of estimation, comparative analysis and application of new evidentially stable national and international standardized cryptographic primitives of asymmetric encryption, electronic signature and key encapsulation protocols at the international and national levels. The methods of research are methods of theoretical and practical cryptology, methods of mathematical and software modeling of cryptographic transformations such as AE, ES, KEP. The goal of the report is: justification of the choice, development and experimental research of modern and post-quantum standardized cryptographic transformations of AE, KEP and ES for cryptographic protection of information (messages) in communication networks, providing communication network users with services of identification, authentication, integrity, confidentiality, availability, non-repudiation, cryptographic survivability and authorization of each user of access in the communication network; analysis of the state of international standardization processes and conditions for the implementation of asymmetric post-quantum cryptography according to the 3rd round of the NIST USA and intermediate post-quantum and quantum studies of the European Union (EU); analysis of the state of national standardization processes and conditions and requirements for the implementation of post-quantum asymmetric cryptography in Ukraine, comparative analysis of post-quantum national and international standards and draft standards of AE, ES, KEP according to requirements for safety, productivity and operational characteristics.

Ключові слова: асиметричне шифрування, електронний підпис, криптографічний примітив, постквантове криптоперетворення, протокол інкапсуляції ключів, стандартизація, NIST.

Вступ

В останні роки спостерігається стійкий прогрес у створенні квантових комп'ютерів. У разі реалізації великомасштабних квантових комп'ютерів вони будуть загрожувати безпеці багатьох широко використовуваних криптосистем з відкритим ключем. Схеми встановлення ключів і електронні підписи, засновані на факторизації, дискретних логарифмах і криптографії на еліптичних кривих, найбільш сильно постраждають. Симетричні криптографічні примітиви, такі як блокові шифри і геш-функції, будуть порушені лише незначно.

У відповідь на ці виклики Національний інститут стандартів та технологій (NIST) ініціював відкритий конкурсний процес для вибору квантовостійких криптографічних алгоритмів. Нові стандарти криптографії з відкритим ключем визначатимуть алгоритми для цифрових підписів, шифрування з відкритим ключем і встановлення ключа. Передбачається, що ці алгоритми будуть здатні захистити конфіденційну інформацію в доступному для огляду майбутньому, в тому числі після появи квантових комп'ютерів.

NIST публічно оголосив про початок подання заявок до процесу стандартизації PQC у грудні 2016 року. До листопада 2017 року було подано 82 алгоритми-кандидати. Незабаром після цього 69 кандидатів, які відповідали як вимогам подання, так і критеріями мінімальної прийнятності, були прийняті в першому раунді процесу стандартизації [1].

Після багаторічного огляду кандидатів NIST вибрав 26 алгоритмів для переходу до 2-го раунду оцінки в січні 2019 року [2]. Ці алгоритми розглядалися як найбільш перспективні кандидати для можливої стандартизації і були обрані на основі як внутрішнього аналізу, так і відгуків спільноти. Після ретельного обговорення NIST вибрав сім фіналістів та вісім альтернативних варіантів, щоб перейти до 3-го раунду в липні 2020 року [3].

Після трьох раундів оцінки та аналізу, NIST вибрав перші алгоритми, які він стандартизує в результаті процесу стандартизації PQC, а також алгоритми, які ще необхідно дослідити – вони будуть досліджуватись у 4-му раунді [4].

Окрім конкурсу NIST, проводиться і дослідження на квантово-стійкі криптографічні алгоритми на національному рівні. Вже запропоновані та знаходяться у роботі стандарти постквантових алгоритмів АСШ, ПІК та ЕП.

Також для протистояння квантовим комп'ютерам і загрозам, які вони за собою несуть, розроблено та прийнято «Меморандум про національну безпеку з просування лідерства США в

галузі квантових обчислень при одночасному зниженні ризиків для вразливих криптографічних систем» від 4 травня 2022 року. У меморандумі викладено політику та ініціативи адміністрації президента США, пов'язані з квантовими обчисленнями. У ньому визначено ключові кроки, необхідні для збереження конкурентної переваги країни в галузі квантової інформаційної науки, а також зниження ризиків, пов'язаних з квантовими комп'ютерами для кібер-, економічної та національної безпеки країни. У ньому вказані конкретні дії, які мають зробити відомства, оскільки Сполучені Штати розпочинають багаторічний процес переведення вразливих комп'ютерних систем на квантово-стійку криптографію [5].

Вимоги NIST до постквантових криптоперетворень

Рівні безпеки NIST США, яким повинні відповідати обрані криптопримітиви:

1 рівень – Принаймні так важко зламати, як AES-128 (вичерпний перебір ключів).

2 рівень – Принаймні так важко зламати, як SHA-256 (пошук колізії).

3 рівень – Принаймні так важко зламати, як AES-192 (вичерпний перебір ключів).

4 рівень – Принаймні так важко зламати, як SHA-384 (пошук колізії).

5 рівень – Принаймні важко зламати, як AES-256 (вичерпний перебір ключів).

NIST попросив заявників зосередитись на рівнях 1, 2 та 3 (4 і 5 призначені для дуже високої безпеки).

Вибрані криптографічні примітиви повинні відповідати визначеній моделі безпеки:

- для АСШ – в умовах дії моделі безпеки IND-CCA2 (Indistinguishability adaptive ciphertext attack), що визначає нерозрізнюваність при атаці на основі адаптивно підбраного шифртексту;

- для ЕП – в умовах дії моделі безпеки EUF-CMA (Existentially unforgeable under adaptive chosen message attacks), що визначає екзистенційну невідтворюваність при атаці на основі вибраного повідомлення;

- для протоколу інкапсуляції ключів – в умовах дії моделі безпеки Canetti-Krawczyk (СК-безпека).

Відбір кандидатів NIST для стандартизації та подальших досліджень

Під час 3-го раунду було отримано деякі криптоаналітичні результати, які мали значний вплив на вибір NIST. Через це з розгляду було виключено ряд криптопримітивів. Так, наприклад, атака на GeMSS різко знизила його безпеку та алгоритм було виключено з розгляду для стандартизації. Алгоритм Rainbow також зазнав значних атак під час 3-го раунду. Перша атака спричинила втрату наборами параметрів від 20 до 55 біт безпеки в моделі RAM, причому набори параметрів з вищим рівнем безпеки втрачали більше бітів безпеки. Потім була ще атака, що призвела до відновлення особистого ключа. NIST також вирішив вилучити FrodoKEM, NTRU Prime та Picnic з розгляду для стандартизації [4].

FrodoKEM в основному вирізняється тим, що він не покладається на структуровані решітки (на відміну від Kyber, NTRU та Saber). У той час як NIST має намір вибрати принаймні один додатковий KEM, не заснований на структурованих решітках, для стандартизації після 4-го раунду, три інші альтернативи KEM (BIKE, HQC і SIKE) краще підходять для цієї ролі, ніж FrodoKEM. NTRU Prime також було висунуто як альтернативний варіант, оскільки він вважався менш перспективним порівняно з фіналістами. Схожа ситуація була і з підписами. Picnic не було обрано, оскільки NIST вирішив стандартизувати SPHINCS+.

Одним із важких виборів, з якими стикнувся NIST, було прийняття рішення між Kyber, NTRU та Saber. Решту обраних кандидатів KEM (BIKE, Classic McEliece, HQC, SIKE) продовжуватимуть оцінювати у 4-му раунді. І BIKE, і HQC засновані на структурованих кодах і будуть придатними як KEM загального призначення, що не ґрунтується на решітках.

NIST мав намір вибрати одного з Dilithium та Falcon, оскільки обидва базуються на структурованих решітках. Зрештою, NIST вирішив вибрати обидві схеми для стандартизації. Генерація ключа та підпису для Falcon, потребує більшої кількості ресурсів, ніж Dilithium, що може зробити Falcon непридатним для впровадження на обмежених пристроях, особливо у випадках, коли вимагається захист від атак бічними каналами. Крім того, NIST визнає, що простіша конструкція ключа та генерації підписів Dilithium допоможе забезпечити безпечні реалізації. З цих причин NIST вибрав Dilithium як основний алгоритм підпису, який він рекомендує для загального використання, і надасть пріоритет його стандартизації [4].

У табл. 1 та 2 наведено алгоритми, які слід стандартизувати та, які переходять до 4-го раунду, відповідно.

Табл.1 Алгоритми, які слід стандартизувати

Шифрування на відкритому ключі/KEM	Цифрові підписи
Crystals-Kyber	Crystals-Dilithium
	Falcon
	SPHINCS+

Табл.2 Кандидати, що переходять до четвертого раунду

Шифрування на відкритому ключі/KEM	Цифрові підписи
BIKE	
Classic McEliece	
HQC	
SIKE	

Квантово-захищені національні стандарти України

У той же час в Україні вже розроблені або знаходяться на стадії прийняття квантово-захищені криптографічні алгоритми. Існуючі стандарти: ДСТУ 7624-2014 (Калина) – блокове симетричне шифрування, ДСТУ 7564-2014 (Купина) – функція гешування, ДСТУ 8845-2019 (Струмок) – потокове шифрування, ДСТУ 8961-2019 (Скеля) – асиметричне шифрування та інкапсуляція ключів. Проекти стандартів (знаходяться у розробці): ДСТУ (Вершина) та ДСТУ (Сокіл) – електронний підпис.

Важливим загальним результатом наших досліджень та розробки є те, що прийняті наші рішення щодо вибору математичних основ побудови перспективних стандартів ЕП “Вершина” та “Сокіл” повністю співпали з подальшими рішеннями NIST США по закінченню 3-го раунду конкурсу на перспективні стандартні алгоритми електронного підпису (ЕП), які були опубліковані раніше, ніж пропозиції NIST США. Проект стандарту ЕП “Вершина” використовує математичні засади Dilithium, а “Сокіл” – математичні засади Falcon.

Алгоритм ЕП “Вершина” може працювати у наступних режимах:

- Вершина-128/64 (нульовий) – 128 біт захисту від класичних атак та 64 біт від квантових атак, захист від спеціальних атак (запас стійкості ДСТУ 7624:2014 (128), AES (128));
- Вершина-256/128 (перший) – 256 біт захисту від класичних атак та 128 біт від квантових атак, захист від спеціальних атак (запас стійкості ДСТУ 7624:2014 (256), AES (256));
- Вершина-384/192 (другий) – 384 біт захисту від класичних атак та 192 біт від квантових атак, захист від спеціальних атак (запас стійкості ДСТУ 7624:2014 (512));
- Вершина-512/256 (третій) – 512 біт захисту від класичних атак та 256 біт від квантових атак, захист від спеціальних атак (запас стійкості ДСТУ 7624:2014 (512)).
- Вершина-с – для сумісності з режимом Dilithium (128, 256 біт).

У залежності від необхідного рівня безпеки та обмежень на техніко-експлуатаційні та техніко-економічні характеристики ЕП «Сокіл» може застосовуватись із забезпеченням 1, 2 та 3 рівнів безпеки та криптоживучості відповідно із забезпеченням 128, 256 та 512 біт безпеки проти класичного криптоаналізу та 64, 128 та 256 біт безпеки проти квантового криптоаналізу, а також із забезпеченням захисту від спеціальних атак.

Теоретична безпека ЕП «Сокіл» підтверджується доказом у моделі QROM на основі складності SIS над NTRU решітками.

Консервативні оцінки складності підробки ЕП «Сокіл» мають такі ж самі значення, як і для ЕП “Вершина”.

Висновки

1. Основними алгоритмами, рекомендованими NIST для більшості випадків використання, є Crystals-Kyber (встановлення ключа) і Crystals-Dilithium (цифрові підписи). Крім того, схеми підпису Falcon і SPHINCS+ також будуть стандартизовані. Кандидати BIKE, Classic McEliece, HQC і SIKE будуть продовжені для подальшого вивчення в 4-му раунді оцінювання.

2. 4-й раунд оцінювання та аналізу відбуватиметься подібно до попередніх раундів. Після завершення 4-го раунду NIST може вирішити вибрати деяких із кандидатів 4-го раунду для стандартизації.

3. В Україні також на національному рівні вже розроблені або знаходяться на стадії прийняття квантово-захищені криптографічні алгоритми. Серед вже розроблених такі стандарти: ДСТУ 7624-2014 (Калина) – блокове симетричне шифрування, ДСТУ 7564-2014 (Купина) – функція гешування, ДСТУ 8845-2019 (Струмок) – потокове шифрування, ДСТУ 8961-2019 (Склея) – асиметричне шифрування та інкапсуляція ключів. Серед проектів стандартів: ДСТУ (Вершина) – електронний підпис, ДСТУ (Сокіл) – електронний підпис.

ЛІТЕРАТУРА

1. National Institute of Standards and Technology (2016) NIST post-quantum cryptography standardization. URL: <https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization> (дата звернення: 26.10.2022).
2. Alagic G. Status report on the first round of the NIST post-quantum cryptography standardization process / Alagic G., Alperin-Sheriff J., Apon D., Cooper D., Dang Q., Liu Y. K., Miller C., Moody D., Peralta R., Perlner R., Robinson A., Smith-Tone D. // (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) 8240. (2019). URL: <https://doi.org/10.6028/NIST.IR.8240> (дата звернення: 26.10.2022).
3. Alagic G. Status report on the second round of the NIST post-quantum cryptography standardization process / Alagic G., Alperin-Sheriff J., Apon D., Cooper D., Dang Q., Kelsey J., Liu Y. K., Miller C., Moody D., Peralta R., Perlner R., Robinson A., Smith-Tone D. // (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) 8309. (2020). URL: <https://doi.org/10.6028/NIST.IR.8309> (дата звернення: 26.10.2022).
4. Gorjan Alagic NIST IR 8413 Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process / Gorjan Alagic, Daniel Apon, David Cooper, Quynh Dang, Thinh Dang, John Kelsey, Jacob Lichtinger, Yi-Kai, Liu Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone // URL: <https://doi.org/10.6028/NIST.IR.8413> (дата звернення: 26.10.2022).
5. National Security Memorandum on Promoting United States Leadership in Quantum Computing While Mitigating Risks to Vulnerable Cryptographic Systems. URL: <https://www.whitehouse.gov/briefing-room/statements-releases/2022/05/04/national-security-memorandum-on-promoting-united-states-leadership-in-quantum-computing-while-mitigating-risks-to-vulnerable-cryptographic-systems/> (дата звернення: 26.10.2022).

ЄСІНА Марина Віталіївна – к. т. н.; доцент кафедри безпеки інформаційних систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: m.v.yesina@karazin.ua; ORCID: 0000-0002-1252-7606.

Наукові інтереси:

– безпека інформації, постквантові криптографічні примітиви.

ГОРБЕНКО Іван Дмитрович – д. т. н., професор; професор кафедри безпеки інформаційних систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: i.d.gorbenko@karazin.ua; ORCID: 0000-0003-4616-3449.

Наукові інтереси:

– постквантова криптографія, механізми захисту інформації.

КАЧКО Олена Григорівна – к. т. н., професор; професор кафедри програмної інженерії; Харківський національний університет радіоелектроніки, проспект Науки, 14, Харків-66, Україна, 61166; e-mail: ekachko@gmail.com; ORCID: 0000-0001-9249-0497.

Наукові інтереси:

– програмні методи захисту інформації, програмне моделювання криптографічних примітивів.

ЗАЦ О.Д., ШМАТКОВ С.І.

МОДЕЛЮВАННЯ ТРАФІКУ КОМП'ЮТЕРНИХ МЕРЕЖ

Мета роботи: дослідження методів моделювання трафіку мереж передачі даних.

Методи дослідження: у роботі розглянуті методи моделювання трафіку на основі Пуассонівського процесу, на основі марківських процесів, авторегресійні моделі, самоподібні моделі. Акцентована увага на методи моделювання трафіку на основі фрактальних моделей, зокрема фрактального точкового процесу.

Результати: проведений огляд моделей і методів, які застосовуються при моделюванні процесу передачі даних у комп'ютерних мережах. Розглянутий алгоритм моделювання самоподібного трафіку.

Висновки: у роботі проаналізовані існуючі моделі трафіку комп'ютерних мереж. Трафік сучасних мереж передачі даних проявляє ефект самоподібності. Самоподібні моделі трафіку дозволяють отримувати репрезентативні характеристики продуктивності мереж за різного характеру даних (наприклад, комп'ютерні дані, цифровий відео потік тощо) і обґрунтовано обирати параметри мережевого обладнання.

Ключові слова: комп'ютерні мережі, моделювання трафіку, самоподібні моделі, фрактальні моделі.

Швидкий і стрімкий розвиток мережевих і комунікаційних технологій, а також підвищення рівня доступності користування цими технологіями призводить до постійного зростання об'єму мережевого трафіку. Щорічно трафік у середньому зростає на 30-40%. При цьому більше 60% трафік складають мультимедійні застосунки потокового аудіо і відео. І саме такий трафік є більш чутливим до затримок у передачі пакетів даних. Таким чином, до мереж передачі даних постійно зростають вимоги до забезпечення якості зв'язку, рівню продуктивності і пропускної спроможності, якості управління трафіком, боротьби з перенавантаженням тощо.

Для того, щоб задовольнити вимоги до роботи мережі передачі даних необхідно визначити раціональний варіант її структури і параметрів. Пошук такого варіанта мережі є складною задачею, пов'язаною з вибором великої кількості різнорідних параметрів мережі – типів комутаційного обладнання, серверів, типів операційних систем та їх параметрів, стеків комутаційних протоколів та ін.

На сьогодні ефективне управління існуючою мережею і створення нової неможливо без детального попереднього моделювання її роботи.

Метою моделювання при цьому є визначення оптимальної топології, адекватний вибір мережевого обладнання, визначення робочих характеристик мережі тощо.

Існує два підходи до моделювання мереж та їх характеристик: імітаційний і аналітичний. За імітаційного моделювання використовують або готові спеціалізовані програмні системи, в яких процес створення моделі спрощений, або створюються програми за допомогою універсальних мов програмування. Системи імітаційного моделювання зазвичай містять набір засобів для підготовки вихідних даних про досліджувану мережу (її топологію і трафік), засоби для статистичної обробки отриманих результатів моделювання. Якість результатів моделювання значно залежить від точності вихідних даних про мережу, які передаються до системи імітаційного моделювання.

Визначити ж характеристики мережі ще на стадії проектування можна за допомогою аналітичного моделювання. При побудові аналітичних моделей мережа і всі процеси, які в ній відбуваються, подаються у вигляді сукупності математичних співвідношень.

Однією з основних характеристик мереж передачі даних є трафік.

У найпростішому випадку трафік можна подати як процес надходження дискретних даних (пакетів, повідомлень, одиничних сигналів тощо) і математично описати як послідовність даних, які надходять, $X_1, X_2, \dots, X_n, \dots$, де $X_0 = 0$. У такому випадку послідовністю даних може бути кількість даних, які надходять, і час між їхніми надходженнями.

Змішаний трафік буде містити більше одного елементу у послідовності даних X_n , яка надходить. Для опису змішаного трафіку використовується невід'ємна випадкова послідовність $B_1, B_2, \dots, B_n$, де $n = 1, 2, \dots, \infty$, B_n – кількість елементів у послідовності даних, які надходять.

Для моделювання трафіку використовуються різні підходи. Моделі трафіку можна класифікувати за характером процесу надходження даних і програмному забезпеченню або

додатку, який здійснює передачу даних. За характером процесу моделі може бути стаціонарними й нестаціонарними. Стаціонарні моделі, у свою чергу, можуть володіти короткостроковою або довгостроковою залежністю. До моделей із короткостроковою залежністю відносять класичні регресійні моделі і моделі, які основані на марківських процесах. Довгостроковою залежністю відрізняються фрактальні моделі. Також за додатком-джерелом трафік може бути класифікований як трафік web, peer-to-peer, потокового відео тощо. На рис. 1 показана класифікація моделей мережевого трафіку.

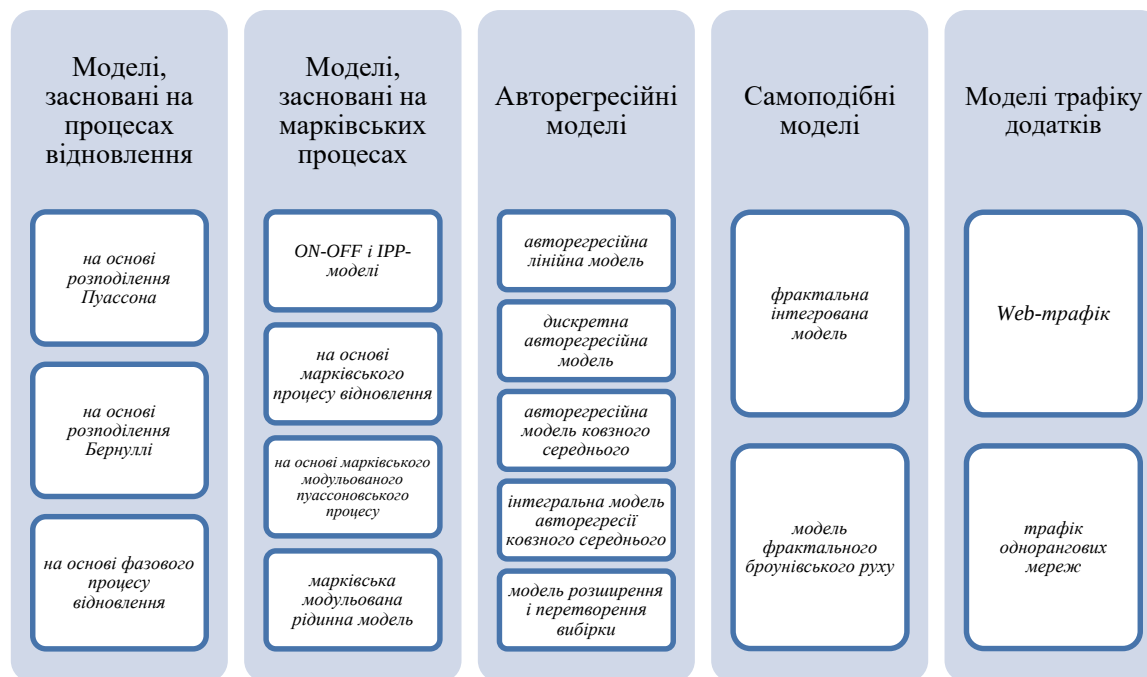


Рис. 1 Класифікація моделей трафіку

Класичними моделями трафіку вважають моделі на основі розподілення Пуассона та марківські моделі. Але трафік сучасних мереж передачі даних проявляє ефект самоподібності.

Ця властивість проявляється на практиці у тому, що статистичні характеристики трафіку «масштабуються» при їх усередненні на різних проміжках часу. Фізично це пояснюється ефектами групування трафіку: пакети, які належать одному ж і тому набору даних більшої протяжності надходять на вузол комутації не випадковим чином, а цілою пачкою. Це явище значно погіршує характеристики при проходженні самоподібного трафіку мережею, навіть тоді, коли середня інтенсивність трафіку набагато нижча максимальної швидкості передачі у даному каналі, оскільки призводить до втрат пакетів через обмеженість буферу, і, відповідно, збільшуються затримки у мережі.

Таким чином більш доречним для моделювання мережевого трафіку є використання самоподібних моделей. Перший тип моделей трафіку – це *фрактальні моделі* процесів, які оперують потоками пакетів і безперервним часом.

Фрактальний точковий процес або FSNDP (Fractal Shot Noise Driven Poisson process) процес відноситься до класу подвійних стохастичних процесів. Змінна у часі інтенсивність цього процесу задається безперервним випадковим процесом – фрактальним дробовим шумом, який отримується шляхом фільтрування класичного пуассоновського процесу.

Процес FSNDP повністю визначається набором параметрів $\langle \mu, \beta, K, A, B \rangle$.

Первісний пуассоновський потік $\{t_i\}$ з постійною інтенсивністю μ є входом для лінійного фільтру з імпульсною функцією

$$h(t) = \begin{cases} Kt^{-\beta}, & t \in (A, B), \\ 0, & t \notin (A, B), \end{cases}$$

де β визначається степенем самоподібності процесу, A, B – невід'ємні обмежувальні параметри, K – додатна константа, яка визначає амплітуду результуючого процесу. Фільтр викликає фрактальний дробовий шум:

$$I(t) = \sum_{i=-\infty}^{\infty} h(t - t_i),$$

розглядається як змінна інтенсивність для другого пуассоновського точкового процесу, виходом якого є потік FSNDP.

Параметр β визначається ступенем самоподібності процесу і пов'язаний з параметром Херста H :

$$H = \frac{3}{2} - \beta.$$

Процес має стійкі самоподібні властивості за умови $A \ll B$ і A , достатньо близьким до нуля. На практиці зазвичай приймають $A = 0 \dots 0,1$ і $B < 10$.

Другий тип моделей засновується на використанні фрактального гаусівського шуму або фрактального броунівського руху. Фрактальний броунівський рух $\{fB_t\}$ – це самоподібний гаусівський процес з параметром самоподібності $0.5 \leq H < 1$.

Висновки. Використання саме самоподібних моделей трафіку дозволить отримувати репрезентативні характеристики продуктивності мереж за різного характеру даних (наприклад, комп'ютерні дані, цифровий відео потік тощо) і обґрунтовано обирати параметри мережевого обладнання.

ЛІТЕРАТУРА

1. Водопьянов С.В. Применение моделей трафика данных для мониторинга компьютерных сетей системы управления воздушным движением. *Проблеми інформатизації та управління*. 2012. № 1 (37). С. 30-35.
2. Карпукhin А.В., Кириченко Л.О., Гривцив Д.И., Ткаченко А.А. Применение методов нелинейной динамики и фрактального анализа для оценивания работы инфокоммуникационных систем с протоколом TCP. *Cloud of Science*. 2014. № 2. С. 258-271.
3. Шыхалиев Р.Г. Анализ и классификация сетевого трафика компьютерных сетей. *Informasiya texnologiyaları problemləri*. 2010. № 2. С. 15-23.
4. Карпукhin А.В., Кириченко Л.О., Радивилова Т.А.. Самоподобие в информационных системах. *Прикладная радиоэлектроника*. 2008. № 1. С. 54-64.
5. Ryu Bo, Steven Lowen. Modeling, Analysis and Simulation of Self-Similar traffic using the Fractal-Shot-Noise-Driven-Poisson Process. 1995.

ЗАЦ Олександр Дмитрович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: zats2021ki51@student.karazin.ua; ORCID: 0000-0002-7623-9187.

Наукові інтереси:

- моделювання складних процесів;
- моделювання комп'ютерних мереж.

ШМАТКОВ Сергій Ігорович – д.т.н., професор, завідувач кафедри теоретичної та прикладної системотехніки, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: s.shmatkov@karazin.ua; ORCID: 0000-0002-6328-988X.

Наукові інтереси:

- моделювання складних процесів;
- системний аналіз;
- методи штучного інтелекту.

АНАЛІЗ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО УПРАВЛІННЯ КОМП'ЮТЕРНИМИ ДАНИМИ

An overview of intelligent analysis methods for further usage in computer data management process. Advantages and disadvantages of their general application are considered. The correctness of their use and efficiency in matters of processing sets of network parameter values were analyzed in the course of further research on the transmission of information packets and their management in the network.

Ключові слова: нейронні мережі і мунні мережі, експертні системи, кластерний аналіз, комп'ютерна мережа.

Вступ

Найбільш поширеними у використанні на сьогоднішній день вважаються алгоритми адаптивної або ж динамічної маршрутизації. “Динамічна маршрутизація” – термін, який в своєму розумінні характеризує мережу, як гнучку та надійну. Інформація про топологію зв'язків в мережі, яка збирається за допомогою протоколів адаптивних алгоритмів, є достатньою для автоматичного оновлення таблиць маршрутизації, після змін конфігурації мережі, та оперативного реагування на всі ці зміни^[1].

Проте швидкий розвиток мереж та їх глобальне об'єднання, в тому числі і в мережу Інтернет, призводить до появи нових атак та викликаних ними аномалій в роботі. Відповідно до цього, ефективність виявлення невідомих, навіть мінімально шкідливих порушень інформаційної безпеки неминуче зменшується, що знижує і оперативне втручання в перебудові маршруту для надійної передачі комп'ютерних даних. Це може призвести до перевантаження системи, підвищення використання ресурсів мережеских пристроїв, виведення їх з ладу та негативного опосередкованого або ж прямого впливу на користувачів.

Використання методів інтелектуального управління даними у роботі комп'ютерних мереж застосовується для покращення ефективності аналізу стану мереж та своєчасного перебудови маршруту передачі інформаційних пакетів даних. Таке поєднання зі звичайними динамічними алгоритмами дозволяє працювати з великими об'ємами даних, що є актуальним при опрацюванні обліку мережеского трафіку. Це призводить до вдосконалення інструментальних засобів та можливості до адаптованого навчання, та може забезпечити неперервну підтримку стабільної та надійної роботи мереж. Загальна обробка отриманої інформації у своїй сукупності сприяє мінімізації виникнення можливих несправностей або невідповідностей процесу передачі інформації.

Методи інтелектуального управління та обробки даних

Кластерний аналіз. Загальна задача цієї групи статистичних процедур – розбиття вибірки необхідних даних на підмножини, що не перетинаються, на основі подібності змінних, що вивчаються^[2]. Серед них виділяються кластери нормальної поведінки. Кожен кластер має складатися зі схожих об'єктів, а об'єкти різних кластерів мають істотно відрізнятися один від одного.

Серед переваг цього методу можна виділити, що задачі кластеризації відносяться до широкого класу задач навчання без вчителя. Серед свого арсеналу методи кластерного аналізу мають досить популярні методи k-середніх та ієрархічного аналізу. Вони відзначаються своєю простотою та швидкістю обробки інформації. Хоча і не існує найкращого критерію якості кластеризації, відомий цілий ряд досить ефективних критеріїв, а також ряд алгоритмів, які не мають чітко вираженого критерію, але все одно здійснюють досить якісну кластеризацію по побудові.

У даного підходу також є ряд недоліків. Перш за все, кількість кластерів зазвичай заздалегідь невідома, що буде значним недоліком в процесі аналізу мережеского трафіку та поведінки самої мережі, враховуючи їх динамічний характер. В такому випадку, використання, наприклад, методу k-середніх стає неможливим, оскільки кількість кластерів має бути визначена спеціалістом заздалегідь (k і є позначенням цього числа).

Методи ієрархічного аналізу не можуть використовуватись у випадку управління комп'ютерними даними через велику кількість наданої інформації. Дендограми, у вигляді яких і представлені кластери, мають велику вірогідність перевантаження та мають тенденцію втрачати наочність. З цією проблемою чудово справляється метод k-середніх, проте чутливість до можливих відхилень у вибірці мережевих даних може критично викривити середнє значення нормального стану мережі.

Алгоритм для аналізу також частіше всього необхідно вибирати експериментально, в залежності від конкретної проблеми та у випадках, коли немає математичної причини віддати перевагу одному алгоритму кластеризації над іншим. Через таку особливість ефективного використання кластерних алгоритмів у випадку аналізу набору мережевих характеристик також підпадає під сумнів: один алгоритм може добре працювати на певному наборі даних, але не працювати для іншого.

Нейронні мережі. В моделюванні використовується система поєднаних та взаємодіючих між собою вузлів - штучних нейронів^[3]. Ця модель називається штучною нейронною мережею та може бути математичною або комп'ютерною.

Внутрішній або реагуючий нейрон має безліч вхідних зв'язків, до яких періодично надходять сигнали від рецепторів або інших внутрішніх нейронів, а на виході - сигнал, що посиляється після опрацювання поточним нейроном отриманих даних. Проте їх об'єднання у мережу здатне виконувати досить складні завдання.

Найпростіший варіант такого поєднання - використання звичайного персептрона з декількома прихованими шарами. Для виявлення аномалій в системі, ШНМ навчаються протягом деякого періоду часу, коли вся спостережувана поведінка вважається нормальною та сприятиме подальшому успішному прогнозу поведінки даних. На входи нейронної мережі зазвичай подається визначений експертом набір параметрів, на основі якого відбуватиметься аналіз. Виходом звичайно є прогноз мережі на майбутній момент часу.

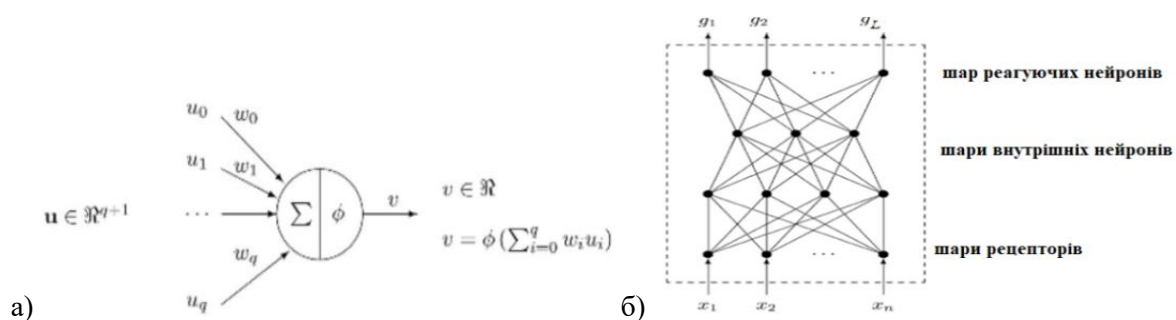


Рис. 1 - а) модель внутрішнього нейрону та б) схема багатошарового персептрону

Можливість навчання - одна з головних переваг використання ШНМ. Навчання полягає в знаходженні коефіцієнтів зв'язків між нейронами. В процесі цього навчання ШНМ виявляє складні залежності між вхідною і вихідною інформацією та здатна виконувати узагальнення. Спроможність розпізнавання "зашумлених" та частково спотворених даних, що не є рідкістю в роботі з комп'ютерними мережами, зменшує ймовірність повернення невірного результату на підставі даних, яких не було в навчальній вибірці. Це також приводить нас до можливості ШНМ адаптуватись до змін її навколишнього середовища. Висока ступінь паралелізму обробки отриманої інформації та низький рівень ресурсоємності гарантують отримання рішення поставленої для системи задачі за прийнятно короткий час. Ці особливості важливі у створенні алгоритму адаптивної маршрутизації та можуть забезпечити надійне функціонування системи, своєчасні виявлення різних класів аномалій та запобігання їх наслідків.

З іншого боку ШНМ володіють рядом недоліків, серед них особливо виділяється їх недетермінованість. Після навчання мережа перетворюється в "чорний ящик", а логіка прийняття рішень нейромережею зовсім схована від експерта. Існуючі алгоритми дають можливість "витягнути" принципи розв'язку задач та перетворити їх на правила для створення експертної системи^[5]. Але такі алгоритми не можна вбудувати до нейромережевих пакетів, а згенеровані дані можуть бути досить об'ємними.

Експертні системи. Експертні системи (ЕС) - складні програмні засоби, що використовують експертні знання конкретної предметної галузі для забезпечення високоефективного пошуку рекомендацій та прийняття рішень^[5]. В основі ЕС лежать принципи технологій штучного інтелекту. Ці методи дозволяють моделювати механізми мислення людини, формуючи певні висновки, а пошук рішень спирається на базу знань (БЗ), що накопичується у процесі побудови та експлуатації системи. Загальна структура механізму ЕС містить в собі такі структури як БЗ та механізм виведення (МВ). Вони в свою чергу також складаються з декількох компонентів: фактів і правил, диспетчера та інтерпретатора відповідно. Диспетчер встановлює послідовність застосування правил нормальної поведінки системи. Інтерпретатор визначає як саме слід застосувати елементи БЗ в поведінці, що спостережується. БЗ та МВ є специфічними елементами програмного забезпечення з використанням ЕС. Але їх робота з майбутнім користувачем не можлива без окремого інтерфейсу.

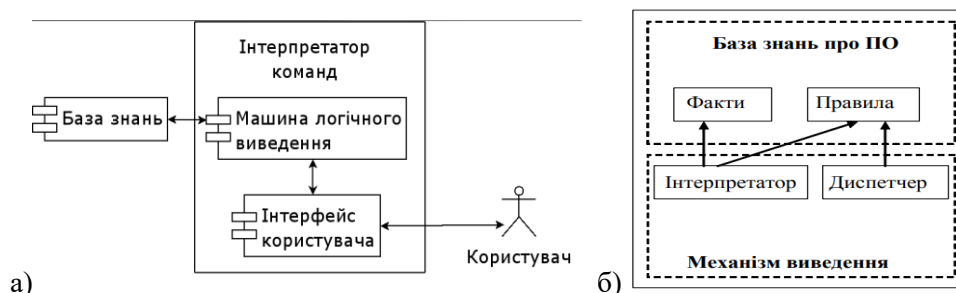


Рис. 2 - а) загальна структура експертної системи та б) схема взаємодії між БЗ та МВ

Достовірність факту або точність правила не завжди є абсолютними та викликають деяку міру невпевненості. Для того, щоб уникнути сумнівів, використовується коефіцієнт довіри. Він означає вірогідність або ступінь впевненості, з якою можна рахувати даний факт або правило справедливими. Це підвищує надійність системи.

Основною особливістю ЕС є накопичення та організація знань. Інформація, що знаходиться в БЗ, зберігається протягом невизначеного довгого часу, на відміну від експерту-людини, у зв'язку зі звичайним людським фактором. За схожою причиною, можна відмітити стійкість і відтворюваність результатів. Це приводить нас до ще однієї особливості ЕС - здатність давати пояснення, щодо запропонованого рішення, і доводити його ґрунтовність, що характеризує ЕС як прозору систему. Можливість одержання й об'єднання експертних знань з багатьох джерел виключає залучення декількох експертів для прийняття рішень та за рахунок цього гарантує низьку вартість експлуатації.

Проте ймовірна невідповідність та уточнення спостережуваної поведінки відносно правил, що викликати протиріччя в рішенні, виключають повну відмову від людини-експерта. Це стосується і побудови та необхідності переопрацювання ЕС через вимогу володіння експертними знаннями кінцевим користувачем. Ще однією проблемою використання ЕС є їх недосконале подання знань про часові та просторові відношення. Також з виявленням нових знань, можливість системи до прийняття рішення не завжди збільшується, адже може викликати впровадження нових правил або перебудови основної задачі для відслідковування появи аномалій.

Імунні мережі. Методологія, створення якої було базоване на принципах біологічних імунних систем, використовується в різноманітних областях досліджень. Штучні імунні мережі (ШІМ) можна визначити як обчислювальну парадигму, що спирається на теорії імунології, спостерігаються імунними функціями та механізмами^[4].

Загальний принцип роботи методів ШІМ полягає в наступному. На вході в систему маємо вибірку даних для розпізнавання S . Створюється випадковий набір "антитіл" A . Для кожного зразка S визначається міра афінності (схожості, що визначається за допомогою математичних відстаней) з конкретним антитілом вибірки A . Антитіла з найвищою афінністю створюють пропорційне значенню схожості число клонів. Кожен клон піддається процесу зміни атрибутів, тобто мутування, ступінь чого залежить від цільової функції. Після цього весь набір клонів одного антитіла порівнюється батьківським. На виході маємо набір M , що буде здатен класифікувати нові для системи зразки.

Основна перевага цього набору методів – вирішення різноманітної кількості задач, що пов'язані з аналізом даних. Методи ШІМ, як і ШІМ^[3], мають ряд характеристик: здатність до навчання, розпізнавання та ухвалення рішень, щодо “чужорідного елементу”, адаптованість. Накопичення нової та використання попередньо вивченої інформації дають змогу знаходити аномалії, розпізнавати образи і сигнали, моделювати системи пошуку і оптимізації, а також системи комп'ютерної безпеки.

Проте застосування імунних мереж також має і свої недоліки. Кількість прикладів нормальної поведінки мережі дуже часто перевищує число прикладів несправної поведінки мережі. Через це процес виявлення мережових аномалій має досить велику обчислювальну складність, що є недопустимим в ході досягнення відповідності критерію швидкого реагування на збій або некоректність роботи.

Висновки

У роботі були розглянуті декілька основних методів інтелектуального управління даними, що використовуються під час побудови алгоритмів аналізу стану комп'ютерних мереж з адаптивною маршрутизацією. Під час аналізу принципів згаданих в роботі методологій було розглянуто ряд їх переваг та недоліків.

Аналіз стану мережі та набору даних мережевого трафіку потребує застосування найбільш ефективних математичних інструментів, що можуть впоратись з великою кількістю даних за оптимально швидкий час та матимуть мінімальну ймовірність викривлення значень нормального стану. Ще одним з важливих критеріїв є невелика складність реалізації, що забезпечить легке впровадження алгоритму майже до будь-якої системи.

Методи засновані на використанні імунних мереж, хоч і являються гарним варіантом для аналізу мережевого трафіку, але не є раціональними саме в силу високої обчислювальної та реалізаційної складності. Методи кластерного аналізу не вимагають дуже великого обсягу часу та досить прості в процесі додавання до систем, але мають велику ймовірність спотворення результатів. Проте враховуючи більш статичний характер часових параметрів мережових засобів, їх використання є прийнятним. Єдиною можливою проблемою в даному випадку може стати тільки підключення нових технічних інструментів. Експертні системи є прийнятним варіантом для поставленої задачі, але в питаннях опрацювання мережових даних їх використання не є доречним, т.я. значна частина сукупності параметрів, відносно яких відбувається аналіз мережі та даних, є часовими.

Таким чином оптимальними варіантами у роботі з великою кількістю даних мережевого трафіку для аналізу виявились саме нейромережові методи. Навіть врахувавши їх недоліки, переваги цієї методології виявились прерогативними.

ЛІТЕРАТУРА

1. Лосев Ю. І., Руккас К. М., Шматков С. І. Комп'ютерні мережі: навч. посіб. / за редакцією Ю. І. Лосева. – Х. : ХНУ імені В. Н. Каразіна, 2013. – 248 с.
2. Czamanski S. and Ablas L. A. Identification of industrial clusters and complexes: a comparison of methods and findings // Urban Studies. –1979. – V. 16. – P. 61-80.
3. Кузсуль Н.Н., Сидоренко А.В., Скакун С.В. Нейросетевая модель пользователей компьютерных систем // Кибернетика и вычислительная техника. – 2004.
4. Castro L., Timmis J. An Artificial Immune Network for Multimodal Function Optimization. Proceedings of IEEE Congress on Evolutionary Computation (CEC'02). 2002. Vol. 1. P. 699–674. DOI: 10.1109/CEC.2002.1007011
5. Попов Э.В. Экспертные системы реального времени//Открытые системы, 2(10), 1995

ЗЕЛЕНСЬКА Ніка Вікторівна – магістр, аспірант кафедри теоретичної та прикладної системотехніки, Харківський національний університет імені В. Н. Каразіна, майдан Свободи 6, Україна, 61022; e-mail: zelenskaya.nika@gmail.com; ORCID: 0000-0002-9289-5061.

Наукові інтереси:

- комп'ютерні мережі;
- методи моніторингу даних.

МОДЕЛЬ УПРАВЛІННЯ РИЗИКАМИ ПРИ РЕАЛІЗАЦІЇ ІТ-ПРОЕКТУ

Мета: виконати аналіз предметної області та оцінити комп'ютерні модель програми управління ризиками при реалізації ІТ-проекту

Методи дослідження: оцінка спроможності ІТ-проекту через дерева рішень та теорію ігор

Результати: у ході аналізу комп'ютерних моделей за допомогою дерева рішень та теорії ігор оцінити життєздатність ІТ-проекту

Висновки: у результаті аналізу комп'ютерних моделей ІТ-проектів прийшли до висновку, що перевірка життєздатності проекту через теорію ігор та дерева рішень є оптимальною.

Ключові слова: модель, ризики, дерево рішень, дерево ризиків, теорія ігор

Вступ

Ризик проекту – це невизначений випадок чи умова, які впливають на досягнення цілі проекту(вплив може бути як позитивний, як і негативний). Ризики з'являються як результат одної чи декількох причин, та впливаючи на один чи декілька аспектів проекту. Причиною виникнення ризика чи ризиків в ІТ-проекті можуть бути додаткова вимога від замовника(як суттєве так і незначне), нехватка бюджету, не передбачені помилки у програмному продукті, тощо. Будь-яка з цих проблем може вплинути на проект, його зміст, вартість, розклад, якість чи виконання. Також ризик чи ризики можуть виникнути через деяких аспектів середовища організації чи проекту, що . Як приклад можна привести віддачу проекту на аутсорс. Проект, який перебуває на аутсорсі, неможливо контролювати безпосередньо.

Управління ризиками ІТ-проекту включає в себе процеси, які пов'язані із здійсненням планування управління ризиками, ідентифікацією, аналізом, плануванням реакції, а також із контролем ризиків у проекті. Метою управління ризиками є збільшення сприятливих подій проекту та уникнення або зменшення ймовірності негативних подій у ході виконання проекту.

У будь-якому ІТ-проекті завжди бувають ризики, маленькі чи великі. Не існує проекту, у якого не буває ризиків, тому у ІТ-проектах управління ризиками займає виключне положення.

Реалізація проекту завжди буде проходити у певній невизначеності, тобто ризику, тому що:

- неповне знання усіх параметрів проекту, а також обставин, ситуацій, необхідних для оптимального рішення, неможливість знання усіх деталей проекту.
 - фактори випадковості, тобто факторами, що неможливо передбачити або калькулювати навіть як ймовірність.
 - суб'єктивні фактори протидії, коли партнери мають протилежні погляди один з іншим.
- Таким чином, ІТ-проекти завжди існують у невизначеності.

Аналіз моделей за допомогою дерева рішень та теорії ігор

На сьогоднішній день інформаційні технології все стрімкіше розвиваються, та займають все більше місця у повсякденному житті людини, розповсюджуючись все ширше і ширше. Ці інформаційні технології називають ІТ-проектами. ІТ-проектом можна назвати діяльність, що зв'язана з використанням чи застосуванням інформаційних технологій.

Реалізація ІТ- проекту являється складним процесом, у якому використовується багата кількість технологій, апаратних засобів та спеціалістів. Як результат, керування проектом складає собою маневрування між ризиками, зменшення їх впливу на весь проект та адаптація к ризикам, якщо вони неминучі.

Нашою ціллю є створення оптимальної моделі управління ризиками при реалізації ІТ-проекту. Протягом цієї роботи ми оцінимо задачі при управлінні ризиками ІТ-проекту. Ми модулюємо процес управління ризиками ІТ-проекту за допомогою таких методів як теорія ігор та дерево рішень.

Дерево рішень(рис 2) представляє собою модель, яка дозволяє велику та складну проблему прийняття рішень в умовах ризику на сукупність менших проблем які можуть бути розглянуті окремо, а потім разом.

На дереві рішень кожна вітка, яка представляє собою або існуючий варіант дій, або можливі наслідки обраної дії, ділиться у визначених точках на сукупність інших віток. Обробляючи модель ІТ-проекту через дерево рішень, ми встановлюємо ту чи іншу вірогідності зв'язані з ризиками, які можуть виникнути у проекті. Після цього вітка можливого варіанту дії, проходячи через виникнення наслідків, розбивається на декілька результатів дій, які також можуть мати точки прийняття рішень.

Дерево ризиків(рис 1) використовується для проведення корисного аналізу ризиків. Дерево ризиків проекту забезпечує виконання повноцінного процесу систематичної ідентифікації ризиків в залежності від рівня деталізації та зв'язку з іншими елементами проекту. Завдяки цьому дереву ми ідентифікуємо, які ризики можуть виникнути при реалізації ІТ-проекту і ідентифікувати їх.

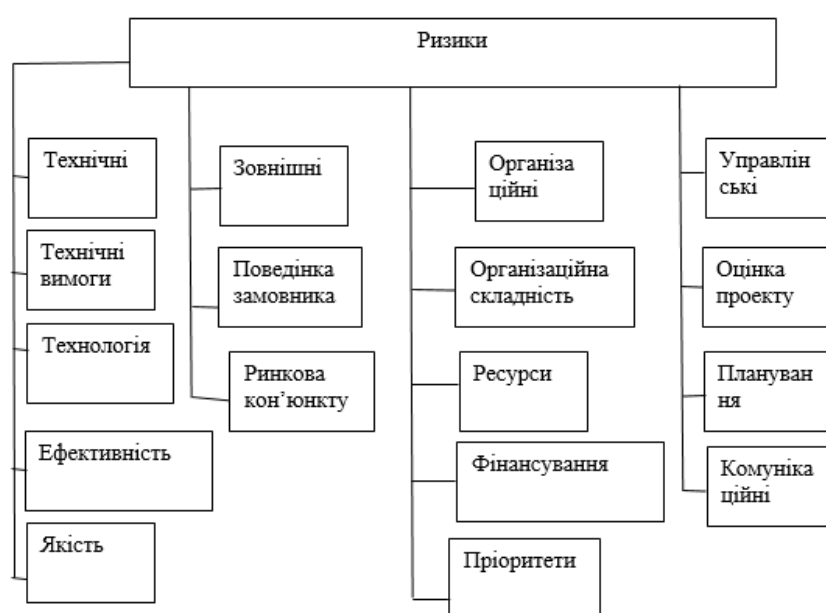


Рис 1. Дерево ризиків ІТ-проекту

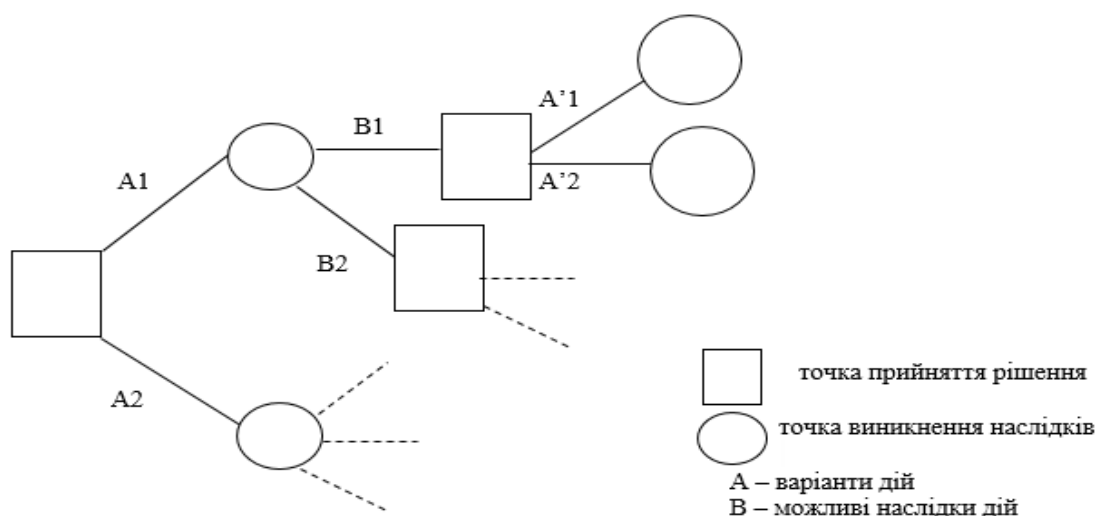


Рис 2. Загальний вид дерева рішень

Визначити стратегії рішень в умовах ризиків можна за допомоги методів теорії ігор. Теорія ігор вивчає конфлікти двох чи більше сторін, які називають іграми. Під вивчення попадають самі ігри, стратегії, які застосовують в іграх, а також моделі поведінки в іграх. Здійснюючи деякі(1, табл 1) з цих методів у програмному забезпеченні, ми можемо за допомогою них виробляти певні стратегії для моніторингу та управління ризиками ІТ-проекту.

$$(1) Y_{\text{гар}} = \max_i \min_j a_{ij}$$

Таблиця 1. Критерій Вальда

Проект	Ефективність у випадках		$\min_i a_{ij}$
	S1	S2	
C1	30	60	30
C2	50	20	20
$\max_i \min_j a_{ij}$			30

При застосуванні критерія Вальда, ми орієнтуємося на найгірші обставини, тому результат може бути гарантованим, чи краще гарантованого.

ЛІТЕРАТУРА

1. Основи проектного управління : підручник Разу М.Л : (2006). – 768 с.
2. Моделі й методи прийняття рішень: навч. посіб. / С.А. Ус, Л.С. Коряшкіна. – Д.: НГУ, 2014. – 300 с.

ІСАЄВ Владислав Сергійович – студент кафедри теоретичної та прикладної схемотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: isaevvladislav06@gmail.com, xa12283974@student.karazin.ua; ORCID :0000-0003-3759-2700

Науковий інтерес:

- Управління проектами

КАНДІЙ С.О.

ПОРІВНЯННЯ АРГУМЕНТІВ БЕЗПЕКИ ПЕРСПЕКТИВНИХ МЕХАНІЗМІВ ІНКАПСУЛЯЦІЇ КЛЮЧІВ

The topic of the work. The work is devoted to the security arguments of promising post-quantum key encapsulation mechanisms CRYSTALS-Kyber, FrodoKEM and DSTU 8961:2019 "Skelya".

The goal of the work. The purpose of the work is to compare the security arguments of promising post-quantum key encapsulation mechanisms CRYSTALS-Kyber, FrodoKEM, DSTU 8961:2019 "Skelya" and to develop recommendations for their use in global and national practice.

Research methods. To achieve the goal, an analysis of existing research on the topic was used and a comparison of the obtained security arguments, taking into account the capabilities of quantum computers, was used.

The results. Received recommendations for the use of post-quantum key encapsulation mechanisms in global and national practice based on a comparison of theoretical and practical security arguments taking into account quantum threats.

Conclusions. DSTU 8961:2019 "Skelya" has good practical safety arguments, but needs additional research in theoretical formalized safety models, such as ROM, QROM. Due to the use of a non-standard field, DSTU 8961:2019 "Skelya" has strong security arguments against quantum attacks that are relevant to CRYSTALS-Kyber. From a practical point of view, CRYSTALS-Kyber is based on less researched issues than DSTU 8961:2019 "Skelya", but this is compensated by stronger theoretical arguments for security. FrodoKEM has the strongest security case among the key encapsulation mechanisms considered, but at the cost of slow performance. It is recommended to use FrodoKEM, DSTU 8961:2019 "Skelya", in systems requiring increased protection, and CRYSTALS-Kyber in systems requiring good speed.

Ключові слова: DSTU 8961:2019 "Skelya", CRYSTALS-Kyber, FrodoKEM, security arguments, NTRU, LWE

Вступ

Механізм інкапсуляції ключів визначається трійкою алгоритмів (Gen, Encaps, Decaps), де

- Gen – поліноміальний за часом алгоритм генерації ключової пари, що приймає параметр безпеки 1^λ та повертає ключову пару (pk, sk).
- Encaps – поліноміальний за часом алгоритм, що приймає відкритий ключ pk та повертає згенерований ключ K та його інкапсуляцію C.
- Decaps – поліноміальний за часом алгоритм, що приймає інкапсуляцію ключа C, секретний ключ sk та повертає ключ K або помилку дешифрування.

Особливістю сучасної криптографії є те, що безпека кожної криптографічної схеми або протоколу підкріплюються формальними математичними доказами, які звісно ж не гарантують, що не існує атак взагалі, проте дозволяють гарантувати, що не існує атак певного виду, які формально визначені через модель атак [2,3]. Для практичних задач вимагається безпека у моделі атак на основі обраних відкритих текстів (CPA) або у моделі атак на основі обраних шифротекстів (CCA)

У моделі атак на основі обраних відкритих текстів (CPA) супротивник обирає відкритий текст, а потім отримує відповідний зашифрований текст. Супротивник використовує отриману інформацію, щоб відновити відповідний відкритий текст для шифротексту, який раніше не бачив.

Під час атаки за допомогою обраного шифротексту (CCA) супротивник може розшифрувати довільні зашифровані тексти, наприклад, за допомогою доступу до обладнання для дешифрування з надійно вбудованим ключем дешифрування. В адаптивній атаці обраного зашифрованого тексту (CCA2), зловмисник має доступ до оракула дешифрування навіть після отримання зашифрованого тексту.

При побудові механізмів інкапсуляції ключів зазвичай обирається CPA безпечна схема асиметричного шифрування (яка ґрунтується на певній складній проблемі), до якої потім застосовується певне перетворення, що робить з неї CCA безпечний механізм інкапсуляції ключів. Розповсюдженим варіантами є перетворення Фуджісакі-Окамото та перетворення Дента. Усі механізми інкапсуляції ключів, що розглядаються в даній роботі (ДСТУ 8961:2019 [1], CRYSTALS-Kyber, FrodoKEM [3]) побудовані саме за таким принципом.

Порівняння перспективних механізмів інкапсуляції ключів

FrodoKEM ґрунтується на проблемі LWE [3,4], яка визначається наступним чином: Нехай $s \in Z_q^n$ - секретний вектор та χ - деякий розподіл ймовірностей. Визначимо LWE розподіл над $Z_q^n \times Z_q$ наступним чином: обирається вектор $a \in Z_q^n$ з рівномірного розподілу, ціле число $e \in Z$ з розподілу χ та формується пара $(a, b) = \langle s, a \rangle + e \bmod q \in Z_q^n \times Z_q$. Проблема LWE полягає у тому, щоб з m пар $(a_1, b_1), \dots, (a_m, b_m)$ знайти секретний вектор $s \in Z_q^n$. Значною перевагою LWE перед іншими проблемами є те, що вона має поліноміальні ймовірнісні редукції типу “від найгіршого до середнього” з складними проблемами з теорії решіток. Це є надзвичайно сильним аргументом безпеки. Якщо буде знайдений алгоритм, що може вирішити LWE за поліноміальний час, то це буде означати також існування алгоритму і для таких добре відомих проблем як SVP та CVP при відповідних параметрах. FrodoKEM використовує перетворення Фуджісакі-Окамото з прямим відхиленням [2], яке має доказати безпеки у ROM (моделі випадкового оракула) та QROM (моделі квантового випадкового оракула), що є важливим теоретичним аргументом безпечності конструкції.

CRYSTALS-Kyber ґрунтується на проблемі Module-LWE [3], яка є різновидом проблеми LWE з додатковою структурою. Безпосередні обчислення виконуються над поліномами замість векторів. Вектори можливо замінити на поліноми через те, що решіткою (за визначенням), фактично, є будь-яка абелева дискретна група. Абелеву дискретну групу можливо вкласти у евклідовий простір, що дає звичайні “геометричні” решітки з додатковою структурою. Обчисленням у CRYSTALS-Kyber відбуваються у полі $Z_q[X]/(X^n + 1)$. Такий вибір поля є стандартним для криптографії на решітках, оскільки воно дозволяє використовувати теоретико-числове перетворення для множення/ділення поліномів і має гарні теоретичні властивості у тому сенсі, що докази безпеки значно спрощується. Багато доказів безпеки (здебільшого редукції до складних проблем на модульних решітках) використовують властивості цього поля. CRYSTALS-Kyber має не такі сильні докази, як FrodoKEM, оскільки складність Module-* проблем, від яких Module-LWE має редукції, не відома. Проте, CRYSTALS-Kyber працює швидше, ніж FrodoKEM, що робить його гарним вибором для багатьох практичних застосувань, для яких FrodoKEM є відносно повільним. При цьому варто зауважити, що існують деякі атаки на поле $Z_q[X]/(X^n + 1)$, які дозволяють отримати значну перевагу на квантових комп’ютерах. Для CRYSTALS-Kyber вони є не актуальними саме через використання модульної структури решітки. Наразі не відомо технік, які б могли узагальнити атаки на випадок модульних решіток, проте цей напрямок потребує додаткових досліджень. CRYSTALS-Kyber використовує власну модифікацію перетворення Фуджісакі-Окамото, для якої автори отримали докази у моделях ROM та QROM.

ДСТУ 8961:2019 спирається на добре вивчену проблему NTRU, яка полягає у тому, щоб для поліному h з деякого поля Галуа знайти поліноми f, g для яких виконується $hf = g \bmod q$ та їх норма (зазвичай евклідова норма, проте у загальному випадку можливе використання інших норм) є меншою за $\sqrt{q}/\gamma$, де γ є загальним параметром, а q є модулем поля. ДСТУ 8961:2019 “Скеля” має доволі не звичайний вибір поля. Річ у тому, що на проблему NTRU в полі $Z_q[X]/(X^n + 1)$ існує серія атак [4], яка дозволяє вирішити проблему NTRU за поліноміальний час, використовуючи структуру підкілець. Зауважимо, що зазначені атаки можливі тільки у випадку у випадку $q > O(n^{2.7})$, який не є релевантним для ДСТУ 8961:2019 “Скеля”. Такі ситуації зустрічається здебільшого тільки у схемах гомоморфного шифрування та більш екзотичних конструкціях. Поле $Z_q[X]/(X^n - X - 1)$, яке використовується у ДСТУ 8961:2019, не має не тривіальних підкілець, що робить подібні атаки не можливими, навіть у випадку, якщо будуть знайдені розширення цих атак на випадок тих параметрів, що використовуються у ДСТУ 8961:2019. Додатково, вибір цього поля захищає від потенційного узагальнення технік, що використовувалися для вирішення Ideal-SVP за поліноміальний час.

Можливо сказати, що вибір поля $Z_q[X]/(X^n - X - 1)$ є додатковим аргументом безпеки для випадку алгебраїчних та квантових атак. ДСТУ 8961:2019 використовує власний гібридний варіант перетворень Дента. Для ДСТУ 8961:2019 ситуація є дещо невизначеною у цьому сенсі доказової безпеки у моделях ROM та QROM. Публікацій, що проводять аналіз саме ДСТУ 8961:2019 у моделях ROM чи QROM не існує. З іншої сторони, ДСТУ 8961:2019 є модифікацією стандарту ANSI X9.98 і відрізняється тільки іншим полем та алгоритмом формування малих поліномів. У цілому, усі докази у QROM або ROM для ANSI X9.98, що не використовують структуру поля, є релевантними і для ДСТУ 8961:2019. Проте, комплексного аналізу стандарту x9.98 також є доволі мало. Особливо це стосується випадку QROM. Хоча окремі результати відомі, проте для ДСТУ 8961:2019 не вистачає комплексного вивчення у моделях ROM та QROM.

Порівняння високорівневої структури механізмів інкапсуляції ключів ДСТУ 8961:2019, CRYSTALS-Kyber та FrodoKEM наведено в таблиці 1.

Табл.1 Порівняння високорівневої структури механізмів інкапсуляції ключів

	CRYSTALS- Kyber	FrodoKEM	ДСТУ 8961:2019
Складна проблема	Module-LWE	LWE	NTRU
Поле	$Z_q[X]/(X^n + 1)$	-	$Z_q[X]/(X^n - X - 1)$
Перетворення	Власна модифікація перетворення Фуджісакі-Окамото	Фуджісакі-Окамото з відхиленням	Власна модифікація перетворення Дента
Докази у ROM	+(сильний доказ)	+(сильний доказ)	+/-
Докази у QROM	+	+	+/-

Конкретні значення безпеки

Найкращі атаки на механізми інкапсуляції ключів зводяться до редукції базису решіток без використання їх алгебраїчної структури. Найкращими відомими алгоритмами редукції решіток є BKZ [3, 4] та його численні модифікації. BKZ є рандомізованим варіантом LLL. На кожному кроці знаходиться найменший вектор на проєктивній решітці меншої розмірності, додається до базису і проводиться LLL редукція. У випадку ДСТУ 8961:2019 найкращою атакою є гібридна атака, що використовує алгебраїчні методи для пришвидшення редукції.

На практиці, для оцінки безпеки криптографічних систем на решітках, як правило, використовується модель core-SVP, яка полягає у тому, що вартість редукції визначається часом роботи алгоритма пошуку малого вектора. Поліноміальна кількість викликів цього алгоритма ігнорується. Модель є доволі старою, проте вона є загальноприйнятим стандартом в криптографії на решітках. Наразі тривають дискусії у науковій спільноті про створення інших моделей, що прив'язані до певних технічних чи фізичних характеристик, проте, ці моделі доволі сирі. У таблицю 2 зведені данні про конкретні оцінки безпеки КЕМ. На Рис 1 наведено порівняння конкретних оцінок (класичних та квантових) безпеки КЕМ у вигляді діаграми.

Табл.2 Конкретні оцінки безпеки перспективних механізмів інкапсуляції ключів

Рівень безпеки	CRYSTALS- Kyber	FrodoKEM	ДСТУ 8961:2019
128	145/132	118/107	-
192	210/191	183/166	-
256	275/250	255/232	265/248
384	-	-	424/408
512	-	-	562/534

Конкретна безпека в бітах

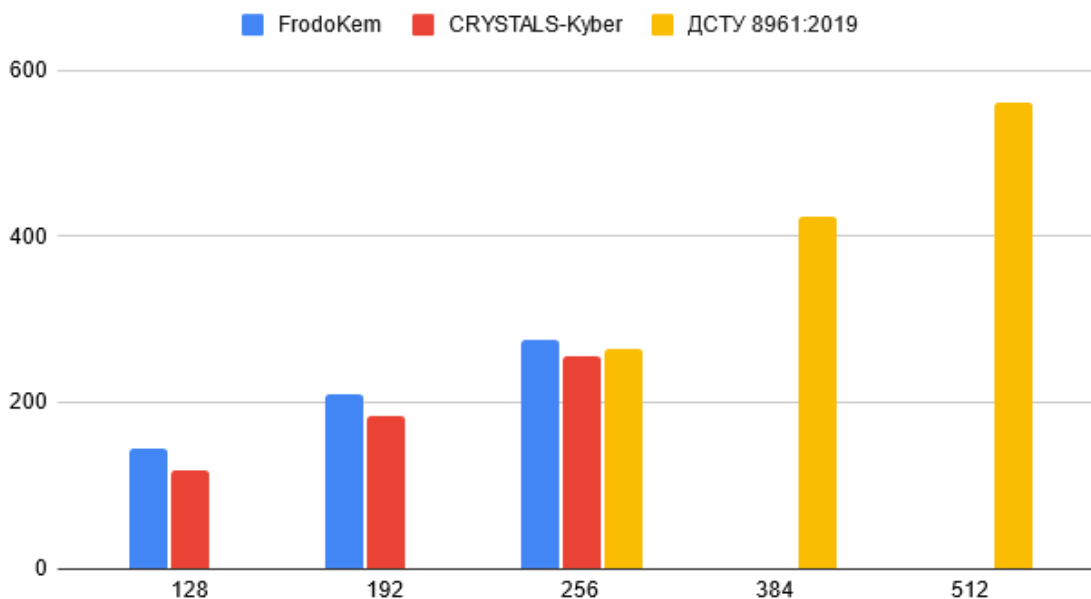


Рис. 1 Порівняння конкретної безпеки в бітах

Висновки

ДСТУ 8961:2019 “Склея” має гарні практичні аргументи безпеки, проте потребує додаткових досліджень у теоретичних формалізованих моделях безпеки, таких як ROM, QROM. Завдяки використанню нестандартного поля, ДСТУ 8961:2019 “Склея” має сильні аргументи безпеки від квантових атак, які є релевантними для CRYSTALS-Kyber, який, з практичної точки зору, ґрунтується на менш досліджених проблемах, ніж ДСТУ 8961:2019 “Склея”, проте це компенсується сильнішими теоретичними аргументами безпеки. FrodoKEM має найсильніші аргументи безпеки серед розглянутих механізмів інкапсуляції ключів, проте ціною за це є низька швидкодія. Рекомендовано використання FrodoKEM, ДСТУ 8961:2019 “Склея” у системах, що вимагають посиленого захисту, а CRYSTALS-Kyber у системах, що вимагають гарної швидкодії.

Література

1. ДСТУ 8961:2019 Інформаційні технології. Криптографічний захист інформації. Алгоритми асиметричного шифрування та інкапсуляції ключів [чинний від 2019-12-21]. Київ, 2019. 64 с. (Інформація та документація)
2. Горбенко І. Д., Горбенко Ю. І. Прикладна криптологія. Теорія. Практика. Застосування: монографія. Харків : Форт, 2012. 880 с.
3. NIST Post-Quantum Cryptography Standardization Project : веб сайт. URL: <https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization> (дата звернення: 28.10.2022)
4. Albrecht M., Bai S., Ducas L. A Subfield. Lattice Attack on Overstretched NTRU Assumptions. CRYPTO, 2016. P. 153–178.

КАНДІЙ Сергій Олегович – Харківський національний університет імені В. Н. Каразіна; аспірант кафедри безпеки інформаційних систем і технологій, факультет комп’ютерних наук, АТ “Інститут Інформаційних Технологій”, технік-конструктор; Україна; e-mail: sergeykandy@gmail.com; ORCID: <https://orcid.org/0000-0003-0552-8341>

Наукові інтереси:

– постквантова криптографія.

УДК 004.9

КАПЛУН В.М., ТОЛСТОЛУЗЬКА О.Г.

СЕКЦІЯ 2,7

РОЗРОБКА КОМП'ЮТЕРНОЇ МОДЕЛІ WEB-ЗАСТОСУНКО З ОБРОБКОЮ ВЕЛИКИХ ОБСЯГІВ ДАНИХ

Theme of the research: development of a computer model of a Web-application for processing large volumes of data to determine the psychophysical abilities of a person.

The main purpose of the research: development of a Web-service model, with the organization of the questionnaire process, testing and processing of the received data. The developed computer model should enable specialists in other fields, not related to programming, in this case psychologists, to apply methods of processing large amounts of data. These specialists should be able to analyze the data stored after user testing at any time, online, in a few clicks. Instead of using offline psychophysical testing programs. Offline programs lack tools for automated analysis of data obtained after testing a large number of people. Up to this point, specialists had to install and run tests on each PC by themselves. After obtaining some results, it is necessary to manually download and merge this data into a common document. And only after that the file was loaded into a separate program for further processing of large amounts of data.

The methods of the research: to develop this model it was decided to use the WebStorm 2020.2.3 environment and the JavaScript programming language. For data storage was used DBMS mongodb. WEB-application is placed on the cloud PaaS-platform - heroku.

Results: a computer model of a web-application with processing of large amounts of data was created. It consists of registration-authorization pages, personal cabinet, pages with tests and pages with data processing tools with limited access, only for specialists.

Conclusions: we can say that the web-application has great potential. Because the developed computer model provides an opportunity to apply methods of processing large amounts of data by specialists of other industries, remote from programming, in this case psychologists.

Key words: *Web, UML, MongoDB, HTML, Javascript, CSS.*

Постановка задачі дослідження

На даний час є значна кількість розроблених та працюючих моделей по тестуванню та збереженню отриманих результатів. Також в сучасному світі з кожною секундою все більше завантажуються інформації в мережу, що спонукає все частіше застосовувати алгоритми обробки великих обсягів даних для облегшення впорядкування та обробки цих даних.

Наразі нам відомо лише про використання офлайн програми по психофізичному тестуванню. Та відсутність широкого інструментарію по автоматизованому аналізу даних отриманих при тестуванні. До даного моменту фахівці змушені самостійно встановлювати та запускати тестування на кожному ПК. Після отримання деяких результатів, необхідно було в ручному режимі ці дані завантажувати та об'єднувати в загальний документ. І лише після цього в окрему програму завантажували файл для подальшої обробки великих обсягів цих даних.

Тому після озвучення всієї проблематики цього питання, мої керівником було поставлено завдання розробити комп'ютерну модель web-застосунку з обробкою великих обсягів даних.

Новизна цієї роботи полягає в тому, що розроблена комп'ютерна модель надає можливість застосування методів обробки даних фахівцями інших галузей, віддалених від програмування, в даному випадку психологів. Цим фахівцям надається можливість в будь який момент, в онлайн режимі, в декілька натискань аналізувати дані які були збережені після тестування користувачів.

На рисунку 1 зображена загальна діаграма «Комп'ютерної моделі WEB-застосунку з обробкою великих обсягів даних».

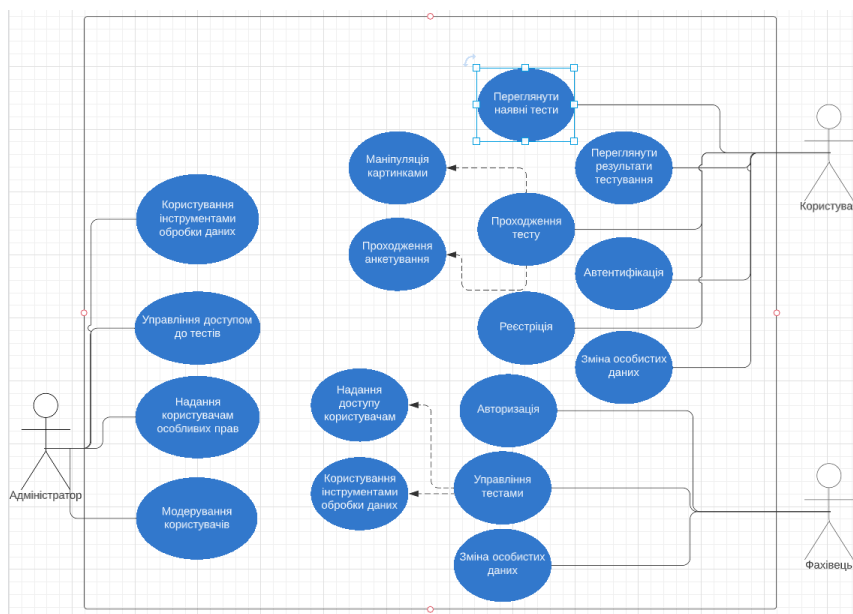


Рис. 1. Діаграма варіантів використання “Комп’ютерної моделі WEB-застосунку з обробкою великих обсягів даних”

Рекомендації по використанню “Комп’ютерної моделі обробки великих обсягів даних на базі отриманих результатів психофізичного тестування”.

Для авторизованого користувача доступна головна сторінка. Є можливість відкривати та проходити бажаний тест. Головна сторінка показана на Рис. 2 містить всю основну інформацію по доступним тестам.

Тест на психофізичні здібності

Ця оцінка вимірює ваші загальні психофізичні здібності. Тест містить чотири розділи:

1. Субтест на словесне міркування
2. Субтест на просторові міркування
3. Субтест на абстрактне міркування
4. Підтест кількісного міркування

Ваші результати відображатимуться після кожного підтесту, а загальний бал буде надано після виконання всіх чотирьох підтестів.

Примітка. Після завершення першого підтесту ви повинні відповісти на кілька запитань про себе, перш ніж ви зможете переглянути свої результати. Дані, які ми збираємо про вас, використовуються для анонімного дослідження. Ми не запитуватимемо вашу електронну адресу чи інші контактні дані. Тест і його результати надаються безкоштовно; від вас не вимагатимуть платити за отримання результатів.

Для початку виберіть один із трьох тестів нижче. Тест відкриється в новому вікні.

1) Усне міркування	Оцінка: 0%
2) Просторове міркування	Оцінка: 0%
3) Абстрактне міркування	<button>Відкритий тест</button>
4) Кількісне міркування	<button>Відкритий тест</button>

Загальні результати

2 з 4 тестів виконано. Виконайте всі чотири тести, щоб отримати загальні результати.

Рис. 2 “Головна сторінка веб сайту”.

Після того як користувач натискає одну із доступних кнопок «Відкрити тест» з’являється сторінка з інструкцією до даного тесту та кнопкою «Почніть».

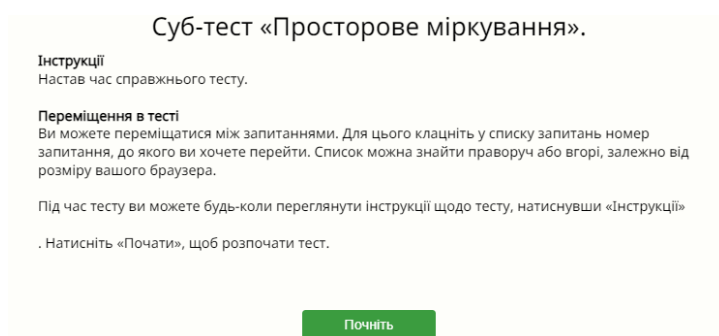


Рис. 3 “Сторінка інструкція до окремого тесту”.

На рисунку 4 зображено один із пунктів тесту. На цій сторінці користувач має змогу переглянути повторно інструкцію, переключатися між пунктами тесту, обирати та відхиляти варіанти.

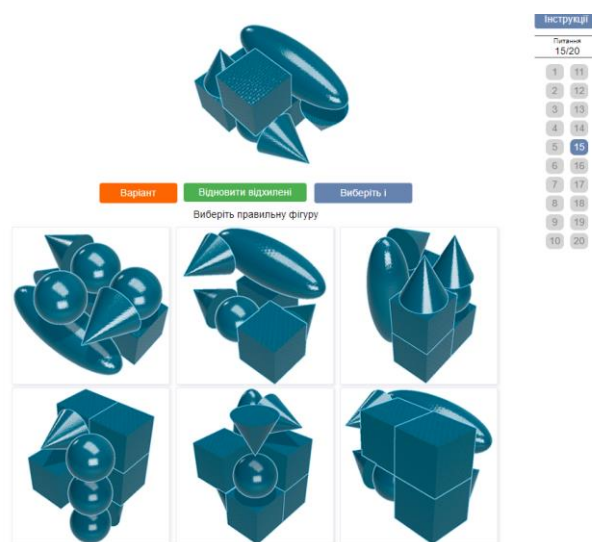


Рис. 4 “Сторінка тестування користувача”.

Після завершення тесту автоматично завантажувється сторінка з результатом пройденого тестування.



Рис. 5 “Сторінка результатів тестування користувача”.

ЛІТЕРАТУРА

1. Introduction to JavaScript Object Notation: A To-the-Point Guide to JSON, USA, 2015, ISBN: 1491929480, 126 pages.
2. Дэвид Хантер, Джефф Рафтер, Джо Фаусетт, Эрик ван дер Влиет, и др. XML. Работа с XML, 4-е издание = Beginning XML, 4th Edition. — М.: «Диалектика», 2009. — 1344 ст. — ISBN 978-5-8459-1533-7.
3. Методи аналізу даних: навчальний посібник для студентів/ В. Є. Бахрушин. — Запоріжжя: КПУ, 2011. — 268 с. ISBN 978-966-414-103-8.
4. Data Analysis for Omic Sciences: Methods and Applications, Volume 82 1st Edition Joaquim Jaumot Carmen Bedia Roma Tauler, USA, 2018, ISBN: 9780444640451, 740 pages.
5. Fayyad, Usama; Piatetsky-Shapiro, Gregory; Smyth, Padhraic (1996). "From Data Mining to Knowledge Discovery in Databases".
6. Learn JavaScript VISUALLY with Interactive Exercises.
7. Макфарланд Девід JavaScript. Детальне руководство; Ексмо - М., 2015. - 608 ст.
8. The Pragmatic Programmer: Your Journey to Mastery.
9. Front-End Developer Handbook — 2019.

КАПЛУН Владислав Миколайович – кафедра прикладної і теоретичної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: xa11867768@student.karazin.ua; ORCID 0000-0001-5787-4786

Наукові інтереси:

– *технології створення та супроводження веб-сайтів;*

ТОЛСТОЛУЗЬКА Олена Геннадіївна – д. т. н., с. н. с. ; Харківський національний університет імені В.Н. Каразіна; м. Харків, майдан. Свободи, 6, 61022; e-mail: elena.tolstoluzka@karazin.ua, ORCID 0000-0003-1241-7906

Наукові інтереси:

– *технологія паралельного розподіленого програмування;*

КИЛИМНИК Д.О.

ОСНОВНІ ВИЗНАЧЕННЯ ТА ХАРАКТЕРИСТИКИ ІНФОРМАЦІЙНО-ВИМІРЮВАЛЬНИХ СИСТЕМ

In this report, the main definitions related to information and measurement systems are considered, the features that broadly characterize them are given. Questions related to the types of such systems (measuring systems, automatic control systems, technical diagnostic systems, image recognition systems, telemetry systems), their main functions are also discussed. The purpose of the report is the analysis and detailed consideration of the use and characteristics of information and measurement systems. Special attention was paid to the consideration of information characteristics and such a mathematical discipline as information theory. During the review, basic theoretical information was given, as well as formulas for entropy, amount of information, etc. were considered. During the writing of the report, the following general scientific research methods were used: empirical and theoretical. As a result of the report, a complete picture of information and measurement systems was obtained, their detailed characteristics were provided. It was determined that the reliable operation of technical equipment is possible in the presence of an information and measurement system or an information and measurement complex installed on it to receive and process information about the state of the research object under the influence of external factors, they are the ones that measure and control physical parameters. In the report, we come to the conclusion that the information and measurement system combines many devices and converters, each of which performs relatively simple functions. At the same time, it is not just a sum of independent devices, but a combination of interdependent devices that jointly participate in the performance of some complex function or a number of functions. The system is characterized by automatic execution of all functions, starting from the collection of information and ending with its display or input into a computer.

Ключові слова: інформаційно-вимірювальна система, вимірювання, контроль, характерні риси, інформація, ентропія.

Робота будь-якого технічного обладнання визначається технологічним процесом, який характеризується сукупністю фізичних параметрів. Надійна робота технічного обладнання можлива при наявності інформаційно-вимірювальної системи (далі ІВС) або інформаційно-вимірювальні комплекси (далі ІВК), встановлених на ньому для отримання і оброблення інформації про стан об'єкта дослідження при впливі на нього зовнішніх факторів. ІВК та ІВС здійснюють вимірювання та контроль фізичних параметрів. Сучасні ІВС і ІВК здатні виконувати функції вимірювання, контролю та діагностики одночасно.

Збільшення контрольних параметрів в ІВС та ІВК при дослідженні об'єкта, підвищення швидкості протікання технологічних процесів привело до збільшення вимірювальної інформації. Ці фактори привели до створення ІВС та ІВК. Під ІВС розуміють системи, які призначені для автоматичного отримання інформації безпосередньо від об'єкта дослідження шляхом вимірювання і контролю, оброблення цієї інформації і відображення у вигляді чисел, графіків, гістограм і, таким чином, характеризують стан цього об'єкта.

Згідно стандарту ІВС – сукупність функціонально об'єднаних вимірювальних, обчислювальних та інших допоміжних технічних засобів для отримання вимірюваної інформації, її перетворення, оброблення з метою подання її оператору в зручному вигляді для подальшого оброблення і прийняття рішення. Таким чином до ІВС входять ІВК, засоби контролю, діагностики, ідентифікації та виконавчі механізми [1].

Вимірювальні системи забезпечують: оператора інформацією про стан об'єкта, математичне оброблення вимірюваної інформації та її відображення на моніторі електронної обчислювальної машини (далі ЕОМ) і вивід її на цифрову реєстрацію.

Візуальне відтворення можливе в аналоговій або цифровій формі за допомогою приладів, які показують та реєструють. Прилади, що показують, застосовуються, коли інформація використовується людиною безпосередньо при її одержанні.

Прилади, що реєструють, застосовуються, коли інформація призначається для наступного використання. При цьому, як правило, інтерес представляє зміну вимірюваних величин у часі, так що прилади, що реєструють, видають при аналоговому відтворенні двох координатні криві (по одній координаті – час, по іншій – вимірювана величина), при цифровому відтворенні – таблиці, у яких кожному дискретному значенню вимірюваної величини відповідає значення часу.

Якщо інформація надходить у регулюючі або керуючі пристрої, в ЕОМ (комп'ютер) або в пристрої довгострокового зберігання, то вона представляється на виході ІВС у вигляді електричних сигналів – аналогових або цифрових. При цьому ті дані, які призначаються не для оперативного використання при їхньому одержанні, запам'ятовуються разом з координатами часу, що вводять у пристрої зберігання також у вигляді цифрових електричних сигналів.

Існує кілька різновидів ІВС.

- **Вимірювальні системи.** Їхня функція складається в одержанні кількісної інформації про значення фізичних величин шляхом прямих, сукупних, непрямих вимірів.
- **Системи автоматичного контролю.** Їхня функція складається у встановленні відповідності між станом об'єкта й заданою нормою й у виробітку судження про даний або (чи) про майбутній стан об'єкта. Ці системи називають системами централізованого контролю. Інформація, видана системою, носить якісний характер і відповідає на запитання чи перебуває об'єкт в заданому режимі. За допомогою таких систем вимірюються фізичні величини, що характеризують стан об'єкта, і результати вимірів порівнюються зі значеннями, що прийняті за норму.
- **Системи технічної діагностики.** Їхня функція складається в контролі стану різних технічних пристроїв, у тому числі пристроїв автоматики, обчислювальної техніки, радіотехніки, у виявленні їхніх відмов і визначенні несправних елементів. Для таких систем характерне застосування спеціальних методів пошуку несправностей.
- **Системи впізнання образів.** Їхня функція складається у визначенні відповідності між досліджуванним об'єктом і заданим образом. Можливі образи: "людина", "буква А" і т.п.
- Особливе місце серед ІВС займають **телевимірювальні системи**, функції яких звичайно такі ж, як у вимірювальних систем і систем автоматичного контролю. Однак вони мають істотну особливість: у них інформація про значення вимірюваних величин передається на більші відстані - від сотень метрів до тисячі кілометрів.

ІВС може існувати й використатися самостійно, але може й входити як складова частина в більшу складну інформаційну систему, що виконує збір і обробку не тільки вимірювальної, але й іншої інформації, або в керуючу систему, що виконує не тільки збір і обробку інформації, але й активний вплив на об'єкти.

Зазначимо характерні риси ІВС:

- ІВС властива централізована структура. Інформаційні потоки спрямовані в один центральний пункт, рідше – у кілька центрів, взаємно координованих. В ієрархічних системах керування інформація направляється спочатку в підлеглі пункти керування, а з них частина інформації/результати обробки надходять у центральний пункт керування.
- Призначення ІВС звичайно полягає в тому, щоб поставляти інформацію для активного використання в тих або інших випадках людської діяльності. Ця особливість проявляється в остаточному підсумку навіть у тих ІВС, які збирають інформацію дослідницького характеру.
- Для ІВС, що збирають інформацію про велику кількість величин, характерне використання загальних блоків, пристроїв або приладів, які по черзі передають, перетворюють, обробляють сигнали від різних датчиків. Кожний такий загальний блок обслуговує, таким чином, безліч окремих інформаційних каналів. Системи, які містять загальні блоки, що обслуговують безліч інформаційних каналів, називаються багатоканальними. Як правило, більшість ІВС є багатоканальними [2].

Приклади загальних блоків (пристроїв): аналого-цифровий перетворювач; пристрій обробки інформації; канал зв'язку.

Із приводу термінів "канал зв'язку" і "інформаційний канал" варто зробити зауваження. По одному каналу зв'язку може бути організована передача сигналів безлічі інформаційних каналів. Таким чином, цілком природно поняття багатоканальної ІВС із одним каналом зв'язку. У цьому випадку той самий термін "канал" використовується в різних значеннях.

Тепер більш детально розглянемо інформаційні характеристики ІВС. Як вже було зазначено метою вимірювання і контролю є отримання інформації про стан контрольованого

об'єкта і його параметрів. Інформація подібно матерії і енергії існує об'єктивно, незалежно від свідомості людини.

Теорія інформації – це математична дисципліна, яка вивчає процеси зберігання, перетворення і передачі інформації. Інформація, отримана в процесі вимірювання називається вимірюваною інформацією. Вимірювана інформація є кількісною інформацією і служить для оцінки кількісних даних про досліджувані об'єкти або явища, незалежно від їх природи, отриманих дослідниками за допомогою засобів вимірювальної техніки. Кількісна інформація може розглядатися тільки в такому ланцюгу: досліджуваний об'єкт – засіб вимірювання – дослідник (оператор).

Для характеристики кількісної інформації використовують логарифмічну міру (Хартлі):

$$I = \log aN \quad (2.1)$$

де I – кількість інформації, отриманої в результаті одного відліку;

N – кількість можливих різних результатів вимірювання конкретним засобом вимірювання;

a – число, яке визначає одиницю кількості інформації (при $a = 2$ одиницю кількості інформації називають біт). 1 біт інформації є в одному розряді двійкового числа. Дійсно, один розряд може приймати одне із двох значень: 0 або 1, тобто $N = 2$. Тоді $I = \log 2 = 1$ біт.

Формула Хартлі справедлива тільки тоді, коли кожне із N повідомлень відображається з однаковою ймовірністю. Справа в тому, що інформацію несе тільки повідомлення невідоме приймачу цього повідомлення. Кількість інформації в раніше відомому повідомленні рівна нулю. В загальному випадку кількість інформації, яка є в одному повідомленні, залежить від ймовірності надходження цього повідомлення:

$$I = \log 2(1/p_i)[\text{біт}] \quad (2.2)$$

де p_i – ймовірність відображення i – го повідомлення.

Із формули (2.2) видно, що чим менша ймовірність отримання повідомлень, тим більше інформації воно несе. Якщо закон розподілу значень вимірюваного параметра не є рівно ймовірними, отримані результати будуть нести різну кількість інформації.

Для інформаційної оцінки системи або комплексів в теорії інформації введено поняття ентропії. **Ентропія** – це середнє значення інформації на одне повідомлення. Для отримання формули ентропії розглянемо джерело інформації, яке видає N незалежних дискретних повідомлень з ймовірністю $p_1, p_2, p_i, \dots, p_n$. Будемо розглядати послідовність із m повідомлень ($m \gg N$). Відповідно з теорією ймовірності кожне i -те повідомлення з'явиться в послідовності m разів [1]. Враховуючи, що кожне i -те повідомлення несе інформацію $I_{1i} = \log 2(1/p_i) = -\log p_i$, всі отримані повідомлення складуть інформацію $I_{mi} = -m p_i \log 2 p_i$. Середнє значення інформації на одне повідомлення, тобто ентропія джерела, буде рівна:

$$H = Lm / m = -p_i \log 2 p_i [\text{біт}]. \quad (2.3)$$

Ця формула отримала назву формули Шеннона. При рівних значеннях $p_1 = p_2 = p_n$ ця формула перетворюється в формулу Хартлі:

$$I = \log 2N \quad (2.4)$$

Тепер наведемо властивості ентропії:

1. Ентропія джерела інформації максимальна, якщо всі повідомлення надходять від цього джерела рівно ймовірно.
2. Ентропія зменшується, якщо є залежність між виходами.

Ентропія об'єднання. Ентропія системи $H(X, Y, \dots)$, яка складається із декількох підсистем, які мають ентропії $H(X), H(Y) \dots$, рівна сумі ентропій, якщо повідомлення підсистем незалежні. Якщо є залежність, то загальна ентропія менша суми ентропій підсистем:

$$H(X, Y, \dots) \leq H(X) + H(Y) + \dots \quad (2.5)$$

Інформаційна надлишковість проявляється в тому, що інформаційний зміст повідомлень є меншим, від можливостей засобів відображення цих повідомлень. Інформаційна надлишковість може бути обумовлена такими факторами:

1. Надлишкове число елементів дискретних сигналів порівняно з необхідним значенням для відображення даних повідомлень. Наприклад. Для передачі інформації використовують 8-розрядний двійковий код. Кількість різних символів, які використовуються в інформації-53, в тому числі: 33 букви, 10 десяткових цифр, 10 допоміжних знаків (кома, крапка, дужки). Для кодування всіх символів достатньо 6 розрядів коду ($2^6 = 64$). Два розряди є зайвими, вони створюють надлишковість.
2. Неоднакова ймовірність відображення окремих повідомлень. Згідно першої властивості ентропії вона при цьому зменшується. Наприклад. Джерело видає 4 різні повідомлення. Якщо повідомлення рівно ймовірні, буде рівна 2 біта. При різних ймовірностях повідомлень – 1/2, 1/4, 1/8 ентропія буде рівна 13/4 біт.

Наявність взаємних зв'язків між повідомленнями також, згідно другої властивості ентропії, зменшує ентропію повідомлень. Інформаційна надлишковість оцінюється виразом:

$$R = I_{\max} - I / I_{\max} = 1 - I / I_{\max}, \quad (2.6)$$

де $I_{\max}$ – максимально можливе значення ентропії для прийнятих засобів відображення повідомлень;

I – дійсне значення ентропії.

Також розглянемо таке поняття як потік інформації. **Потік інформації** – це середня кількість інформації, яку видає джерело за одиницю часу:

$$\Phi = I / T = H / t_1 [\text{бim} / \text{с}] \quad (2.7)$$

де H – ентропія джерела,

t_1 – час, за який джерело видає одне повідомлення.

Пропускна здатність інформаційного каналу Інформаційний канал – це будь-яка система, по якій передається інформація [1]. Пропускна здатність – це максимальна швидкість, з якою канал може передавати інформацію:

$$C = (I / T)_{\max} = I_{\max} / T [\text{бim} / \text{с}] \quad (2.8)$$

Якщо за час T канал може передати m повідомлень, причому передача кожного повідомлення проходить за час t_1 , то формулу (2.9) можна записати в такому вигляді:

$$C = mH_{\max} / (mt_1) = H_{\max} / t_1 \quad (2.9)$$

де $H_{\max}$ – максимальна ентропія, яку можуть мати засоби відображення інформації.

Можемо підсумувати, що інформаційно-вимірювальна система поєднує безліч приладів і перетворювачів, кожний з яких виконує відносно прості функції. При цьому вона являє собою не просто суму незалежних приладів, а об'єднання взаємозалежних приладів, що беруть участь спільно у виконанні деякої складної функції або ряду функцій. Для системи характерно автоматичне виконання всіх функцій, починаючи від збору інформації й кінчаючи її відображенням або введенням в ЕОМ.

ЛІТЕРАТУРА

1. Паламар М., Стрембіцький М., Паламар А. Проектування комп'ютеризованих вимірювальних систем і комплексів: навч. посіб. Тернопіль: ТНТУ, 2018. 7-13 с.
2. Орнатський П. Теоретичні основи інформаційно-вимірювальної техніки. Київ, "Вища школа", 1982. 455 с.

КИЛИМНИК Діана Олександрівна – студентка 4 курсу факультету комп'ютерних наук, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: kylymnyk.n@gmail.com, xa12850471@student.karazin.ua; ORCID: 0000-0003-0387-1664.

Наукові інтереси:

– системи автоматизованого збору даних.

ON THE DYNAMICS OF A SHORT BUNCH OF RELATIVISTIC ELECTRONS IN PLASMA

The features of radiation and dynamics of short 1D bunches of relativistic electrons propagating in a homogeneous isotropic plasma are considered. To estimate the amplitude of the wakefield, a model is used that takes into account the dynamics of individual particles in a self-consistent electric field. When the bunch size in the direction of its motion exceeds several wavelengths of the excited plasma oscillations, no field accumulation occurs in its volume, and the bunch dynamics is determined by the development of a two-stream dissipative-type instability. When the bunch size is smaller than the wavelength of the driven wave the spontaneous self-profiling of the bunch takes place as a result of which a wakefield with the maximum possible amplitude is formed behind the bunch. The wakefield maximum in the laboratory frame of reference is localized at the place of their initial formation.

Keywords: bunches of charges particles, two-stream instability, wakefield, JCUDA technology.

Introduction

The desire to accelerate charged particles in the potential of the driven wave excited by a dense bunch of electrons caused the appearance of a lot of theoretical and then experimental works. The dynamics of short electron beams in plasma was also of considerable interest for a number of applications even before that. It has been shown in a number of works (see, for example, [1,2]) that bunches of rapidly moving electrons are capable to keep stability due to the reversal of the Coulomb forces.

The most useful work that shed light on the structure of both the bunch internal electric field and the wakefield generated by the bunch in plasma was [3], where the dynamics of such a bunch driven by these fields was also clarified. A self-consistent description of bunch particles dynamics was used for numerical simulation of bunch behavior in one-dimensional [3, 4], two-dimensional [5], and three-dimensional descriptions [6]. Namely these numerical experiments confirmed the existence of the focusing action of internal beam fields. This paradoxical conclusion was interesting because the Coulomb repulsion forces between bunch electrons should be compensated by collective effects in the system “plasma+electron bunch”.

In many cases, the authors considered the electron beams (usually modulated) with size comparable to hundreds of plasma oscillation wavelengths. Therefore, it was important to investigate the stability and intensity of such formations. We note that the use of relativistic electrons gave reason to believe that the bunches are sufficiently stable due to the large relativistic mass of the electrons. Therefore, many problems were not solved self-consistently, assuming that the bunch retains its shape.

But the realization of the need to take into account the dynamics of processes inside the bunch of moving electrons forced us to return to a theoretical description of this phenomenon. First, the idea that the bunch emits a sufficiently intense electromagnetic field energy flow allows to use the parameter θ [7]: the ratio of the energy flow from the bunch to the total energy of the field inside its volume. Usually in electronics, $\theta < 1$ corresponds to regime of reactive generation, and $\theta > 1$ to regime of dissipative generation.

The bunches, which longitudinal (in the direction of its movement) size was comparable to the length of the wake wave or did not exceed ten such lengths, had the value of this parameter much less than unity. This manifests the similarity of such a process with dissipative beam instability. However, since all particles of the beam-bunch radiate plasma waves, and their neighbors find themselves in the integral field of this radiation, this regime is more similar to the superradiance and then it actually turned out to be. Therefore, a number of works that were carried out at Kharkov University were mainly devoted to studying the development of the process of superradiance of a fast electron bunch moving through the plasma [8].

The model chosen in the previous works of this cycle of studies represents the electrons of the bunch in its own frame of reference, which makes it possible not to take into account the magnetic field of the radiation, and for not very dense bunches one could make sure that such a description is quite correct [8, 9]. In this paper, we discuss the radiation efficiency of beam-bunches of different lengths. In addition, we will try to estimate the intensity of the field that arises behind a bunch of electrons moving at a relativistic speed.

The model for estimation of the wakefield generated by a moving electron bunch

The one-dimensional model developed in [3], due to its simplicity, allows to estimate how the amplitude of the wakefield is related to the parameters of the moving bunch-beam. The equations describing the behavior of one-dimensional electron bunch propagating through a dense plasma can be written as

$$\frac{d\xi}{d\tau} = \nu, \quad \frac{d\nu}{d\tau} = E(\xi), \quad (1)$$

$$E(\xi) = -\frac{2}{N} \sum_{\alpha}^N f_{\alpha} \cos[2\pi g_{\alpha}(\xi - \xi_{\alpha})] \Theta(\xi_{\alpha} - \xi), \quad (2)$$

where $2\pi\xi = K_0(z - V_0t)$, $\nu = K_0(V - V_0)/2\pi\gamma_L$, $\gamma_L^2 = e^2 K_0 M / m_e$, $g = (1 + \Delta\nu)^{-1}$, $\Delta = 2\pi\gamma_L / K_0 V_0$, $\tau = \gamma_L t$ and M is a total number of particles in the bunch, $E = eK_0 E / 2\pi m_e \gamma_L^2$, E is the electric field strength, f_{α} is a statistical weight of a large particle, modelling the beam, $\delta_D = V_0 / b$, b is the bunch length, $\theta = \delta_D / \gamma_0$, $\gamma_0 = (n_{b0} / 2n_{p0})^{1/3} \cdot \omega_{pe}$, e, m are the electron charge and rest mass, n_{b0}, n_{p0} are the non-perturbed dense of the beam and plasma correspondingly, $\omega_{pe} = \sqrt{4\pi e^2 n_{p0} / m_e}$ is the Langmuir frequency of plasma waves, excited in the wake of the bunch. Note that in an infinite beam-plasma system where the maximum growth rate of beam-plasma instability is equal to $\gamma_0 = (n_{b0} / 2n_{p0})^{1/3} \cdot \omega_{pe}$, the maximum of the field amplitude reaches the value of $E = m_e \gamma_0^2 / eK_0$ (in dimensional units).

In the case of extended bunches with the same number of particles, the largest amplitude of the wake radiation achieved in the process of instability development weakly depends on the initial longitudinal size of the bunch. For the system of units used in this model it is equal to $E = eK_0 E / m_e \gamma_L^2 \approx 1$. Note that if all particles of the bunch are collected at a point, then the amplitude of the wakefield would be twice as large $E = eK_0 E / m_e \gamma_L^2 = 2$, which is easy to see from Eq. (2). In this case, the degree of radiation coherence is 25%. It is important to note that if the beam is short enough, the field energy is taken out of the bunch volume in time $b / V_0 < \gamma_L^{-1}$ (e.g. $\theta > 1$). For sufficiently short beams (more precisely, bunches) with a large value of θ , the maximum amplitude of the wakefield in units used in the model (1)-(2) will be equal to $E = eK_0 E / m_e \gamma_L^2 \approx 1$. Moreover, if θ is sufficiently large (in this case, the beam size is relatively small), then the amplitude of the wakefield for the same number of particles in the bunch will be the same.

For small values of θ or for extended beams, the maximum field amplitude at the boundary of such a beam may be greater, e.g. $E = eK_0 E / 2\pi m_e \gamma_0^2$.

The wakefield of a short bunch of relativistic electrons

In the laboratory frame of reference, the radiation field of individual bunch electrons is an interference of oscillations with different phases. It is rational to believe that there are possible configurations and a velocity distribution of a bunch of emitters, which allow to achieve a small spread in the phases of the fields emitted by individual particles of the bunch in separate, generally speaking, rather extended regions.

Such velocity and density profiling of the bunch was proposed to be created artificially in order to achieve the maximum value of the wakefield [1,2], at least in certain regions of the plasma space. It was assumed that such a profiled bunch would be stable for a sufficient time. It should be noted that in addition to the short time of existence of the desired profile, such profiling was very difficult for experimenters to create. The purpose of the profiling was to create the largest possible electric field for accelerating particles

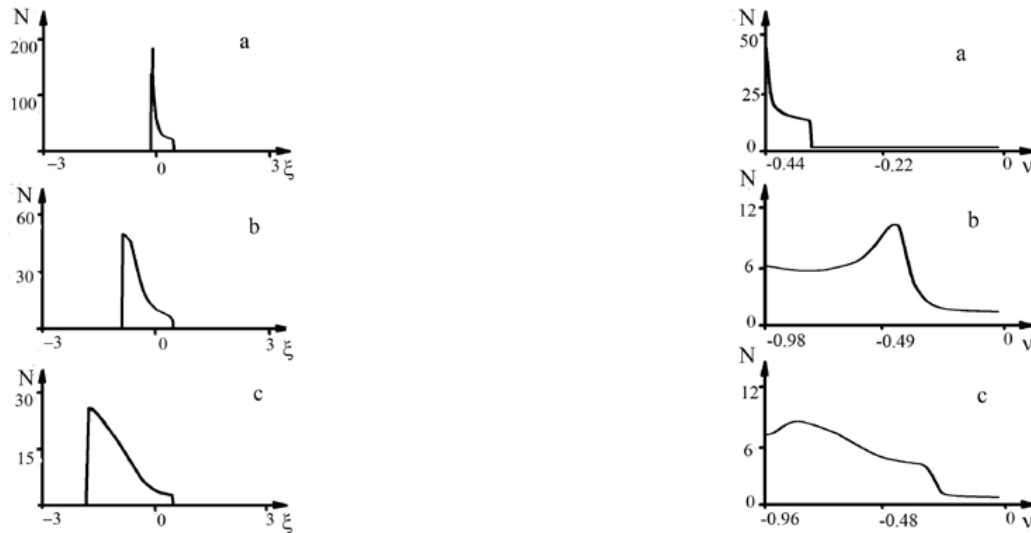


Fig. 1. The space distribution of electrons (left) and velocity distribution (right), [20]
a) $\tau = 1$; b) $\tau = 2$; c) $\tau = 3$

of another beam. It is all the more surprising that such profiling, as it turned out [8], occurs spontaneously in the bulk of the beam, and indeed, a region of the field maximum arises behind the beam, i.e. this region remains localized in the laboratory reference frame.

Of greatest interest are the regimes of excitation of the largest values of the wakefield, which are achieved in the case of short bunches, the size of which is less than the length of the wakewave. It follows from the calculation results that for bunches whose longitudinal size is smaller but still comparable with the wavelength, the phenomenon of bunch self-profiling is detected, the radiation efficiency of which in a certain region behind the bunch reaches a value of E for a given number of particles, which already noticeably exceeds most often observable, i.e. equal to one. Indeed, for bunches whose length exceeds several wavelengths of the emitted waves, the value of the field amplitude in the given normalization is always less than or on the order of unity. In the case of a bunch, all the particles of which come together at a point, the amplitude of the wakefield in the chosen representation (1) - (2) turns out to be twice as large and reaches its absolute maximum - a value equal to 2 [8].

The bunch as a whole is decelerated, forming characteristic triangular distributions of particles in the configuration and phase spaces (see Figs. 1, 2).

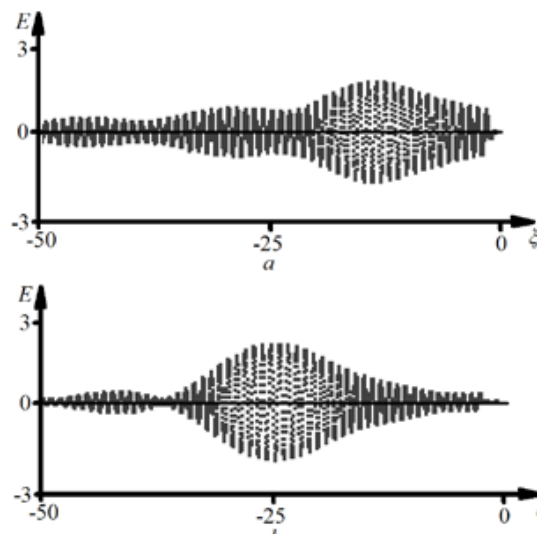


Fig. 2. The wakefield E in the rest frame of the bunch at times a) $\tau = 2$ and b) $\tau = 3$ [20]

The formation of such triangular distributions leads to the interference of the fields of the particles in a rather extended region of space of the laboratory frame of reference. It can be seen from Fig. 2 that the region of the field's absolute maximum $E(\xi) \approx 2$ is shifted with the speed of the bunch in the opposite direction from it. Note, the beam is located on the right side of the figure in the vicinity of $\xi = 0$.

This model describes qualitatively the dynamics of the bunch and its radiation, however, it makes it easy to detect the effect of beam self-profiling. The amplitude of the wakefield in this region is practically equal to the maximum possible for a bunch, the particles of which are collected in a very small region.

It is important to note that calculations in the beam reference frame are somewhat simpler than in the laboratory reference frame. And taking into account the fact that we are considering a one-dimensional model and the wake field is longitudinal, the transition to the laboratory frame of reference does not change the amplitude of the electric field, and all the above estimates of the wakefield maximum remain valid

REFERENCES

1. P. Chen P., Dawson J. M., Huff R. W. and Katsouleas T. Acceleration of electrons by the interaction of a bunched electron beam with plasma // Phys Rev. Lett., 1985. – V.54. – P. 693–696.
2. Katsouleas T. Physical Mechanisms in the plasma wake-field accelerator.// Phys.Rev. A., 1986. – V. 33. – P. 2056–2064.
3. Kuklin V. M., One-dimensional moving bunches of charged particles in plasma. Ukr. Phys. Journal 1986. – V. 31. – No. 6. – P. 853–857.
4. Kuklin V. M., Moiseev S. S., Panchenko. I. P. On the issue of transportation and radiation of short one-dimensional bunches of charged particles in plasma // Preprint № 1314, Space Research Institute. M., 1987. – 15 p.
5. Alterkop B. A, Zheksemin S. R, Rukhlin V. G. Tarakanov V. P. Two-dimensional dynamics of a compensated electron bunch in a dense plasma // Preprint of the Institute of High Temperatures of the Academy of Sciences of the USSR, 1986. – N 6–193. – P. 35–45.
6. Kuklin V. M., Moiseev S. S., Panchenko. I. P.; 3-D short Beam Dynamics // Reprint № 1619 of Institute of Space Research, M. 1989. – 11 p.
7. Kondratenko A.N., Kuklin V.M., Tkachenko V.I. Nonlinear theory of the beam instability in collisional plasma // Izv. Vuz. Radiofizika, 1978. – V. 21. – N 10. – P. 1535–1537 (in Rus.); About anomalous beam energy losses under development of dissipative beam instability. // Ukr. Phys. Journal 1979. – V. 24. – N 4. – P. 559–561 (in Rus.);
8. Kirichok A.V., Kuklin V.M., Mischin A.V., Pryimak A.V. Modelling of superradiation processes driven by an ultra-short bunch of charged particles moving through a plasma/ VANT, 2015 N 4–series “Plasma Electronics and New Methods of Acceleration” p.255-257.
9. Balakirev V. A., Karbushev N. I., Ostrovsky A. O., Tkach Yu. V. Theory of Cherenkov's Amplifiers and Generators on Relativistic Beams. – 1993. –Kiev: Naukova Dumka, 1993.–208 p.

КИРИЧОК Олександр Віталійович – к.ф.м.н., доцент, докторант кафедри штучного інтелекту та програмного забезпечення; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: sandyrcs@gmail.com; ORCID: 0000-0002-2283-1070

Наукові інтереси:

– нелінійна фізика. фізика плазми.

МІШИН Олександр Вікторович – ст. викладач кафедри штучного інтелекту та програмного забезпечення; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: o.v.mishyn@karazin.ua sandyrcs@gmail.com

ORCID: 0000-0001-7478-757X

Наукові інтереси:

– комп'ютерне моделювання фізичних процесів у плазмі.

КОВАЛЕНКО А. В., БУЛАВІН Д. О.

ВИКОРИСТАННЯ РОЗШИРЕНОГО АНАЛІЗУ ДОХОДУ ТА ВИТРАТ ДЛЯ УПРАВЛІННЯ ЯКІСТЮ ІТ-ПРОЕКТУ

В наш час комп'ютерні технології все більше й більше «захоплюють» навколишній світ, тому важливість їх якості повинна зростати та чітко перевірятися. Незначні помилки в подальшому можуть призвести до великих проблем з часом, та іншими ресурсами. Підвищення якості продукту є важливою складовою кожної компанії, оскільки це невід'ємна частина стратегії, що впливає на ефективність проекту та прибуток від нього. В останній час важливим критерієм якості є думка клієнта. Якщо він задоволений, то якість продукту вважається задовільною чи гарною.

В наш час все частіше використовують математичні моделі, які дозволяють «привести» різночасові грошові потоки (як дохід, так і витрати) до одного моменту часу шляхом так званого дисконтування.

Математична модель різниці аналізу доходу і витрат (1):

$$PV(B - C)_{iT} = \sum_{i=1}^T \frac{B_{it}}{(1+I)^t} - \sum_{t=1}^T \frac{C_{it}}{(1+I)^t} \quad (1)$$

де B_{it} – це сумарний дохід за окремий період t для варіанта i ,

C_{it} – сумарні витрати за період t для варіанта i ,

I – значення процентної ставки.

Математична модель відношення доходів і витрат (2):

$$PV(B/C)_{iT} = \left[\sum_{i=1}^T \frac{B_{it}}{(1+I)^t} \right] / \left[\sum_{t=1}^T \frac{C_{it}}{(1+I)^t} \right] \quad (2)$$

Завдяки цьому аналізу можна розраховувати різні варіанти розвитку подій в ході виконання проекту, такі як нехватка часу або людських ресурсів, непередбачувані зміни будь-якого характеру. На основі цього можна зобразити приклад графіку різниці аналізу витрат та прибутку, див. рис 1.

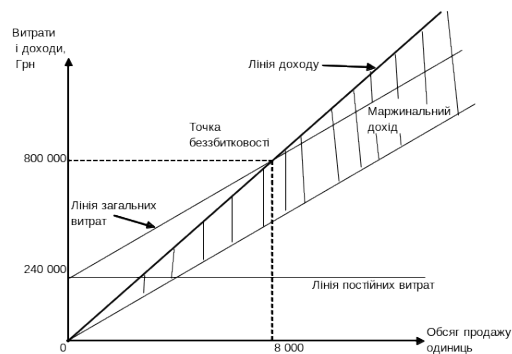


Рисунок 1 – Графічне представлення витрат та доходів

Аналіз економіки якості повинен починатися в першу чергу з процесів зв'язку цінності й витрат для користувача. Слід зазначити, що очікуваний результат цінності виконавця і вартість готового продукту можуть не співпадати з очікуванням замовника. Це пояснюється тим, що базуючись на категоріях цінності (необхідність кінцевого продукту і параметрів його якості, відповідають очікуванням певного сегменту ринку) і витрати, які визначають ефективність створення цінностей і визначають результат проекту.

В результаті аналізу необхідно визначити, яку частину доданої до продукту цінності за для покращення процесу необхідно передати користувачу без миттєвого повернення прибутків виробнику, а яку частину перенести до вартості проекту. Таким чином, при аналізі з ціллю

покращення процесу для забезпечення збігу очікувань виробника та споживача необхідно вирішити питання про розподіл витрат на забезпечення цінності продукту між виробником і споживачем. Зазвичай це пов'язано з аналізом конкурентоспроможності продукту/проекту на рівні якості.

Таким чином, при аналізі економіки якості обов'язково необхідно враховувати цінність цілі, яку компанія/команда припускає досягти безпосередньо своїм процесом (цінність для користувача і для виробника) а також ефективність, з якою буде досягнута ця ціль (забезпечення цінностей для користувача в процесі втілення якості) і витрат на досягнення цілі.

Процес визначення цілі має бути зорієнтований на визначення «правильних речей» для продажу необхідної цінності для користувача. Процес втілення (виконання) цілі повинен бути направлений на реалізацію з мінімальними втратами і витратами на невідповідність.

Першорядне завдання процесу виконання – досягти цільових величин з найменшими загальними затратами, зокрема шляхом мінімізації витрат на неякісну продукцію. Результативність і ефективність цілей мають бути розглянуті окремо для кожного під процесу.

У системі управління оцінка ефективності проектів є найвідповідальнішим етапом. Від того, наскільки об'єктивно і всебічно проведена ця оцінка, залежать терміни повернення вкладеного капіталу, варіанти альтернативного його використання, потік прибутку, що додатково генерується, в майбутньому періоді.

Слід розглянути базові принципи і методичні підходи, які використовуються в сучасній практиці оцінки ефективності проектів.

1. Оцінка ефективності реальних проектів повинна здійснюватися на основі складання обсягу витрат, з однієї сторони, і сум і термінів повернення капіталу, з іншої. Це загальний принцип формування системи оціночних показників ефективності, згідно з результатами будь-якої діяльності, яка повинна бути поєднана з витратами (використаними ресурсами) на її здійснення. Переважно до діяльності він реалізується шляхом сполучення прямого і зворотного потоків капіталу.

2. Оцінка обсягу витрат має охоплювати всю сукупність використовуваних ресурсів, що пов'язаних з реалізацією проекту. У процесі оцінки мають бути враховані всі прямі та непрямі витрати коштів (власних та позикових), матеріальних та нематеріальних активів, трудових та інших видів ресурсів. Як свідчить сучасна практика найчастіше оцінка обсягу витрат не відбиває непрямі витрати, пов'язані з підготовкою проекту до реалізації, формуванням необхідного обсягу ресурсів, контролю над реалізацією проекту тощо. Це не дозволяє здійснювати порівнянню оцінку ефективності проектів.

3. У процесі оцінки суми витрат та чистого грошового потоку мають бути приведені до справжньої вартості. На перший погляд здається, що витрати по відношенню до чистого грошового потоку завжди виражені у дійсній вартості, оскільки значно передують йому. У реальній практиці це не так - процес інвестування в більшості випадків здійснюється не одномоментно, а проходить низку етапів. Тому, за винятком першого етапу, всі наступні суми витрат повинні наводитись до справжньої вартості (з диференціацією кожного наступного етапу інвестування). Так само повинна призводитися до реальної вартості і сума чистого грошового потоку (за окремими етапами його формування).

4. Вибір дисконтної ставки у процесі наведення окремих показників до цієї вартості має бути диференційований щодо різноманітних проектів. У процесі такої диференціації мають бути враховані рівень ризику, ліквідності та інші індивідуальні характеристики реального інвестиційного проекту.

З урахуванням вищезазначених принципів розглянемо методи оцінки ефективності реальних проектів на основі різних показників.

Чистий наведений доход [Net Present Value, NPV] дозволяє отримати найбільш узагальнену характеристику результату інвестування, тобто його кінцевий ефект у абсолютній

сумі. Під чистим наведеним доходом розуміється різниця між наведеними до цієї вартості сумою чистого грошового потоку за період експлуатації проекту PV та сумою витрат на його реалізацію INV. Розрахунок цього показника при одноразовому здійсненні витрат здійснюється за такою формулою (3):

$$NPV = \sum_{k=1}^n \frac{CF_k}{(1+r)^k} - INV \quad (3)$$

де $CF_1, CF_2, CF_k, \dots, CF_n$ – це річні грошові надходження протягом n -років,

n – кількість періодів інвестування,

k – номер періоду вкладення,

INV – початковий вклад,

r – ставка дисконтування,

$PV = \sum_{k=1}^n \frac{CF_k}{(1+r)^k}$ – загальна накопичена величина дисконтованих надходжень.

Якщо проект передбачає не разову інвестицію, а послідовне фінансування ресурсів протягом кількох років (m -років), то формула розрахунку NPV модифікується (4):

$$NPV = \sum_{k=1}^n \frac{CF_k}{(1+r)^k} - \sum_{j=1}^m \frac{INV_j}{(1+i)^j} \quad (4)$$

де INV_j – це інвестиції щодо окремих інтервалів загального періоду реалізації проекту,

i – прогнозований середній рівень інфляції.

Очевидно, що якщо $NPV > 0$, то проект слід ухвалити; при $NPV < 0$ проект має бути відкинута, як збитковий; за $NPV = 0$ проект не прибутковий, але й збитковий. Незалежні проекти з позитивним значенням показника чистого наведеного доходу дозволяють збільшити капітал підприємства та його ринкову вартість. Із системи взаємовиключних інвестиційних проектів приймається той із них, за яким значення показника чистого наведеного доходу є найвищим.

При прогнозуванні доходів за роками необхідно враховувати всі види надходжень як виробничого, так і невиробничого характеру, які можуть бути асоційовані з даним проектом. Так, якщо після закінчення періоду реалізації проекту планується надходження коштів у вигляді ліквідаційної вартості обладнання або вивільнення частини оборотних коштів, вони мають бути враховані як доходи відповідних періодів.

У роботі було розглянуто розширений метод аналізу витрат і доходів для управління якістю, існують різні причини, які впливають на якість проекту. Ці причини мають різні наслідки та зовсім по іншому впливають на прибутки. Усі розглянуті показники оцінки ефективності реальних проектів перебувають між собою у тісному взаємозв'язку та дозволяють оцінити цю ефективність із різних сторін. Тому при оцінці ефективності реальних проектів їх слід розглядати у комплексі.

ЛІТЕРАТУРА

1. Разу М. Л. Управління проектом. Основи проектного управління: учебник. КНОРУС, 2006. 5-е изд., 768 с.
2. Горбовцов Г.Я. Управління проектом: Навчально-практичний посібник. – М.: Изд. центр ЕАОИ, 2007. – 279 с.
3. Мілошечич Д. М60 Набір інструментів для управління проектами/ Драган З. Мілошечич; Пер. с англ. Мамонтова Е.В.; Під ред. Невідомого С.И. . – М.: Компанія Айтї; ДМК Пресс, 2008. – 729 с.: ил.
4. Управління інноваційними проектами і програмами: навчальний посібник / В.В. Биковський, Е.С. Міщенко, Е.В. Биковська и др. – Тамбов : Видавництво ГОУ ВПО ТГТУ, 2011. – 104 с.
5. Кліффорд Ф. Грей, Эрик У. Ларсон. Управління проектами: Практичний посібник/ Пер с англ. – М.: Видавництво «Діло и Сервіс», 2003. – 528 с.

КОВАЛЕНКО Анастасія Василівна – кафедра прикладної і теоретичної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: xa11867840@student.karazin.ua; ORCID 0000-0002-5814-5371.

БУЛАВІН Дмитро Олексійович – к. т. н., доцент кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна; м. Харків, майдан. Свободи, 6; 61022; e-mail: d.bulavin@karazin.ua, ORCID 0000-0002-4840-4763.

Наукові інтереси:

– застосування методів теорії ухвалення рішень, штучні нейронні мережі, генетичні алгоритми, управління проектами.

КОНЮШЕНКО Р.В., СВАТОВСЬКИЙ І.І.

ДОСЛІДЖЕННЯ ЗАХИЩЕНОСТІ КАНАЛІВ ПЕРЕДАЧІ ДАНИХ І УПРАВЛІННЯ БЕЗПІЛОТНИМИ ЛІТАЛЬНИМИ АПАРАТАМИ

Вступ

БПЛА – Безпілотний літальний апарат літакового чи гелікоптерного плану без екіпажу на борту, політ якого здійснюється або за по завчасно змодельованим програмам або шляхом віддаленого управління по різних каналах зв'язку [1].

На сучасному етапі розвитку БПЛА призначені для вирішення широкого спектру задач, наприклад, спостереження (розвідка), нанесення ударів, транспортування вантажів, цілевказання, ретрансляція даних при їх дистанційним управлінні оператором, чи шляхом автономних дій по завчасно закладеним програмам.

БПЛА це достатньо прогресивна платформа для успішних завдань, і в майбутньому вони лише будуть розвиватися і прогресувати у своїй конструкторських ідеях та являється однією з передових технологій майбутнього без якої неможливо уявити майбутнє, а саме силових структур різних відомств.

Метою цієї роботи є дослідження захищеності каналів передачі даних і управління БПЛА, визначення вразливих місць у методах шифрування даних та забезпечення надійного захищеного зв'язку з відповідними літальними апаратами.

Система радіозв'язку

Система радіозв'язку БПЛА представляє собою сукупність різних ліній, які в свою чергу використовуються для передачі великих за об'ємом даних телеметричного обладнання або простих, але важливих даних управління.

Тому інформацію для передачі через лінію зв'язку поділяють за об'ємом даних, рівень важливості, рівня крипто захисту.

Найпоширенішим рішенням для управління і обміну інформації з БПЛА є створення так званих напрямків важливості цієї інформації:

- Напрямок вгору визначається від пункту управління до БПЛА і включає в себе передачу команд управління польотом, та командами управління спец апаратурою і засобами корисного навантаження, що розміщуються на БПЛА

- Напрямок вниз визначається власне БПЛА до пункту управління і включає в себе телеметричну інформацію підсистем БПЛА та ці ж дані, що характеризують спец апаратуру, наприклад: кут нахилу польоту, температуру систем, завантаженість обчислювальної техніки, підтвердження виконання команд і тому подібне [3, 4].

А також високошвидкісна лінія передачі даних від спец апаратури і технічних засобів і контролерів корисного навантаження.

Вищезгадані лінії можуть відрізнятися і методами побудови, наприклад, кожан лінія може організовуватися в своєму окрему частотному діапазоні, або ж навіть використовувати різні режими з ретрансляцією або ж її відсутністю, зазвичай, так чи інакше це залежатиме від типу і важливості цих даних.

Різними службами різних країн і різних сценаріїв використання БПЛА з різними методами зв'язку, тож спираючись на малу культуру використання БПЛА країн агресорів зазначених в доктрині про безпеку України будемо спиратися на найпростіші методи.

Використання всієї ширини радіочастот при прямій видимості від пункту управління до БПЛА, використання інших літальних апаратів як ретрансляторів між БПЛА та пунктом управління, а також використання систем супутникового зв'язку.

Через використання країною агресором своїх БПЛА на не підвладній території та високій ймовірності втрати малих БПЛА пропустимо використання технологій зв'язку доступних

цивільним особам, наприклад: мережі мобільних операторів, мережі Bluetooth та Wi-Fi і тому подібне.

При неотриманні команд управління польотом від пункту управління, як вже згадувалося раніше, в основній своїй масі БПЛА переходять в автономний режим, де можуть бути реалізовані як найпростіші алгоритми, такі як повернення додому, прямолінійний політ, баражування, політ на малих висотах з імітацією польоту великих птахів, так і більш складних (в залежності від наявності в пам'яті електронних карт місцевості або інших навігаційних систем на БПЛА).

Система управління

За способами управління БПЛА поділяються на:

- автономні;
- напівавтономні;
- керовані.

В залежності від складності і специфіки виконання завдань, що ставляться перед БПЛА, обирається один з способів управління. Зазвичай, великі БПЛА типу «літак» в загальних принципах і ієрархіях літальних апаратів відповідає таким як пілотовані, на відміну від малих, які в свою чергу мають трьохрівневе ієрархічне компонування систем управління:

- 1) Нижній – рівень, що відповідає за окремі пристрої, механізми, датчики, сканери і тому подібне.
- 2) Середній – рівень, що відповідає власне за сам процес польоту спираючись на бортові контролери та модулі, які відповідають за їхню комунікацію, ввід та вивід сигналів управління.
- 3) Верхній – рівень інтерфейсу та управління безпосередньо оператором БПЛА, де зв'язок будується між ним та програмою польоту через інтерфейс, що зв'язує ще й середній рівень через його контролери [2, 3].

Нижній рівень управління утворюється за допомогою установки, що приводить всю конструкцію у рух, датчиків навігаційних систем, додаткове обладнання, перспективне для виконання певних цілей і завдань, таке як радіолокаційні станції, обладнання спостереження і тому подібне.

Середній рівень, зазвичай, відповідає бортовим механізмам та засобам управління. Бортова система управління малих БПЛА розміщується на бортових обчислювальних машинах, які в свою чергу управляються відкритими або спеціалізованими операційними системами, останні в свою чергу все більше і більше вибирають нішу операційних систем для БПЛА та вже створюються лише для цього.

На сьогодні існує багато рішень, де на одній платі розміщується і сама обчислювальна техніка і контролери (висоти, температури, вологості і тому подібне). В свою чергу ця плата розміщується у відповідний відсік, корпус або ж то навіть в найпримітивніших моделях взагалі ніяк не захищається від зовнішніх подразників і інших непереборних сил. Найпопулярнішим і найдешевшим з таких на ринку є системні плати на платформі Rasbery, де на заводі виготовлювачі буде повішено обладнання відповідне від побажань замовника, мається на увазі процесор, що відповідає потребам відповідної операційної системи та її обчислювальним завданням, яка гратиме лише роль інтерфейсу між приймачем радіозв'язку та іншими платами, які замовник додає відповідно окремо, які в свою чергу створюються вже під усіма грифами секретності відповідного виготовлювача та виконують запрограмоване для них завдання.

Верхній рівень управління відповідає архітектурі «безпілотник – пункт управління», що утворюється телеметричним обладнанням, системою збору даних про стан апарату, системою зв'язку та власне сам пункт управління та його обладнання, де обробляються телеметрія, визначається стан та місцезнаходження, модулюються або змінюється програма польоту і відповідно від попередніх даних система повертає команди назад на борт БПЛА.

Найбільш критичним показником для вирішення задач протидії БПЛА є технічні характеристики найслабшої ланки літального апарату, і такою є система зв'язку.

Також потрібно відмітити, що переважна кількість БПЛА не мають можливості взагалі на будь яке прийняття рішень щодо ризиків, загроз, різких змін рельєфу, погоди і тому подібне. Зазвичай вони мають просто якийсь надісланий з пункту управління маршрут він може бути не обов'язково прямим, але дуже часто не враховуються жодні непереборні сили, такі як рельєф,

флора та фауна тих територій на яких відбувається політ, а також антропогенних об'єктів: лінії електропередачі, будинки так їхні надбудови, інші літальні апарати.

Тому найпрогресивнішим зараз завданням перед передовими виробниками являється на збалансованість між можливістю управління БПЛА, за допомогою штучного інтелекту, швидкодії реагування на загрози, та дешевизною використання в апаратах, які достатньо ймовірно так чи інакше буде втрачений під час виконання завдань.

Найпростіший методі придушення засобами РЕП

При підготовці до придушення каналів навігації БПЛА зазвичай спираються на те, наскільки складна система управління та на те на яку кількість датчиків спирається навігаційна система та власне на природу черпання інформації цих датчиків:

- Навігаційна система, що заснована тільки на апаратурі користувачів найбільш розповсюджених систем керування – така система характерна для найменших і найпростіших БПЛА – квадрокоптерів.
- Проста інтегрована навігаційна система, що заснована на комплексуванні даних різних механічних інерційних навігаційних систем – така система характерна для великого спектру малих БПЛА для професійного використання в різних сферах людського життя.
- Інтегрована навігаційна система, що заснована на комплексуванні даних декількох навігаційних пристроїв: механічних інерційних навігаційних пристроїв, барометричних висотомірів, радіо чи лазерних висото метрів – характерна як для БПЛА малого і середнього класу.
- Інтегрована навігаційна система, що заснована на основі комплексуванні даних декількох навігаційних пристроїв: професійних авіаційних інерційних систем навігації, комплексів різних висотомірів, VOR/DME (Very high frequency Omni directional radio Range / Distance Measuring Equipment), що все в синхронній і злагодженій роботі дає результат на рівні звичайних пілотованих літальних апаратів і характерна для великих і тяжких БПЛА [2].

Також, говорячи про спроби придушення каналів навігації треба пам'ятати, що так чи інакше факт будь якої дії на цей канал передачі інформації між системами управління (системами навігації безпосередньо на літальному апараті, системами, що отримують навігаційні дані від пунктів управління і тому подібне) можливий лише за використання БПЛА з великої кількості передачі даних лише одного каналу для навігації, адже у більш складних системах, де використовують декілька систем навігації, можлива перебудова управління в основі якого буде найменш вразлива для даного сценарію протидії засобами РЕП.

Також, потрібно зауважити, що у більш складних системах управління на великих чи навіть середніх БПЛА використовуються навігаційні дані з інерційно навігаційних систем авіаційного типу, що базуються на лазерних чи волоконно оптичних гіроскопах, що власне взагалі може виключити канал передачі навігаційних систем з можливих для впливу засобами РЕП.

Потрібно зазначити, що в порівнянні з завданням радіоелектронного придушення навігаційної системи БПЛА, завдання придушення радіоліній ПУ - БПЛА і БПЛА - ПУ не є принципово новою і фактично зводиться до формування на вході приймача, що придушуються, такого сигналу, який просто не дозволить забезпечити прийом даних з потрібною завадостійкістю і теоретичною достовірністю. Дана задача є класичною в теорії РЕП, а вирішуються простим використанням у радіолініях типів сигнально-кодових конструкцій, типів переданих даних (тип переданих даних визначає необхідний рівень достовірності прийому), а також сигнальних, енергетичних, просторових та інших параметрів радіоліній.

Висновки

Як висновок, можна стверджувати, що БПЛА залежне від зовнішніх джерел інформації, чи то дані управління польотом чи то дані навігаційної системи, тому як найслабшу ланку було визнано власне радіозв'язок та навігаційну систему літального апарату, тому вони найчастіше піддаються зовнішньому втручанню та потребують максимальну кількість акумулювання технологій, які б могли б вберегти БПЛА від непереборних сил, такі як автономний рух в межах відсутності зв'язку зі станціями керування, чи то навіть нового підходу до побудови каналів зв'язку як мобільні станції.

Рекомендаційними діями для підвищення захищеності каналів зв'язку БПЛА є використання інших літальних апаратів з направленим радіосигналом, які буде важче зафіксувати в тих межах, де вони будуть працювати, або ж використання радіоефіру таким чином, щоб радіосигнал не перевищував «шум» в радіоефірі, також при великому ризику пеленгування радіосигналу рекомендовано взагалі позбутися будь якого витоку радіосигналів, а отже БПЛА мусить працювати автономно під командами штучного інтелекту, але подібні системи керування дорого вартісні і потребують великої кількості датчиків і обчислювальних ресурсів, тому літальні апарати такого типу, будуть здебільшого дорого вартісні та великі за розміром, але дуже високоефективні. Поєднавши всі ці рекомендаційні дії, можна отримати БПЛА максимально захищений від впливу зовнішніх непереборних сил, які націлені на відрізання від пункту управління або ж на перехоплення управління над БПЛА.

ЛІТЕРАТУРА

1. Special'naya radiosvyaz'. Razvitie i modernizaciya oborudovaniya i ob'ektov (Special radio communication). Pod red. A. L. Buzova, S. A. Bukashkina. Moscow: Radiotekhnika. 2017. 448 p.
2. Sovremennye tekhnologii radiomonitoringa v sputnikovyh sistemah svyazi i retranslyacii (Modern technologies of radio monitoring in satellite communication and retransmission systems). Pod red. A. V. Kuzovnikova. Moscow: Radiotekhnika. 2015. 216 p.
3. Barnard J. Small UAV command-control and communication issues// IEEE on communicating with UAV's. 2007. P. 75-85.
4. Haas E. Aeronautical channel modeling// IEEE Transactions on Vehicular Technology. 2002. V. 51. No 2. P. 254-264.

КОНЮШЕНКО Ростислав Вікторович – студент кафедри безпеки інформаційних систем і технологій; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: xa11867677@student.karazin.ua; ORCID: 0000-0001-9594-7798.

Наукові інтереси:

– апаратне складова комп'ютерних систем, процедурне програмування, алгоритмізація.

СВАТОВСЬКИЙ Ігор Іванович – доцент кафедри безпеки інформаційних систем і технологій, кандидат технічних наук, старший науковий співробітник; Харківський національний університет імені В. Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61000; e-mail: i.svatowsky@karazin.ua; ORCID: 0000-0002-1836-5599.

Наукові інтереси:

– завадостійке кодування, завадостійкі інформаційно-телекомунікаційні системи, методи підвищення захищеності передачі цифрової інформації по радіоканалах

ВИКОРИСТАННЯ МЕТОДІВ СТАТИСТИЧНОГО АНАЛІЗУ В РЕКОМЕНДАЦІЙНИХ СИСТЕМАХ

Recommender systems is a separate subclass of software systems in data science that allow to process, analyze user data, while finding new relations between user data and actions, as well as supposing personal recommendations for an individual user. The development of recommender systems is combined with the use of additional similarity features, clustering methods and artificial intelligence models. A scoring matrix is used to display user ratings and further analysis. It is possible to analyze user ratings using similarity functions. Similarity function is an important component of the recommendation process which provides an opportunity to evaluate the measure of the relation between objects. In addition, similarity functions are most often represented by a set of different functions, the choice of which depends on the specific conditions of the problem. Examples of the main similarity functions used in recommender systems are given. An example of the application of similarity functions is collaborative filtering. Collaborative filtering suggests, based on user ratings, what a particular product's score will be from a particular user. This approach includes two methods, such as user-based collaborative filtering and item-to-item based collaborative filtering. This article gives an example of the use of similarity functions in the collaborative filtering method in practice, using an appropriate rank matrix. This paper also contains all the necessary formulas that are used during data analysis. The results are provided in the text and the advantages and disadvantages of each approach are also mentioned. Conclusions are drawn regarding the use of these methods in recommender systems.

Ключові слова: рекомендаційна система, функція подібності, оцінка користувача, спільна фільтрація.

Рекомендаційна система – це окремий клас програмних систем, бізнес-логіка якої полягає у підборі та пропонуванні користувачеві відповідних релевантних рекомендацій, що ґрунтуються на власних знаннях щодо користувача та результатах його взаємодії з програмною системою. Релевантними можна вважати такі рекомендації, що розташовується у порядку того, що більше підходить користувачу в даний конкретний момент часу. Релевантність включає в себе користувацькі дані, відповідний контекст і вірогідні або отримані оцінки.

Користувацькі оцінки є одними із основних видів даних, які використовуються у даній програмній системі. Їх досить зручно та ефективно зберігати у вигляді декартового добутку або у вигляді матриці, у рядках якої зазначені користувачі, а в стовпчиках – товари. Таке відображення дозволяє дослідити ставлення кожного з користувачів до відповідного товару, що виражено оцінкою. Приклад такої матриці наведено у Таблиці 1 [1, с. 183].

Табл. 1 – Приклад матриці оцінок користувачів

	Товар 1	Товар 2	Товар 3
Клієнт А		2	3
Клієнт В	4	5	2
Клієнт С	3	4	5
Клієнт D	2	4	5
Клієнт E	1	5	4

Відмінності та подібності знаходяться у обернено пропорційному співвідношенні, тобто зі збільшенням відмінностей збільшується подібність, а зі зменшенням відмінностей подібність збільшується. Функція подібності являє собою мірою того, наскільки схожі два об'єкти, що вимірюються. Нехай є 2 елементи k_1 та k_2 , подібність яких визначимо як міру функції $\text{sim}(k_1, k_2)$, яка, в свою чергу, може бути представлена як коефіцієнт Жаккара, Серенсена, коефіцієнт кореляції Сімпсона або косинусна подібність. Найбільше у рекомендаційних системах застосовуються саме дві останні міри. [2, с. 60]

Коефіцієнт кореляції Сімпсона може бути розрахований за наступним співвідношенням:

$$\text{sim}(x, y) = \frac{\sum_{i=1}^p (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^p (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^p (y_i - \bar{y})^2}} \quad (1)$$

де $\bar{x}$ – середнє арифметичне значення вибірки,

x_i – значення i -го елементу вибірки x ,

y_i – значення i -го елементу вибірки y .

Використовуючи косинусну подібність, можна розглядати вибірки даних як окремі вектори, між якими розглядається міра кута, що вказує на подібність або відмінність значень вибірки. Косинусна подібність розраховується із використанням наступної формули:

$$\text{sim}(x, y) = \cos(\theta) = \frac{x \cdot y}{\|x\| \cdot \|y\|} = \frac{\sum_{i=1}^n x_i \cdot y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (2)$$

де x_i – значення i -го елементу вибірки x ,

y_i – значення i -го елементу вибірки y .

Дані функції подібності найбільше застосовуються у алгоритмах спільної фільтрації, суть яких полягає у знаходженні схожих користувачів і рекомендації того, що подобається схожим користувачам. У даному підході головний акцент зроблено на класифікації користувачів за схожими типами, відповідно до уподобань яких кожен користувач отримує особисту рекомендацію. Спільна фільтрація поділяється на два підвиди, а саме на основі користувачів та на основі об'єктів. [3]

Спільна фільтрація на основі користувача являє собою метод, що полягає у прогнозуванні елементів, котрі можуть сподобатись користувачу, базуючись при цьому на оцінках, які залишили інші користувачі зі схожими з цільовим користувачем уподобаннями.

Спільна фільтрація на основі товарів є методом дослідження взаємозв'язку між парами товарів і дозволяє розрахувати відсутню оцінку за допомогою оцінок, які користувач поставив іншим товарам. [4, 5]

Розглянемо приклади використання функцій подібності у спільній фільтрації. Спочатку розглянемо спільну фільтрацію на основі товарів. У Таблиці 1 відсутнє значення оцінки Клієнта А для Товару 1. Теоретично розрахуємо можливе значення даної оцінки, використовуючи функцію оцінки та наступну формулу:

$$m(C, I_i) = \frac{\sum_j m(C, I_i) \cdot s_{ij}}{\sum_j s_{ij}} \quad (3)$$

де $m(C, I_i)$ – оцінка C -го користувача i -го товару,

s_{ij} – значення функції подібності.

Відповідно до даних з Таблиці 1:

$$m(A, I_1) = \frac{2 \cdot 0.9073 + 3 \cdot 0.922}{0.9073 + 0.922} = 2.5$$

Розглянемо спільну фільтрацію на основі користувачів для цього ж самого випадку, однак тепер залежність буде розглядатись саме між користувачами, а не товарами. Теоретично розрахуємо вірогідне значення оцінки із використанням функції оцінки за наступною формулою:

$$m(C, I_i) = \bar{m}(C) + \frac{\sum_{i \in U} \text{sim}(C, i) \cdot m(i, I_i)}{\sum_{i \in U} |\text{sim}(C, i)|} \quad (4)$$

де $m(C, I_i)$ – оцінка C -го користувача i -го товару,

$\bar{m}(C, I_i)$ – середня оцінка C -го користувача усіх товарів,

$\text{sim}(C, i)$ – значення функції подібності.

Відповідно до даних з Таблиці 1, отримаємо:

$$m(C, I_i) = 2.5 + \frac{-0.3273 - 0.5 - 0.5455 - 0.5604}{|-0.982| + |0.5| + |0.3273| + |-0.24|} = 2.1$$

Отже, у результаті проведення розрахунків, значення можливої оцінки Товару 1, яку залишить Клієнт А, за спільною фільтрацією на основі товарів дорівнює 2.5, тобто оцінка однаково близька як до 2, так і до 3, тоді як за фільтрацією на основі користувачів становить 2.1, що наближає оцінку ближче до 2. Невелика розбіжність вірогідної оцінки пов'язана із аналізом зв'язків між товарами за першим методом, і аналізом зв'язків між користувачем за другим методом. У обох підходах використовуються функції подібності. Вибір метода напряму залежить від кількості оцінок, залишених користувачами, так як за умови холодного старту, тобто відсутності оцінок нових користувачів, рекомендаційній системі буде досить складно запропонувати їм рекомендацію. Також береться до уваги той факт, що не всі користувачі завжди залишають хоч якусь оцінку, що ускладнює процес генерування рекомендацій. Через дані можливі недоліки перевагу в використанні має метод спільної фільтрації на основі товарів.

ЛІТЕРАТУРА

1. Kim Falk Practical Recommender Systems. Manning Publications Co., 2019, 432 с.
2. Charu C. Aggarwal Recommender Systems: The Textbook. Springer International Publishing Switzerland, 2016, 518 p.
3. Introduction to Recommender Systems [Електронний ресурс]. – Режим доступу: <https://towardsdatascience.com/introduction-to-recommendersystems-6c66cf15ada>.
4. Collaborative Filtering – ML [Електронний ресурс]. – Режим доступу: <https://www.geeksforgeeks.org/collaborative-filtering-ml/?ref=lbp>.
5. Frank Kane Building Recommender Systems with Machine Learning and AI. National Publishing United Kingdom, 2020, 439 p.

КОСТЮЧЕНКО Артем Дмитрович – здобувач вищої освіти, студент групи КС-42; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: xa12850455@student.karazin.ua; ORCID: 0000-0002-3940-8797.

ЧУБ Ольга Ігорівна – к.е.н., доцент кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: o.i.chub@karazin.ua; ORCID: 0000-0002-1216-856X.

КУЛИК В.М, БАКУМЕНКО Н.С.

ПРОГНОЗУВАННЯ НАВАНТАЖЕНЬ В КОМП'ЮТЕРНИХ МЕРЕЖАХ ЗА ДОПОМОГОЮ АВТОРЕГРЕСІЙНИХ МЕТОДІВ

The main theme is the forecasting of loads in computer networks using autoregressive methods.

The purpose of the work is to get acquainted with the basic concepts related to computer networks, analyze the concept of monitoring, identify its goals and tasks, and distinguish the concept of monitoring loads in computer networks. Also, the goal was to get acquainted with the methods of monitoring traffic and loads in computer networks and learn about the existing autoregressive models for forecasting time series.

The analysis of autoregression models, the concept of monitoring, its organizational functions, tasks, and goals, as well as the analysis and comparison of methods of monitoring traffic and loads in computer networks, were used as **research methods**.

The results of the report are the systematization of information, which is the basis for further analysis of the topic and the identification of goals, tasks, and methods of monitoring loads in computer networks. Also, 4 autoregressive models that can analyze time series with data for load forecasting and prevention are reviewed.

In a conclusion, it can be said that monitoring loads in computer networks is a necessary measure that will allow forecasting and prevention of overloads. Otherwise, users and administrators of the computer network will observe a drop in bandwidth, which will lead to slower work. Forecasting can be done using time series analysis methods, namely autoregression models, which allow the efficient description of stationary stochastic processes.

Ключові слова: комп'ютерні мережі, навантаження в комп'ютерних мережах, методи прогнозування, авторегресійні методи, моніторинг трафіку.

Моніторинг навантажень у комп'ютерних мережах

Постійне збільшення користувачів послуг високошвидкісного інтернету створює проблеми інтернет провайдерам, адже зростаюча кількість підключених користувачів підвищує навантаження на сервери.

Через це, провайдери вимушені займатися прогнозуванням завантаженості мереж, адже своєчасне виявлення потенційної проблеми дозволить уникнути неминучих збоїв у роботі. Для цього, фахівці використовують різні алгоритми прогнозування, які базуються на вхідних даних вибірки, що формується з часових рядів.

Перевантаження в комп'ютерних мережах з'являється тоді, коли мережа передає кількість даних, яка майже досягає або дорівнює граничному показнику пропускної здатності. Відповідно, наслідком перевантажень є зниження пропускної здатності мережі.

Для того, щоб ефективно відслідковувати потенційні перевантаження в комп'ютерних мережах, спеціалісти використовують різні методи та засоби моніторингу перевантажень. Перш ніж ознайомитися з ними, слід визначити що являє собою поняття моніторингу. Гузій М. М. зазначає, що «це систематичний збір і обробка інформації, яка може бути використана для поліпшення процесу ухвалення рішення, а також побічно для інформування громадськості або прямо, як інструмент зворотного зв'язку в цілях здійснення проектів, оцінки програм або вироблення політики» [1]. Процес моніторингу має одну або більше функцій, які зазначені на рисунку 1.

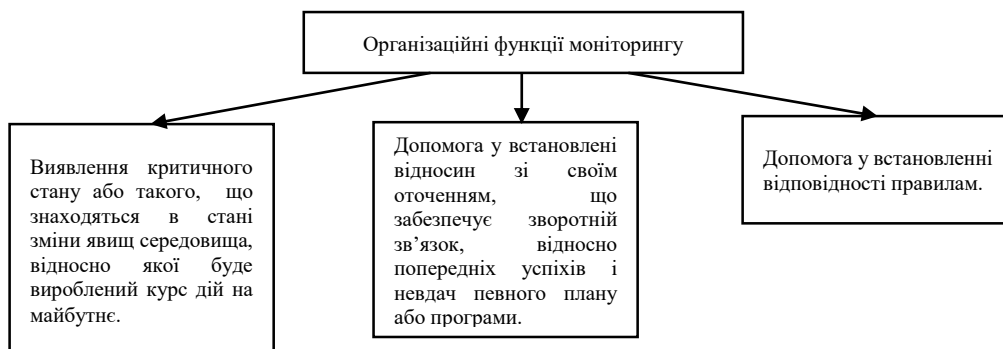


Рис. 1 Організаційні функції моніторингу

В контексті комп'ютерних мереж, поняття моніторингу має дещо інше трактування. Вокорокос Л. визначив його наступним чином: «це робота системи, яка виконує постійне спостереження за комп'ютерною мережею у пошуках повільних або несправних систем і яка, при виявленні збоїв, повідомляє про них мережевого адміністратора» [2]. Також, моніторинг комп'ютерних мереж переслідує дві основні цілі, кожна з яких відноситься до перевантажень:

- виявлення перевантажень;
- виявлення аномалій.

Також, моніторинг комп'ютерних мереж повинен виконувати наступні завдання:

- контроль завантаження каналів зв'язку;
- виявлення тонких місць мережі, неоптимальності зміни мережевого обладнання;
- прогнозування стану мережі.

Слід враховувати всі вищезазначені завдання, адже результати їх виконання допомагають керувати станом роботи системи в конкретний проміжок часу та прогнозувати поведінку у майбутньому.

Методи та засоби моніторингу комп'ютерних мереж

Методи моніторингу перевантажень у комп'ютерних мережах можна поділити на дві групи: пасивні та активні. Перша група характеризує моніторинг, під час якого програми або мережеве обладнання періодично надсилають дані про свій стан монітору, який зберігає їх для подальшого використання. Прикладом є протокол SNMP. Друга група характеризує моніторинг, під час якого монітору треба надіслати запити програмам або мережевому обладнанню та отримати на них відповіді перш ніж він отримає дані. Прикладом є протокол ICMP. В таблиці 1 зазначені переваги та недоліки пасивного і активного методів моніторингу комп'ютерних мереж.

Табл.1 Переваги та недоліки пасивних та активних методів моніторингу трафіку і КМ

Переваги		Недоліки	
Пасивний	Активний	Пасивний	Активний
більша кількість показників створює велику інформативність моніторингу	точний підхід, який спрямований на різні об'єкти	несумісність деякого обладнання зі стандартами SNMP	може створювати велику завантаженість на мережу
створює менше навантаження на мережу	можливість виконувати моніторинг без доступу до обладнання	необхідність наявності адміністративного доступу до програм або мережевого обладнання для конфігурування параметрів моніторингу	через меншу кількість показників має меншу інформативність моніторингу
	малі інтервали семплювання	великі інтервали семплювання	брандмауери можуть впливати на процес моніторингу

Огляд авторегресійних моделей

Загальну модель авторегресії – ковзного середнього вперше описав Пітер Уїтл в дисертації «Перевірка гіпотез в аналізі часових рядів» в 1951 році. Потім, Джордж Бокс та Гвілім Дженкінс популяризували її в 1970 році в спільній книзі. Загалом, авторегресійні моделі можна поділити на 4 види:

- ARIMA-базовані моделі;
- моделі ковзного середнього порядку q (MA(q)-моделі);
- моделі авторегресії порядку p (AR(p)-моделі);
- авторегресійні моделі зі ковзними середніми в залишках (ARMA(p, q)-моделі).

Розглянемо їх більш детально.

ARIMA-базовані моделі мають наступні етапи побудови:

1. ідентифікація пробної моделі;
2. оцінювання параметрів моделі та діагностичну перевірку адекватності моделі;
3. використання моделі для прогнозування [3].

Ця модель має наступний характерний запис:

$$\varphi(B)(1-B)^d X_t = \theta(B)\varepsilon_t, \quad (1)$$

де $\varphi(\bullet)$, $\theta(\bullet)$ є поліномами степеню p та q ;

B – лаговий оператор:

$$(B^j X_t = X_{t-j}, B^j \varepsilon_{t-j}, j = 0, \pm 1, \dots), \quad (2)$$

d – порядок взяття послідовної різниці:

$$(\Delta X_t = X_{t-1} - X_t = (1-B)X_t, \Delta^2 X_t = \Delta^2 X_{t+1} - \Delta X_t = (1-B)^2 X_t, \dots) [4], \quad (3)$$

Моделі ковзного середнього порядку q (MA(q)-моделі) можна записати наступним чином:

$$\varepsilon_t = \sum_{j=1}^{\infty} \pi_j \varepsilon_{t-j} + \delta_t, \quad (4)$$

де коефіцієнти π_j ($j = 1, 2, \dots, j$) певним чином виражаються через параметри $\theta_1, \dots, \theta_q$;

символи $\theta_1, \dots, \theta_q$ використовуються для позначення кінцевого набору параметрів β ;

δ_t – послідовність випадкових величин, що утворює білий шум.

Умова оборотності MA(q)-моделі (тобто умова збіжності ряду $\sum_{j=1}^{\infty} \pi_j$) формується в термінах характеристичного рівняння моделі (4) в наступний спосіб. Усі корені характеристичного рівняння

$$1 - \theta_1 z - \theta_2 z^2 - \dots - \theta_q z^q = 0 \quad (5)$$

повинні лежати поза одиничним колом, тобто $|z_j| > 1$ для всіх $j = 1, 2, \dots, q$ [5].

Моделі авторегресії порядку p (AR(p)-моделі), де $p \geq 3$ утворюють підмножини у класі загальних лінійних моделей, самі становлять досить широкий клас моделей. Якщо в загальній лінійній моделі вважати всі параметри π_j , крім перших p коефіцієнтів, рівними нулю, ми приходимо до визначення AR(p)-моделі:

$$\varepsilon_t = \sum_{j=1}^p \alpha_j \varepsilon_{t-j} + \delta_t, \quad (6)$$

де α – деякий числовий коефіцієнт, що не перевищує абсолютну величину одиниці ($|\alpha| < 1$);

δ_t – послідовність випадкових величин, що утворює білий шум [5].

Авторегресійні моделі зі ковзними середніми в залишках (ARMA(p, q)-моделі)

Для отримання економічної параметризації іноді буває доцільно включити в модель як члени, що описують авторегресію, так і члени, що моделюють залишок у вигляді ковзного середнього [5]. Такі лінійні процеси записуються наступний чином:

$$\varepsilon_t = \alpha_1 \varepsilon_{t-1} + \dots + \alpha_p \varepsilon_{t-p} + \delta_t - \theta_1 \delta_{t-1} - \dots - \theta_q \delta_{t-q} \quad (7)$$

та називаються процесами авторегресії – ковзного середнього порядку (p, q)(ARMA(p, q)). Далі буде розглянута стаціонарність та оборотність ARMA(p, q)-процесів [5]. Процес (7) можна також записати як:

$$\varepsilon_t = \sum_{j=1}^p \alpha_j \varepsilon_{t-j} + \bar{\delta}_{qt}, \quad (8)$$

де $\bar{\delta}_{qt} = \delta_t - \theta_1 \delta_{t-1} - \dots - \theta_q \delta_{t-q}$.

Висновок

Виходячи із вищезазначеної інформації, моніторинг навантажень в комп'ютерних мережах є необхідною мірою, яка дозволить прогнозувати та запобігати перевантаження. Прогнозування можна робити за допомогою методів аналізу часових рядів, а саме авторегресійних моделей, що дозволяють ефективно описувати стаціонарні стохастичні процеси.

ЛІТЕРАТУРА

1. Аналіз технологій моніторингу комп'ютерних мереж. Гузій М. М., канд. техн. наук, ДОЦ. Станіславова О. В. Кадет М.В URL: jrn1.nau.edu.ua/index.php/SBT/article/download/5091/5353 (дата звернення: 01.11.2022).
2. Vokorokos L., Adam N., Balarz A., Application of intrusion detection systems in distributed computer systems and dynamic networks, Computer Science and Technology Research Survey, 2008, pp. 19–24.
3. *Practical Time Series Analysis: Prediction with Statistics and Machine Learning*, Aileen Nielsen
4. Крюков Ю. А. ARIMA – модель прогнозування значень трафіка [Електронний ресурс] / Ю. А. Крюков, Д. В. Чернягин // 2. – 2011. – URL: http://www.isa.ru/jitcs/images/documents/2011-02/41_49.pdf (дата звернення 01.11.2022).
5. Бідюк, П. І. Аналіз часових рядів [Електронний ресурс] : навчальний посібник / П. І. Бідюк, В. Д. Романенко, О. Л. Тимошук ; НТУУ «КПІ». – Електронні текстові дані (1 файл: 8,42 Мбайт). – Київ : НТУУ «КПІ», 2010.
6. Box G.E.P., Jenkins G.M., Reinsel G.C. Time Series Analysis: Forecasting and Control. 3rd edition, Prentice Hall, 1999.

КУЛИК Вячеслав Миколайович – студент групи КІ-51 факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: slavakulik0104@gmail.com; viacheslav.kulyk@student.karazin.ua; ORCID: 0000-0002-0232-3351.

Наукові інтереси:

– аналіз та прогнозування часових рядів.

БАКУМЕНКО Ніна Станіславівна – к.т.н., доцент кафедри ТПС; факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: n.bakumenko@karazin.ua; ORCID: 0000-0003-3496-7167.

Наукові інтереси:

– інтелектуальний аналіз даних.

ЛАГУТІН А.Г., СТРИЛЕЦЬ В.Є.

СУЧАСНІ ЗАСОБИ І ТЕХНОЛОГІЇ СТВОРЕННЯ ВЕБ-САЙТІВ

Веб-сайт – це сукупність електронних документів приватної особи або організації, що об'єднана під однією адресою. Сукупність сайтів складає Всесвітню павутину.

Спочатку веб-сайти були лише збіркою статичних документів. На тепер сучасні сайти в більшості є динамічними та інтерактивними. Це стало можливим завдяки втіленню веб-застосувань – готового програмного комплексу для вирішення завдань веб-сайту.

Зазвичай, великі сайти використовують кілька доменів, щоб логічно відокремити різні види послуг (mail.google.com, news.google.com, maps.google.com). Іноді, окремі домени виділяють для іншої мовної версії. Наприклад, google.com і google.com.ua логічно є сайтом Google з різними мовами інтерфейсу, але технічно це різні сайти.

HTML – мова розмітки тексту

Всесвітня павутина складається з веб-сторінок, які створено у форматі HTML (HyperText Markup Language, «мова розмітки гіпертексту»). HTML – це фундаментальна, базова технологія Інтернету.

HTML не є мовою програмування, це мова розмітки тексту, що використовує спеціальні оператори – теги (tag) чи інша назва дескриптори (descriptor) для розмітки текстового документа. Ці позначки вказують в якому вигляді буде виведено текстовий чи інший елемент у вікні браузера.

HTML дозволяє формувати на сторінці сайту текстові блоки, додавати до них зображення, організовувати таблиці, керувати відтворенням кольору, додавати до дизайну сайту звуковий супровід, організовувати гіперпосилання з переходом до інших розділів сервера або звертатися до інших ресурсів Інтернету і компонувати всі ці елементи між собою. Документи, що створено лише засобами HTML мають розширення .htm або .html.

Однією з основних функціональних особливостей мови HTML, завдяки якій вона і отримала свою назву, є гіперпосилання.

Гіперпосилання (Hyperlink) — це базовий функціональний елемент HTML-документу, який реалізовує зв'язок певного об'єкту веб-сторінки з іншим об'єктом. Для гіперпосилання може використовуватися як фрагмент тексту, так і графічний об'єкт, а сам гіперзв'язок можна встановлювати як між об'єктами одного сайту, так і між об'єктами, що розміщені на різних сайтах Інтернету.

HTML є мовою, що лише інтерпретується, тому, для виконання коду, його не потрібно компілювати. Інтерпретатор мови втілено в браузер, і він «компілює» код безпосередньо під час відкривання документа. Якщо в коді сторінки виявлено помилку, інтерпретатор, зазвичай, не видає відповідного попередження, а просто ігнорує весь «помилковий» рядок, що може зіпсувати зовнішній вигляд завантаженої сторінки. Це є важливим для розробників, тому слід бути уважним під час складання HTML-коду і ретельно тестувати результати своєї роботи.

CSS – каскадна таблиця стилів

CSS (Cascading Style Sheets) — це технологія опису зовнішнього вигляду документа, що створено засобами HTML, XML і XHTML.

CSS використовується для завдання кольорів, шрифтів, розташування елементів сторінки тощо. До появи CSS оформлення веб-сторінок вказувалося безпосередньо в HTML-коді сторінки. Проте, з появою CSS стало можливим принципове розділення змісту і представлення документа. За рахунок такого нововведення стало можливим легке застосування єдиного стилю оформлення для кількох сторінок сайту, а також швидка зміна цього оформлення.

Переваги:

- Застосування кількох варіантів дизайну сторінки для різних пристроїв перегляду. Наприклад, для відображення на екрані монітора - дизайн буде розраховано на велику ширину. У разі друкування документа не буде роздруковане меню сайту, а у разі

перегляду у мобільному комп'ютері чи телефоні, меню буде виведено після вмісту сторінки.

- Зменшення часу завантаження сторінок сайту за рахунок перенесення правил представлення даних до окремого CSS-файлу. В цьому випадку браузер завантажує лише структуру документа і дані, що містяться на сторінці. CSS-файл з правилами представлення цих даних завантажується браузером лише один раз і зберігається в кеші браузера.
- Простота подальшої зміни дизайну. Не потрібно виправляти кожну сторінку, достатньо лише змінити кілька правил у CSS-файлі.
- Додаткові можливості оформлення. Наприклад, за допомогою CSS-правил можна застосувати обтікання певного блоку текстом або зробити так, щоб меню фіксовано знаходилося в певному місці при перегортанні сторінки.

Недоліки:

- Різні браузери можуть в різний спосіб інтерпретувати правила CSS, і відповідно, по різному відображати одні і ті ж фрагменти сторінки.

JavaScript – мова сценаріїв

JavaScript – це мова, що призначена для написання сценаріїв для інтерактивних HTML-сторінок. Мова JavaScript не має жодного відношення до мови Java. Java розроблено фірмою SUN, а JavaScript – фірмою Netscape Communication Corporation. Первинною назвою було «LiveScript», але, з огляду на велику популярність мови Java, назву «LiveScript» з комерційних міркувань було змінено на «JavaScript».

JavaScript не призначено для створення автономних застосунків. Програмний код на JavaScript вписується безпосередньо в текст HTML-документа і інтерпретується браузером в міру завантаження цього документа. За допомогою JavaScript можна динамічно змінювати текст завантаженого HTML-документа і реагувати на події, які пов'язані з діями відвідувача або змінами стану документа чи вікна.

Важливою особливістю JavaScript є об'єктна орієнтованість. Програмісту є доступними численні об'єкти, такі, як документи, гіперпосилання, форми, фрейми тощо. Об'єкти характеризуються описовою інформацією (властивостями) і можливими діями (методами).

PHP - мова створення сценаріїв

PHP (Personal Home Page) - мова створення сценаріїв, яка давно переросла свою назву. Перша версія PHP була створена Расмусом Лерддорфом в 1994 р. і була набором інструментів для відстеження поведінки відвідувачів сайту. З часом PHP з набору інструментів перетворилася на повноцінну мову програмування, а її назву було змінено як PHP HyperText Preprocessor (препроцесор гіпертексту PHP).

PHP - це серверна мова. Конструкції PHP, що вставлено в HTML-текст, виконуються сервером при кожному відвідуванні сторінки. Результат обробки конструкцій разом із звичайним HTML-текстом передається браузеру.

Основними конкурентами PHP є ASP (Active Server Pages) від компанії Microsoft і ColdFusion від компанії Allaire.

У порівнянні з ними PHP має ряд переваг, зокрема:

- Висока продуктивність. PHP-програми працюють швидше, ніж ASP.
- Функціональність. Розробку PHP-програми можна відокремити від власне розробки веб-сторінки, що спрощує працю і для програміста і для дизайнера.
- Ціна. Мова PHP є абсолютно безкоштовною.
- Простота у використанні. Програмісти, що мають досвід програмування на поширених мовах швидко зрозуміють синтаксис PHP.
- Переносимість. Один і той же PHP-код можна використовувати як в середовищі Windows, так і на платформах UNIX.

Аїах - технологія для взаємодії з сервером без перевантаження сторінок

Аїах (Asynchronous Javascript And XML) це підхід до створення веб-застосовань за допомогою наступних технологій:

- Стандартизоване представлення засобами XHTML і CSS.
- Динамічне відображення і взаємодія з користувачем за допомогою DOM.
- Обмін і обробка даних у вигляді XML и XSLT.
- Широке застосування мови сценаріїв JavaScript.
- Асинхронні запити за допомогою об'єкту XMLHttpRequest.

В стандартному веб-застосуванні обробкою всієї інформації займається сервер, браузер відповідає лише за взаємодію з користувачем, передачу запитів і виведення отриманих даних у форматі HTML.

В Аїах-застосуванні між користувачем і сервером з'являється ще один посередник – програмний механізм (рушій) Аїах. Він визначає, які запити можна обробити з боку клієнта, а які необхідно виконувати на сервері.

ЛІТЕРАТУРА

1. Стеценко І.В. Моделювання систем: навч. посіб. Черкаси: ЧДТУ, 2010. 399 с.
2. Олифер В., Олифер Н. Комп'ютерные сети. Принципы, технологии, протоколы : Учебник для вузов. Питер: 5-е изд. СПб, 2016. 992 с.
3. Лосев Ю. І., Руккас К. М., Шматков С. І. Комп'ютерні мережі: навч. посіб. / за редакцією Ю. І. Лосева. Харків : ХНУ імені В. Н. Каразіна, 2013. 248 с.
4. Ляшенко В.П., Кобильська О.Б., Дям'янченко О.П. Математичні моделі теплообміну з умовами імпедансного типу у багатошарових областях. *Вісник Кременчуцького національного університету імені Михайла Остроградського*. 2017. Вип. 6/2017 (106). С. 37–43.
5. Боков І. П., Бондаренко Н. С., Стрельнікова О. О. Дослідження поведінки узагальнених переміщень, отриманих із використанням теорії $\{m,n\}$ -апроксимації. *Вісник Харківського національного університету ім. В.Н. Каразіна серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління»*. 2018. Вип. 38. С. 14–24.
6. Курило І. А., Грудська В. П., Спінул Л. Ю., Щерба М. А. Розрахунок перехідних процесів у лінійних електричних колах із зосередженими і розподіленими параметрами : навч. посіб. Київ: НТУУ “КПІ”, 2013. 289 с. URL : toe.fea.kpi.ua/te_sait/posibniki/per_proc.pdf (дата звернення: 11.08.2018).
7. Bulanchuk G., Ostapenko A. Modeling of the viscous fluid flow around rotating circular cylinders with the lattice Boltzmann method at moderate Reynolds numbers. *Bulletin of V. Karazin Kharkiv National University, series Mathematical Modelling. Information Technology. Automated Control Systems*. 2017. Issue 36. P. 27–37. DOI: 10.26565/2304-6201-2019-41-09. URL: <https://periodicals.karazin.ua/mia/article/view/10086/9614> (Last accessed: 20.11.2018).
8. eLearning Industry [Internet source], available at the URL: <https://elearningindustry.com/> (last accessed on 21.10.2019).
9. HTML and CSS: Design and Build Websites by Jon Duckett
10. You Don't know JS: Scope & Closures by Kyle Simpson

ЛАГУТІН Артем Геннадійович – студент кафедри теоретичної і прикладної системотехніки Харківського національного університету імені В.Н. Каразіна.

Наукові інтереси:

– веб-розробка, програмна інженерія.

СТРИЛЕЦЬ Вікторія Євгенівна – доцент кафедри теоретичної та прикладної системотехніки Харківського національного університету імені В.Н. Каразіна, кандидат технічних наук, доцент.

Наукові інтереси:

– Математичні моделі і методи діагностування складних систем; системи штучного інтелекту; методи машинного навчання.

МАКСИМЕНКО В.Ю., ТОЛСТОЛУЗЬКА О.Г.

ОНЛАЙН-ТРЕНАЖЕР ДЛЯ ВИЗНАЧЕННЯ КОГНІТИВНИХ ЗДІБНОСТЕЙ ЛЮДИНИ

Вступ

Для успішного сприйняття значущої нам інформації з навколишнього світу нам необхідно вміння управляти своєю увагою. Увага – це процес відбору та поточного контролю актуальної на даному етапі інформації про конкретні об'єкти зовнішнього та внутрішнього середовища. У всіх сферах людської діяльності об'єкти, процеси та явища несуть інформацію, для сприйняття якої нам слід вміти перемикаати та розподіляти свою увагу. Тому його розвитком потрібно займатися всім людям, які бажають жити повноцінним життям та досягати високих результатів.

В умовах цифровізації суспільства та освіти, а також впровадження інформаційних технологій у всі сфери людської діяльності особливу актуальність набуває створення онлайн-тренажера для оцінки, тренування перемикаання та розподілу уваги.

Постановка задачі

Для створення такого онлайн-тренажера я розглянув властивості уваги та вибрав дві особливо значущих, якими є перемикаання та розподіл уваги. В основу лягли таблиці Горбова-Шульте, призначені для оцінки обраних властивостей. Онлайн-тренажер буде реалізований у вигляді веб-застосунку.

Суть онлайн-тренажера: Користувачеві надається таблиця (7×7) з 49 клітин. У клітини у випадковому порядку вписані цифри на червоному фоні (від 1 до 24) та цифри (від 1 до 25) на чорному фоні. Відповідно до інструкції випробовуваний повинен дуже швидко і без помилок знаходити цифри на чорному фоні (у порядку зростання від 1 до 25) та цифри на червоному фоні (у порядку зменшення від 24 до 1).

Для онлайн-тренажера складатиметься з трьох етапів:

- 1 етап: показати чорні числа у зростаючому порядку від 1 до 25
- 2 етап: показати червоні числа у спадному порядку від 1 до 24
- 3 етап: поперемінно знаходити та показувати числа у чорних клітинках від 1 до 25 у порядку зростання, а числа в червоних клітинках від 1 до 24 у спадному порядку.

Для опису роботи програми було складено блок схеми кожного етапу. Блок схема для першого етапу представлена на рисунку 1.

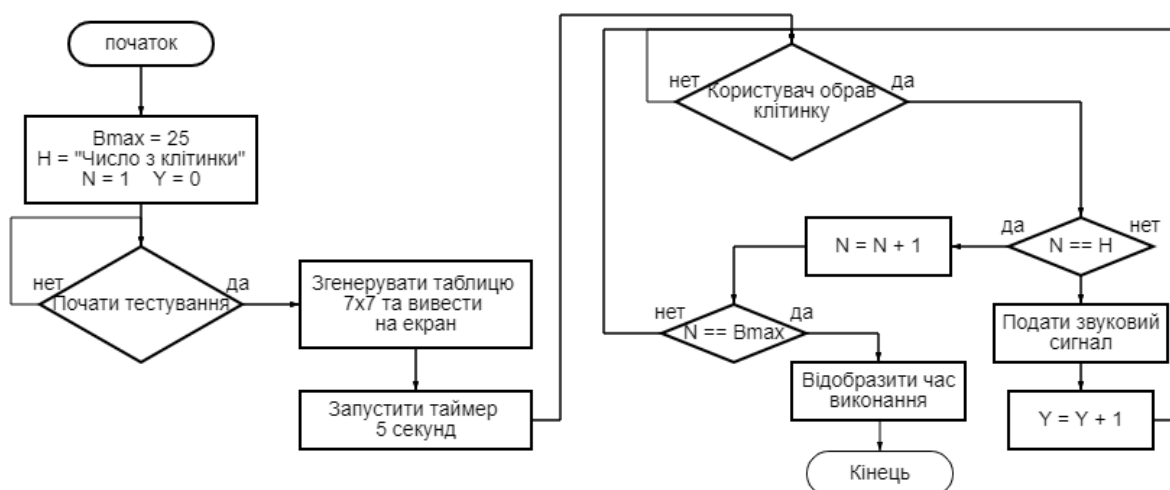


Рис. 1 Блок-схема онлайн тренажера.

Змінні, які будуть використані у програмі:

- V_{max} – 25 у комірці з чорним тлом;
- N – номер ходу (відповідає числу, яке має бути обрано наступним);
- H – число, що знаходиться у вибраному осередку;
- Y – у помилок.

Основна частина

1 етап.

Завдання: показати чорні числа у порядку зростання від 1 до 25

На екран виводиться таблиця Горбова-Шульте, де представлені числа від 1 до 25 у чорних клітинках та від 1 до 24 у червоних клітинках (при кожному новому запуску тесту цифри вибудовуються у випадковому порядку). Після натискання клавіші «Почати тестування» таблиця змінює свій масштаб (стає більше). На екрані з'являється зворотний звіт (5 секунд). За допомогою курсору миші випробовуваний вказує на числа в чорних клітинках від 1 до 25 в порядку зростання. Якщо випробовуваний, припускається помилки, видається звуковий сигнал (так відбувається доти, доки не буде вказано правильне число). Тест закінчується після того, як випробовуваний вказав усі числа від 1 до 25. Після закінчення тесту на екран виводиться час, витрачений на виконання завдання. Також на екран виводиться кнопка «Продовжити тест» кнопка призначена для переходу до 2 етапу.

Що необхідно фіксувати під час виконання 1 етапу:

1. Час, витрачений виконання першого етапу (t_1);
2. Кількість помилок під час виконання 1 етапу (Y_1).

2 етап.

Завдання: показати червоні числа у порядку спадання від 1 до 24

На екран виводиться таблиця Горбова-Шульте, де представлені числа від 1 до 25 у чорних клітинках та від 1 до 24 у червоних клітинках (при кожному новому запуску тесту цифри вибудовуються у випадковому порядку). Після натискання клавіші «Почати тестування» таблиця змінює свій масштаб (стає більше). На екрані з'являється зворотний звіт (5 секунд). За допомогою курсору миші випробовуваний вказує на числа в червоних клітинках від 1 до 24 у порядку зменшення. Якщо випробовуваний, припускається помилки, видається звуковий сигнал (так відбувається доти, доки не буде вказано правильне число). Тест закінчується після того, як випробовуваний вказав усі числа від 1 до 24. Після закінчення тесту на екран виводиться час, витрачений виконання завдання. Також на екран виводиться кнопка «Продовжити тест» призначена для переходу до 3 етапу.

Що необхідно фіксувати при виконанні 2 етапи:

1. Час, витрачений виконання другого етапу (t_2);
2. Кількість помилок при виконанні 2 етапу (Y_2).

3 етап.

Завдання: поперемінно знаходити та показувати числа у чорних клітинках від 1 до 25 у порядку зростання, а числа в червоних клітинках від 1 до 24 у спадному порядку.

На екран виводиться таблиця Горбова-Шульте, в якій представлені числа від 1 до 25 у чорних клітинках та від 1 до 24 у червоних клітинках (при кожному новому запуску тесту цифри вибудовуються у випадковому порядку). Після натискання клавіші «Почати тестування» таблиця змінює свій масштаб (стає більше). На екрані з'являється зворотний звіт (5 секунд). За допомогою курсору миші випробовуваний вказує по чергову на числа в чорних клітинках від 1 до 25 у порядку зростання, а на числа в червоних клітинках від 1 до 24 у порядку зменшення. Якщо випробовуваний, припускається помилки, видається звуковий сигнал (так відбувається до тих пір, поки не буде вказано правильне число). Тест закінчується після того, як випробовуваний вказав всі числа у правильному порядку. Після закінчення тесту на екран виводиться час,

витрачений виконання завдання. Також на екран виводиться кнопка «Закінчити тестування» призначений для перегляду результатів тестування.

Що необхідно фіксувати при виконанні 3 етапу:

1. Час, витрачений виконання другого етапу (t_3);
2. Кількість помилок під час виконання 3 етапу (Y_3).

Висновки

Таким чином, у роботі представлена методика діагностики та розвитку перемикання та розподілу уваги на основі онлайн-тренажера. Спроектований онлайн-тренажер планується виконати у вигляді веб-застосунку. Користувачі зможуть визначити свої обсяг, розподіленість і перемикання уваги.

ЛІТЕРАТУРА

1. Безпалько О. В. Соціальна педагогіка в схемах і таблицях: Навчальний посібник. – К.: Центр навч. літератури, 2003. – 137 с.
2. Бондар В. І. Дидактика: Підручник. – К.: Либідь, 2005. – 264 с.
3. Вітвицька С. С. Основи педагогіки вищої школи: Методичний посібник. – К.: Центр навч. літератури, 2003. – 316 с.
4. Власова О. І. Педагогічна психологія: Навчальний посібник. – К.: Либідь, 2005. – 400 с.
5. Вознюк Н. М. Етико-педагогічні основи формування особистості: Навчальний посібник. – К.: Центр навч. літератури, 2005. – 196 с.
6. Волкова Н. П. Педагогіка: Навчальний посібник. – К.: Академія, 2003. – 576 с.
7. Солсо Р.: Когнітивна психологія. 2005. — 28с.

МАКСИМЕНКО Владислав Юрійович – студент; Харківський національний університет імені В.Н.Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: ха11867733@student.karazin.ua; ORCID: 0000-0002-1052-2027.

Наукові інтереси:

– комп'ютерні моделі веб-сервісів.

ТОЛСТОЛУЗЬКА Олена Геннадіївна – професор кафедри теоретичної та прикладної системотехніки, д. т. н., с. н. с.; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: elena.tolstoluzka@karazin.ua; ORCID: 0000-0003-1241-7906.

Наукові інтереси:

– технології автоматичного проектування паралельних програм.

МОДЕЛІ І МЕТОДИ ОБРОБКИ ТЕКСТОВИХ ДАНИХ ЗАСОБАМИ МАШИННОГО НАВЧАННЯ

Тема роботи. Моделі і методи обробки текстових даних засобами машинного навчання.

Мета роботи. Ознайомлення та вивчення найбільш розповсюджених моделей, методів та засобів для обробки текстових даних у машинному навчанні.

Методи дослідження. Аналіз, пояснення, класифікація.

Результати. Результатом виконання роботи являється огляд і загальний опис найбільш розповсюджених моделей і методів обробки текстових даних у машинному навчанні та приклади їх застосування.

Висновки. У роботі були розглянуті моделі та методи для обробки текстових даних засобами машинного навчання. Дана робота ознайомлює читача з основами роботи та реалізації цих методів, їх перевагами та висвітлює області їх практичного використання. У роботі описано взаємодію деяких методів одне з одним для покращення результатів їх роботи. Також дана робота може бути початковим керівництвом до вибору методів або моделей для людей, які тільки починають роботу у сфері машинного навчання та допомогти у виборі більш ефективних методів для виконання практичних задач.

Ключові слова: NLP, машинне навчання, обробка тексту, розпізнавання, нейронні мережі.

Вступ

Natural Language Processing – обробка природної мови – підрозділ інформатики та штучного інтелекту, присвячений тому, як комп'ютери аналізують природні (людські) мови. NLP дозволяє застосовувати алгоритми машинного навчання для тексту та мовлення.

NLP використовується для розуміння структури та значення людської мови шляхом аналізу різних аспектів, таких як синтаксис, семантика, прагматика та морфологія. Потім інформатика перетворює ці лінгвістичні знання в засновані на правилах алгоритми машинного навчання, які можуть вирішувати конкретні проблеми та виконувати бажані завдання.

Візьмемо, наприклад, Gmail. Електронні листи автоматично класифікуються як рекламні, соціальні, основні або спам завдяки завданню NLP під назвою вилучення ключових слів. «Зчитуючи» слова в темах і пов'язуючи їх із задалегідь визначеними тегамі, машини автоматично дізнаються, яку категорію призначати електронним листам.

Задачі обробки природної мови у машинному навчанні

Людська мова наповнена двозначностями, що неймовірно ускладнює написання програмного забезпечення, яке точно визначає передбачуване значення текстових або голосових даних. Омоніми, омофони, сарказм, ідіоми, метафори, винятки з граматики та вживання, варіації в структурі речень — це лише деякі з порушень людської мови, на вивчення яких люди потребують роки, але програмісти повинні навчити програми, що керуються природною мовою, розпізнавати та з самого початку точно зрозуміти, чи будуть ці програми корисними.

Завданням NLP являється обробка текстових та голосових даних таким чином, щоб комп'ютер міг зрозуміти інформацію яку він отримує. Нижче перераховані деякі з найпоширеніших видів задач:

- Розпізнавання мовлення, яке також називають text-to-speech, — це завдання перетворення голосових даних у текстові. Розпізнавання мовлення потрібне для будь-якої програми, яка виконує голосові команди або відповідає на запитання. Що робить розпізнавання мовлення особливо складним, так це те, як люди розмовляють — швидко, невиразно перебираючи слова, з різним наголосом та інтонацією, з різними акцентами та часто з використанням неправильної граматики.
- Позначення частини мови, яке також називають граматичним тегом, — це процес визначення частини мови певного слова чи фрагмента тексту на основі його

використання та контексту. Частина мови визначає англійське слово «make» як дієслово у вислові «I can make a paper plane» і як іменник у вислові «What make of car do you own?»

- Усунення неоднозначності слова – це вибір значення слова з кількома значеннями за допомогою процесу семантичного аналізу, який визначає слово, яке має найбільший сенс у даному контексті. Наприклад, усунення неоднозначності слів допомагає розрізнити значення слів омографів «Ми знайшли човен на березі річки» від «Птах сидів на високій березі».
- Розпізнавання іменованих сутностей, визначає слова або фрази як корисні сутності. NEM визначає «Кентуккі» як місцезнаходження або «Фред» як ім'я людини.
- Аналіз настроїв намагається витягнути з тексту суб'єктивні якості — ставлення, емоції, сарказм, збентеження, підозру.
- Генерація природної мови іноді описується як протилежність розпізнаванню мовлення або text-to-speech; це завдання перекласти структуровану інформацію людською мовою.

Моделі та методи обробки природної мови у машинному навчанні

Токенізація є однією з найбільш основних і простих технік НЛП при обробці природної мови. Токенізація є важливим кроком під час попередньої обробки тексту для будь-якої програми NLP. Тривалий текстовий рядок береться та розбивається на менші одиниці, звані токенами, які складають слова, символи, числа тощо. Ці токени є будівельними блоками та допомагають зрозуміти контекст під час розробки моделі НЛП. Більшість токенизаторів використовують «порожній простір» як роздільник для формування tokenів. Залежно від мови та мети моделювання в НЛП використовуються різні техніки токенизації

Стемінг або лемматизація. Наприклад, коли ми виконуємо пошук продуктів в онлайн магазині, скажімо, ми хочемо бачити продукти не лише за тим словом, яке ми ввели в рядок пошуку, але й за іншими можливими формами слова, яке ми ввели. Дуже ймовірно, що ми захочемо побачити результати пошуку, які містять в своєму описі слова «комп'ютерний/комп'ютерна», якщо ми введемо «комп'ютер» у вікні пошуку. Техніка стеммінгу або лемматизації НЛП спрямована на створення кореневих слів із цих варіацій слова. Стеммінг — це досить грубий евристичний процес, який намагається досягти вищезазначеної мети шляхом відсікання кінців слів, що може призвести або не призвести до значущого слова в кінці. Лематизація, з іншого боку, є більш складною технікою, яка спрямована на те, щоб робити речі належним чином за допомогою словникового запасу та морфологічного аналізу слів. Видаляючи флексивні закінчення, повертається основа або словникова форма слова, яка називається лемою.

Видалення стоп-слів – етап попередньої обробки, який настає відразу після стеммінгу або лематизації, полягає в видаленні «не потрібних» слів. У будь-якій мові багато слів є лише наповнювачами і не мають жодного значення. Здебільшого це слова, які використовуються для зв'язку речень (сполучники – «тому що», «і», «оскільки») або використовуються для вказівки зв'язку слова з іншими словами (прийменники – «під», «над», «в», «при») . Ці слова складають більшу частину людської мови і не дуже корисні при розробці моделі NLP. Однак видалення стоп-слів не є певною технікою NLP, яку слід застосовувати для кожної моделі, оскільки це залежить від завдання. Наприклад, під час виконання класифікації тексту, якщо текст потрібно класифікувати за різними категоріями (класифікація за жанрами, фільтрація спаму, автоматичне створення тегів), тоді корисно видалити стоп-слова з тексту, оскільки модель може зосередитися на словах, які визначають значення текст у наборі даних. Для таких завдань, як узагальнення тексту та машинний переклад, видалення стоп-слів може не знадобитися.

Вилучення ключових слів – коли ви читаєте фрагмент тексту на телефоні, газеті чи книзі, ви мимовільно переглядаєте його – ви здебільшого ігноруєте слова-заповнювачі та знаходите в тексті важливі слова, а все інше відповідає контексту. Функція вилучення ключових слів виконує те саме, що й пошук важливих ключових слів у документі. Вилучення ключових

слів — це техніка НЛП аналізу тексту для отримання значущої інформації щодо теми за короткий проміжок часу. Замість того, щоб переглядати документ, можна використати техніку вилучення ключових слів, щоб стиснути текст і витягнути відповідні ключові слова. Техніка вилучення ключових слів дуже корисна в додатках NLP, де компанія хоче визначити проблеми клієнтів на основі оглядів або якщо ви хочете визначити цікаві теми з останніх новин.

Word2Vec — це модель нейронної мережі, яка вивчає асоціації слів із великого об'єму тексту. Word2vec можна навчити двома способами: за допомогою моделі загального пакета слів (CBOW) або моделі Skip-Gram. У моделі загального пакету слів контекст кожного речення береться як вхід, а слово, що відповідає контексту, має бути передбачено як вихід. Розглянемо приклад речення: «День яскравий і сонячний». У наведеному вище реченні слово, яке ми намагаємось передбачити, є сонячним, використовуючи вхідні дані як середнє значення векторів, закодованих одним кодом слів: «День яскравий». Цей вхідний сигнал після проходження через нейронну мережу порівнюється з одним закодованим вектором цільового слова «сонячний». Розраховується збиток, і таким чином дізнається контекст слова «сонячний». Модель Skip Gram працює якраз протилежно описаному вище підходу: ми надсилаємо вхідні дані як одноразово закодований вектор нашого цільового слова «сонячний», і він намагається вивести контекст цільового слова. Для кожного вектора контексту ми отримуємо розподіл ймовірностей із V ймовірностей, де V — розмір слова, а також розмір вектора, закодованого одним оперативним кодуванням.

Аналіз настроїв — також відомий як емоційний штучний інтелект або дослідження думок, є одним із найважливіших методів NLP для класифікації тексту. Мета полягає в тому, щоб класифікувати текст такі види тексту як пости у соціальних мережах, статтю у ЗМІ, огляд фільму чи будь-який текст в Інтернеті, в одну з трьох категорій: позитивний/негативний/нейтральний. Аналіз настроїв найчастіше використовується для виявлення ворожих висловлювань у соціальних мережах і виявлення проблемних клієнтів і негативних відгуків.

Моделювання тем — це статистичний метод NLP, який аналізує набір текстових документів, щоб знайти приховані в них теми. Це алгоритм машинного навчання без вчителя, тобто йому не потрібно позначати ці документи. Ця техніка дозволяє організовувати та узагальнювати електронні архіви в масштабі, який був би неможливий за участі людини. Основна ідея полягає в тому, що кожен документ має кілька тем, і кожна тема розподілена за фіксованим словником слів. Скажімо, у нас є колекція документів з науковими публікаціями. Документ, який ми зараз розглядаємо, пов'язаний з наукою, точніше з предметом біології. У цьому документі є кілька тем. Ці теми широко пов'язані з генами, біологією, неврологією та інформатикою. Будь-яка з цих тем може бути найважливішою в будь-якому з документів нашого корпусу, але наш метод допомагає виявити, що для поточного документа найбільш значущими є теми, пов'язані з генами та біологією і на основі цих висновків ми можемо віднести документ до певної тематики або області наукових досліджень.

Суммаризація тексту. Ця техніка NLP використовується для стислого та короткого резюмування тексту плавним і зв'язним способом. Резюмування корисно для отримання корисної інформації з документів без необхідності читати слово в слово. Цей процес займає дуже багато часу, якщо виконується людиною, автоматичне підсумовування тексту радикально скорочує час. Існує два типи прийомів суммаризації тексту:

- Резюме на основі вилучення: у цій техніці деякі ключові фрази та слова в документі витягуються для створення резюме. Жодних змін до оригінального тексту не вноситься.
- Резюме на основі абстракції: у цій техніці з оригінального документа створюються нові фрази та речення, які містять найбільш корисну інформацію. Мова та структура речень

резюме не збігаються з оригінальним документом, оскільки ця техніка передбачає перефразування.

Висновки

У роботі були розглянуті моделі та методи обробки текстових даних засобами машинного навчання. Дана галузь машинного навчання покриває великий спектр задач у різних галузях. Аналіз текстових даних може використовуватись як для наукових досліджень, особливо це релевантно у соціологічних дослідженнях для аналізу настроїв у суспільстві або політології, так і для комерційних потреб таких як виявлення проблемних клієнтів, аналіз відгуків для виявлення недоліків продукції, машинного перекладу, тощо. У роботі були розглянуті найбільш розповсюджені методи та моделі обробки текстових даних, поверхневий аналіз їх роботи та сфери їх використання. Дана робота висвітлює загальні поняття обробки текстових даних тож може бути у певному сенсі керівництвом для початківців у галузі машинного навчання та використовуватись як дорожня мапа для подальшого поглибленого вивчення та роботи у даній сфері.

ЛІТЕРАТУРА

1. Уосерман Ф. Нейрокомп'ютерна техніка: Теорія та Практика; Neural Computing. Theory and Practice. — Мир, 1992 – 240 с.
2. Michael Taylor Neural Networks: A Visual Introduction for Beginners – Independence publisher, 2017 – 248 с.
3. Segaran T. Programming Collective Intelligence – O'Reilly Media, 2008 – 362p.
4. Sebastian Raschka Python Machine Learning: Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow – Packt Publishing, 2019 – 770 p.
5. Occhipinti A., Rogers L., Angione C. – A pipeline and comparative study of 12 machine learning models for text classification – School of Computing, Engineering and Digital Technologies Teesside University, 2022.

МАЛИГА Ігор Євгенійович – аспірант; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: igormalyga@gmail.com; oleksiy33@gmail.com; ORCID: 0000-0002-5708-7739.

Наукові інтереси:

– обробка натуральної мови у машинному навчанні.

АНАЛІЗ МЕТОДІВ ПОДІЛУ КАНАЛІВ У ВИРІШЕННІ ПИТАНЬ МНОЖИННОГО ДОСТУПУ В КОМП'ЮТЕРНІЙ МЕРЕЖІ

The work is devoted to the analysis of methods of channel division in solving multiple access issues in a computer network. Its purpose is to find methodologies and tools for the most effective solution of problems related to multiple access of subscribers to the communication channel. The study was conducted by analyzing information sources of existing methods of channel division such as FDMA, TDMA and CDMA. The accompanying tools of the CDMA methodology are also considered. The result is the determination of the optimal separation method for solving the problem of multiple access of subscribers in networks. In the conclusion to the work, it was determined that, given the disadvantages and advantages, the use of the CDMA methodology and its set of tools (the DSSS method, Walsh orthogonal codes and Barker sequences) is optimal for achieving the given task.

Ключові слова: кодовий поділ каналів, ортогональні функції, псевдовипадкові послідовності, автокореляційна функція кодів Баркера.

Вступ

В ході еволюції автоматизованих систем управління технологічними процесами (АСУ ТП), інженери стикалися з великою кількістю задач і проблем передачі інформації, що заважають коректній та комфортній роботі з цими системами. Протягом років проблема множинного доступу абонентів залишається однією з головних в цьому переліку. Виявлена проблематика особливо виразна, коли станції починають передавати дані в інформаційний канал одночасно. Тоді їх сигнали з великою ймовірністю будуть накладатись один на одного. Через це базова приймальна станція не зможе відрізнити ці сигнали.

В різний час були запропоновані різні варіанти вирішення цієї задачі. Серед них виділяють FDMA (Frequency Division Multiple Access) – розділення за частотою, TDMA (Time Division Multiple Access) – розподіл за часовими інтервалами та CDMA (Code Division Multiple Access) – кодовий розподіл інформаційних каналів^[1].

Методи поділу каналів.

Частотний поділ каналів. Принцип FDMA полягає в поділі загального діапазону частот на окремі частотні канали. В одному частотному діапазоні знаходиться тільки один абонент [1]. У діапазоні частот кожного каналу виділяється відносно вузька (20-25 КГц) смуга. Між цими смугами є додаткові захисні частотні інтервали, які мінімізують взаємовплив каналів.

Перевагами цього методу є простота реалізації та висока якість передачі даних. Проте через вузькі смуги та чутливість інформаційного каналу, виявляються і недоліки методу. Мала стійкість до перешкод та шуми можуть з легкістю викривити дані, що передаються. Одноосібне використання кожного каналу абонентами також спричиняє досить вагому проблему. Поки початковий запит не закінчений, канал закритий для інших сеансів зв'язку. Це призводить до нерационального використання ділянок діапазону частот, особливо на одній території.

Часовий поділ каналів. Робота методу TDMA полягає в розподіленні різних часових інтервалів передачі між абонентами, які знаходяться в одному частотному діапазоні^[1]. Тобто TDMA покладається на те, що інформація поділяється на декілька мілісекундних пакетів. Спочатку виділяється один частотний канал на короткий відрізок часу, а потім переходить до іншого. Цифрові відліки від одного передавача займають різні часові інтервали в декількох діапазонах одночасно. TDMA використовує технологію доступу, в якій три користувача. Кожен їх використовує несучу частоту 30 КГц.

Серед переваг цієї методології перш за все можна виділити ефективну та адаптивну передачу даних, за рахунок такого розподілу інформації. TDMA пропонує можливість передачі зі швидкістю від 64 кбіт/с до 120 Мбіт/с та забезпечує суттєву економію в обладнанні, просторі та обслуговуванні базової станції, що є важливим фактором збільшення розмірів осередків. Та, на відміну від методів CDMA, відокремлює користувачів у часі, що гарантує, безперешкодність інших одночасних передач.

Проте TDMA методологія має і недоліки. Серед них можливість обриву сигналу в процесі переходу абоненту з одного осередку до іншого. Це може викликатись тим, що всі часові інтервали каналу, до якого переходить абонент, вже будуть зайняті. Аналогічно, якщо всі часові інтервали в осередку, в якому знаходиться користувач, вже зайняті, користувач не отримає сигнал відповіді станції. TDMA також піддається багатопроменеві спотворення. Сигнал, що надходить від вежі, може виходити з будь-якого з кількох напрямків. Він може відскочити від кількох різних будівель до прибуття, що може спричинити перешкоди. Один із способів обійти це втручання – встановити обмеження за часом на систему. Система буде призначена для прийому та обробки сигналу протягом певного терміну. Після закінчення терміну дії система ігнорує сигнали. Чутливість системи залежить від того, як вона обробляє багатопроменеві частоти.

Кодовий поділ каналів. CDMA є формою комунікації розподіленого спектра, в якій вузькосмуговий сигнал розподіляється по більш широкому діапазону частот [1, 2]. Це забезпечує стійкість до перешкод та спільне використання загального діапазону декількома сигналами різних абонентів. Ключовою ідеєю CDMA є виділення корисного сигналу при ігноруванні усіх інших шумових. У CDMA кожен бітовий інтервал розбивається на m періодів, які називаються елементарними сигналами або чіпами. Кожній станції відповідає унікальний m -бітний код, який має назву елементарної послідовності. Наприклад, біполярна система -1 і $+1$.

Щоб передати один біт, станція посилає свою елементарну послідовність. Щоб передати біт зі значенням 0, потрібно відправити замість елементарної послідовності її доповнення (всі одиниці послідовності змінюються на нулі, а всі нулі – на одиниці). Ніякі інші комбінації передавати не дозволяється. Збільшити кількість інформації, яку необхідно передавати, можна тільки за рахунок збільшення в m раз пропускної здатності. Таким чином, для CDMA-станцій потрібна в m разів більша пропускна здатність, ніж для інших.

Кожна станція має свою унікальну елементарну послідовність. Всі елементарні послідовності попарно ортогональні. Нормування скалярного добутку двох різних елементарних послідовностей дорівнює 0:

$$S \cdot T \equiv \frac{1}{m} \sum_{i=1}^m S_i T_i = 0, \quad (1)$$

де S – вектор довжини m ; T – доповнення S , m – кількість чіпів.

На основі (1), можна зробити висновок, що нормований скалярний добуток елементарної послідовності самої на себе дорівнює 1:

$$S \cdot S = \frac{1}{m} \sum_{i=1}^m S_i S_i = \frac{1}{m} \sum_{i=1}^m S_i^2 = \frac{1}{m} \sum_{i=1}^m (\pm 1)^2 = 1 \quad (2)$$

Кожен з m доданків суми (2) дорівнює 1, тому вся сума дорівнює m . Зверніть також увагу, що $S \cdot S = -1$ [5].

Протягом кожного бітового інтервалу станція може передавати 1, посилаючи свою елементарну послідовність, або передавати 0, посилаючи доповнення до послідовності, або може мовчати і нічого не передавати.

$$\begin{aligned} A &= (-1 -1 -1 +1 +1 -1 +1 +1) \\ B &= (-1 -1 +1 -1 +1 +1 +1 -1) \\ C &= (-1 +1 -1 +1 +1 +1 -1 -1) \\ D &= (-1 +1 -1 -1 -1 -1 +1 -1) \end{aligned}$$

Рис. 1 - Приклад передачі послідовностей сигналів "станцій" A, B, C та D

Коли дві або більше станції намагаються здійснити одночасно передачу, їх біполярні сигнали лінійно складаються. Наприклад, якщо передачі одного елементарного сигналу три станції передають $+1$, а одна передала -1 , то результаті вийде $+2$.

Щоб відновити початковий бітовий потік кожної станції, приймач повинен заздалегідь знати елементарні послідовності всіх передавачів, з якими він працює. Відновлення здійснюється шляхом обчислення нормованого скалярного добутку прийнятої послідовності та елементарної послідовності тієї станції, початковий сигнал якої відновлюється. Якщо прийнята елементарна послідовність S і приймач намагається зрозуміти, що передала станція з елементарною послідовністю C , проводиться обчислення нормованого скалярного твору $S \cdot C$.

Методи кодового поділу каналів. Ортогональні коди і послідовності Уолша. Ортогональні кодові послідовності поділяються на ортогональні і квазіортогональні [4]. Основна ознака послідовності – коефіцієнт взаємної кореляції. Мінімальне значення взаємної

кореляційної функції забезпечують коди, в яких коефіцієнти кореляції для будь-яких пар послідовностей є негативними (трансортогональні коди). Коефіцієнт взаємної кореляції ортогональних послідовностей завжди дорівнює нулю. Є декілька способів генерувати ортогональні послідовності найбільш розповсюджений – за допомогою послідовностей Уолша довжиною 2^n .

Послідовності Уолша – коди, які мають властивості ортогональності і мультиплікативності, можуть бути згенеровані декількома способами, одним із них є матриці Адамарта^{[2][4]}.

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}, H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Рис. 2 - Матриці Адамарта

Початковою є послідовність виду $H_1 = \{0\}$. Підставляючи її в матрицю H_{2n} отримуємо нову матрицю більшого розміру. Багаторазове повторення процедури дозволяє сформувати матрицю будь-якого розміру, для якої характерна ортогональність усіх стовпців і рядків. Такий спосіб формування послідовностей реалізований в стандарті IS-95, де довжина послідовностей Уолша обрана 64. Різниця між матрицею Адамарта і послідовностями Уолша складається в тому, що в послідовностях Уолша використовуються уніполярні сигнали виду $\{0,1\}$ ^[4].

Ортогональним кодам притаманні два суттєвих недоліки. Першим з них є обмеження максимального числа послідовностей їх довжиною або обмеження адресного простору. Для розширення ансамблю сигналів разом з ортогональними послідовностями використовують квазіортогональні послідовності. В проекті стандарту CDMA2000 використовують запропонований метод генерації квазіортогональних кодів шляхом множення послідовностей Уолша на спеціальну маскуючу функцію^{[2][4]}. Цей метод дозволяє за допомогою однієї такої функції отримати набір квазіортогональних послідовностей. Ще одним недоліком ортогональних кодів є те, що ортогональність здобувається лише в фіксованій точці у часі. Тому ці коди використовуються лише в синхронізованих системах.

Коди Баркера. Коди Баркера – це спеціальні коди, кожен з елементів яких має значення або 1, або -1^[3]. Вони підбрані так, що автокореляційна функція має високий максимум і мале значення бічних пелюсток. Ці послідовності слугують для того щоб серед перешкод точно виділити початок передачі і вловити такт.

Існує дев'ять послідовностей Баркера: по дві довжиною в 2 і 4 символи, і по одній довжиною в 3, 5, 7, 11 і 13 символів відповідно. Послідовності більше 13-ти символів не існує.

Послідовності Баркера використовують для синхронізації і фазування в системах передачі, через їх хорошу АКФ (автокореляційну функцію).

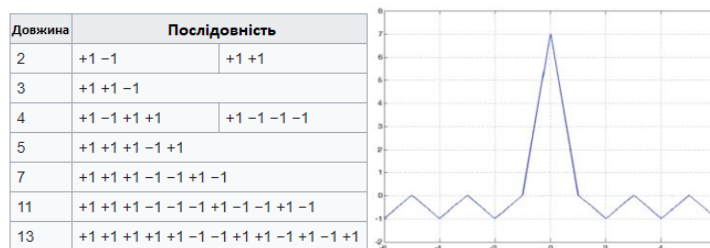


Рис. 3 - Послідовності Баркера та приклад АКФ кода Баркера

Порівняльний аналіз методів розширення спектру. Технологія розширення спектру дозволяє збільшити завадостійкість коду для сигналів малої потужності за рахунок збільшення спектру сигналу, що дуже важливо в мобільних додатках^{[2][5]}.

В CDMA системах можна використовувати два методи розширення спектру: FHSS (псевдовипадкова перебудова несної частоти) і DSSS (пряма послідовність для розширення) [2, 5].

У методі FHSS^{[2][5]} протягом визначеного інтервалу часу передача ведеться на змінній несній частоті. Несна частота змінюється узгоджено з номером, який генерує алгоритм псевдовипадкових чисел і залежить від початкового числа. Однаковий алгоритм працює на приймачі і передачі, що дозволяє їм в однаковий час перемикається на однакову несну частоту. Метод намагається зайняти весь діапазон частот і підібрати для кожного каналу такий алгоритм генерування псевдовипадкових послідовностей, щоб при кожному перемикаванні вони переходили на різні частоти. В одному діапазоні може працювати не більше число каналів, ніж мається підканалів. На кожному моменті переходу, для постійної синхронізації передаються синхробіти. Таким чином загальна корисна швидкість зменшується. Якщо частота зміни підканалів нижче, ніж швидкість передачі даних в каналі, то такий режим називають повільним розширенням спектра інакше ми маємо справу з швидким розширенням спектра.

У методі DSSS^{[2][5]} частотний діапазон розширюється не за рахунок постійних перемикань з частоти на частоту, а за рахунок того, що кожен біт інформації замінюється N бітами, тому тактова швидкість передачі сигналів та спектр сигналу збільшуються в N раз.

Мета кодування методом DSSS та ж, що і методом FHSS, - підвищення завадостійкості. Вузкосмугова перешкода буде спотворювати тільки певні частоти спектра сигналу, так що приймач з великим ступенем ймовірності зможе правильно розпізнати передану інформацію.

Метод швидкого розширення спектра більш стійкий до перешкод, оскільки вузкосмугова перешкода не призводить до втрати біта, так як його значення передається кілька разів в різних частотних підканалах. У цьому режимі не проявляється ефект міжсимвольної інтерференції, адже до приходу сигналу, затриманого вздовж одного з шляхів, система встигає перейти на іншу частоту. Метод повільного розширення спектра такою властивістю не володіє, але зате він простіше в реалізації і має менші накладні витрати.

Висновки. В роботі було розглянуто декілька існуючих методів поділу каналів в процесі експлуатації бездротових комп'ютерних мереж з множинним доступом абонентів. Під час аналізу було розглянуто основні принципи їх роботи, ряд переваг та недоліків.

Попередній аналіз методів множинного доступу абонентів показав, що використання методів FDMA і TDMA не є доцільним. FDMA ділить загальну частоту каналу між усіма абонентами. Це призводить до нераціонального використання ресурсу частоти і побудови бар'єрів між каналами. TDMA може втратити дані при зриві синхронізації БС. Цей недолік не є критичним через можливість динамічного перерозподілу кадрів, проте вразливість до багатопромених спотворень збільшує ймовірність отримати спотворені дані. Оптимальним вибором є CDMA завдяки можливості раціонального використання ресурсу частоти каналу і захисту від перешкод. Для забезпечення цих характеристик було вирішено використовувати метод DSSS, ортогональні коди Уолша і послідовності Баркера.

ЛІТЕРАТУРА

1. Ипатов В.И. Широкополосные системы и кодовое разделение сигналов. Принципы и приложения. «Техносфера»: 2007, 488 с.
2. Hsiao-Hwa Chen. The Next generation CDMA technologies. *John Wiley & Sons Ltd.* 2007. 478p.
3. Barker, R. H. Group Synchronizing of Binary Digital Sequences. *Communication Theory*. London: Butterworth. 1953. 273 – 287.
4. Walsh, J. A closed set of normal orthogonal functions. *American Journal of Mathematics*. Vol. 45, No. 1. 1923. 5-24
5. Таненбаум Э., Уэзеролл Д. Компьютерные сети 5-е изд. СПб.: Питер, 2012. 960 с.: ил. ISBN 978-5-459-00342-0

МАНИЦЬКИЙ Сергій Сергійович – студент магістратури 1-го курсу кафедри теоретичної та прикладної системотехніки, Харківський національний університет імені В. Н. Каразіна, майдан Свободи 6, Україна, 61022; e-mail: serhii.manytskyi@student.karazin.ua; ORCID: 0000-0002-4348-3677.

Наукові інтереси: – комп'ютерні мережі; множинний доступ.

МЕДВЕДЕНКО В.О.

АНАЛІЗ ЗАСТОСУВАННЯ КОНЦЕПЦІЇ WEB 3.0 ТА ЇЇ ВПЛИВУ НА МЕРЕЖУ ІНТЕРНЕТ

Мета: проаналізувати застосування концепції Web 3.0 та її вплив на мережу Інтернет.

Методи дослідження: теоретичний аналіз.

Результати: у ході аналізу застосування концепції Web 3.0 та її впливу на інтернет з'ясував що Web 3.0 досить нова концепція яка має розвиток і перспективи у найближчому майбутньому.

Ключові слова: Web, Web 3.0, Semantic Web, Семантичне павутиння, блокчейн, децентралізовані застосунки

Вступ

За останні двадцять років інтернет дуже змінився. Ми перейшли від Internet Relay Chat (IRC) до сучасних платформ соціальних мереж. Від базових цифрових платежів до найскладніших банківських онлайн послуг. Ми застали час абсолютно нових інтернет технологій, такі як криптовалюта та блокчейн. Інтернет став життєво важливою частиною людської взаємодії та зв'язку і продовжує розвиватися. До цих пір ми бачили Web 1.0 та 2.0, але чого саме нам слід очікувати від Web 3.0?

Історія Web

Поняття Web 1.0 можна розглядати як "Інтернет тільки для читання". Іншими словами, раніше Інтернет дозволяв нам шукати інформацію і читати її. Було дуже мало шляхів для взаємодії з користувачами або для участі в складанні змісту. Проте, на першому етапі розвитку мережі Інтернет це було саме те, чого хотіло більшість власників веб-сайтів: мета їх сайту полягала в тому, щоб позначити свою присутність в Інтернеті і зробити інформацію про себе доступною для всіх в будь-який час.

Web 2.0, – друге покоління мережевих сервісів Інтернету. На відміну від першого покоління сервісів (the mostly read-only Web) Веб 2.0 (the wildly read-write Web) дозволяє користувачам спільно діяти – обмінюватися інформацією, зберігати посилання та мультимедійні документи, створювати та редагувати публікації, тобто відбувається налагодження соціальної взаємодії. Тому технології Веб 2.0. ще називають соціальними сервісами Інтернету.

Web 3.0

Поняття «Web 3.0» одне із тих, які кожен трактує собі по-різному. Так, згідно з визначенням Джейсона Калаканіса, «Web 3.0 – високоякісний контент та сервіси, які створюються талановитими професіоналами на технологічній платформі Web 2.0» Інше визначення ідентифікує Web 3.0 як «Семантичне павутиння» (Semantic Web).

Головна думка Semantic Web базується на впровадженні метамови, яка описує зміст сайтів для організації автоматичного обміну між серверами. Інакше кажучи, ці прикріплені метадані повинні бути зрозумілі системою, піддані машинній обробці, логічно співвіднесені один з одним, проаналізовані і включені в результати запиту.

Більшість проектів які позиціонують себе як проект з використанням концепцій Web 3.0 працюють на блокчейн технологіях, тому можна виділити такі переваги:

- **Спільне прийняття рішень**

Рішення про спірний контент, наприклад, про такий, що закликає до насильства, прийматиметься спільно учасниками блокчейну.

- **Безпека інформації**

Застосунки Apps, якими ми користуємося зараз, перетворюються на DApps – децентралізовані застосунки, які складніше зламати.

- **Криптогаманець для входу**

Ніяких логінів і паролів. Усі операції в мережі можна проводити під ніком, без реальних імені та прізвища. Щоб зайти до мережі, достатньо адреси вашого криптовалютного гаманця.

- **Штучний інтелект**

Ми вже знайомі з ботами обслуговування клієнтів, але це лише початок. Штучний інтелект може як представляти нам дані, так і сортувати їх, що робить його універсальним інструментом для Web 3.0. Найкраще те, що штучний інтелект вчитиметься і вдосконалюватиметься, скорочуючи роботу, необхідну для розвитку людини в майбутньому.

Табл.1 Порівняння WEB 1.0, WEB 2.0 та WEB 3.0

	WEB 1.0	WEB 2.0	WEB 3.0
Зміст	Пасивна взаємодія для користувача	Платформи спільноти та контент користувачів	Право власності для творців контенту
Технології	HTML	Динамічний HTML, Javascript	Блокчейн, штучний інтелект, машинне навчання
Віртуальні середовища	Немає	Деяке базове використання 3D	3D, VR, AR
Реклама	Нав'язлива (банери і т.д.)	Інтерактивна	Таргетинг на основі поведінки користувача
Зберігання даних	Зберігання на серверах окремих вебсайтів	Належить великим технологічним гігантам	Розподілено серед користувачів
Аудиторія	Індивідуальні користувачі	Конкретні спільноти користувачів	Взаємопов'язані користувачі на різних платформах та пристроях

Аналіз застосування

Концепція WEB 3.0 вже має широке застосування в багатьох сферах Інтернет послуг, і вирішує більшість типових проблем сучасних веб-застосунків. Провівши теоретичний аналіз, я визначив, що наприклад більшість месенджерів повністю контролюють ваш діалог та можуть при першому ж запиті, або зламі отримати доступ до всіх ваших даних та діалогів. Тому за допомогою децентралізації, яка є однією з особливостей Web 3.0, у купі з шифруванням, можна вирішити цю та інші проблеми пов'язані з конфіденційністю користувачів.

Проблеми пов'язані зі зберіганням даних, наприклад витік даних у LinkedIn виявив імена та паролі 117 мільйонів користувачів LinkedIn. Uber також мав подібну ситуацію, коли було викрадено дані 57 мільйонів клієнтів їх бази даних. Існуючі системи зберігання даних також є централізованими, це означає, що даними можна маніпулювати або використовувати для виконання інших видів діяльності, таких як продаж даних третім сторонам в рекламних цілях. Рішенням даної проблеми є використання децентралізованих систем зберігання даних. Наприклад, веб-застосунок Sia, який працює наступним чином: розбиває файл на тридцять сегментів, потім розподіляє їх відповідним чином, а також шифрує файл під час передачі.

Дуже важливо сказати про браузери Web 3.0. Поточні браузери не повністю захищені, коли йдеться про захист користувачів. У браузерах може статися витік інформації, якщо вони не добре розроблені, відповідно до чинних стандартів безпеки. Наприклад, ваш браузер зберігає ваше місцезнаходження, інформацію про обладнання та програмне забезпечення, інформацію про підключення, інформацію в соціальних мережах і т.д. Користувачі також уразливі завдяки надбудовам, які можуть запропонувати ці сучасні браузері.

Варто відзначити, що великий акцент на анонімність дій користувача робиться вже зараз. Більшість компаній (Yandex, Google, Opera) випускає браузері із вбудованим анонімайзером, який не запам'ятовує дані (паролі, логіни, адреси електронної пошти тощо), введені користувачем, і забороняє сайтам відстежувати дії відвідувачів через cookie. Сам користувач може активувати або дезактивувати анонімайзер у будь-який час. Проте не можна це сказати про використання ваших даних в комерційних цілях, наприклад для таргетованої реклами, тому браузері Web 3.0 мають рішення цих проблем.

Рішення полягає у використанні децентралізованого браузера, який забезпечує найкращу екосистему з використанням технології блокчейн. Наприклад, браузер Brave дозволяє вам продавати ваші дані для використання в таргетованій рекламі, або зберегти їх анонімними. Ще одна важлива частина браузерів Web 3.0 - метамова, за допомогою якої ви зможете отримувати більш якісні результати запиту, тому що система буде розуміти, що саме ви маєте на увазі. Важливо розуміти що всі сучасні проекти які використовують концепції Web 3.0 ще розвиваються, і не є досконалими.

Сьогодні не можна сказати що технології Web 3.0 мають великий вплив на всю мережу інтернет, тому що це лише майбутнє, яке вже сьогодні починає наставати, тому можна з упевненістю стверджувати, що Web 3.0 — це наступний логічний етап розвитку ідей, що зумовлюють поняття Web 2.0. Тим не менш, як такого Web 3.0 ще не існує. Та маса вимог, які висувуються сьогодні до мережі, ідеї, що кристалізуються серед Web 2.0, знайдуть свою реалізацію у Web 3.0. Важливо розуміти, що Web 3.0 не є надбудовою над другою версією.

ЛІТЕРАТУРА

1. Александр Фельдман. Интернет третьего поколения [URL: https://www.e-xecutive.ru/management/practices/945802-internet-tretego-pokoleniya](https://www.e-xecutive.ru/management/practices/945802-internet-tretego-pokoleniya) (Последний доступ: 06.11.2022)
2. В.А. Шачкова. Возможности системы Web 3.0 в современном медиапространстве: проблемы и перспективы [URL: https://cyberleninka.ru/article/n/vozmozhnosti-sistemy-web-3-0-v-sovremennom-mediaprostranstve-problemy-i-perspektivy](https://cyberleninka.ru/article/n/vozmozhnosti-sistemy-web-3-0-v-sovremennom-mediaprostranstve-problemy-i-perspektivy) (Последний доступ: 06.11.2022)
3. Binance academy. What Is Web 3.0 and Why Does It Matter? [URL: https://academy.binance.com/en/articles/the-evolution-of-the-internet-web-3-0-explained](https://academy.binance.com/en/articles/the-evolution-of-the-internet-web-3-0-explained) (Last accessed: 06.11.2022).
4. Євгенія Смерека. Що таке web 3.0, чому великі компанії в нього інвестують і коли настане епоха інтернету нового покоління. [URL: https://mc.today/uk/web-3-0/](https://mc.today/uk/web-3-0/)

МЕДВЕДЕНКО Владислав Олексійович – студент 1-го курсу магістратури; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: xa12284100@student.karazin.ua; vladshaxtar@gmail.com; ORCID: 0000-0001-6502-030X.

Наукові інтереси:

– веб розробка, розробка мобільних застосунків.

МІРОШНИЧЕНКО Д.О., ТОЛСТОЛУЗЬКА О.Г., ПЕТРИЧЕНКО М.О.

МОДЕЛЬ АВТОМАТИЗАЦІЇ РОЗГОРТАННЯ ІНФРАСТРУКТУРИ ДЛЯ СИСТЕМИ УПРАВЛІННЯ КОНТЕНТОМ В ХМАРНОМУ СЕРЕДОВИЩІ

Актуальність: в ході розвитку інформаційних технологій в останні роки великої популярності набули хмарні технології. До таких хмарних технологій можна з упевненістю віднести хмарні платформи, які надають можливість розгортати серверну інфраструктуру у хмарах, а не безпосередньо на фізичних серверах, а також автоматизувати цей процес. У цій статті представлено модель одного з можливих варіантів автоматизації розгортання такої інфраструктури.

Мета: ця робота присвячена розробці та демонстрації моделі одного з можливих варіантів автоматизації розгортання інфраструктури для системи управління контентом у хмарному середовищі. Оскільки можливих варіантів програмного забезпечення, за допомогою яких можна реалізувати подібну модель є декілька, то розроблена модель має підходити під різні інструменти розробки.

Методи дослідження: модель представлено у даній статті було розроблено за допомогою методології IDEF0. Для практичної реалізації даної моделі було використано засоби інструментарію Docker для управління ізольованими контейнерами.

Результати: розроблено модель автоматизації процесу розгортання інфраструктури для системи управління контентом в хмарному середовищі, де всі важливі етапи розбиті на окремі блоки, а також реалізовано один з блоків розробленої моделі.

Висновки: визначено, що дана модель автоматизації, розроблена за методологією IDEF0, цілком підходить для поставленої задачі, адже включає в себе всі важливі етапи автоматизації розгортання інфраструктури у хмарі без прив'язки кожного етапу до об'єктів, які можуть забезпечити реалізацію.

Ключові слова: автоматизація, модель, хмара, блок, інфраструктура.

Вступ

В ході розвитку інформаційних технологій в останні роки великої популярності набули хмарні технології. Нині існує тенденція витіснення більшості звичайних технологій хмарними, що цілковито може бути втілено у майбутньому. До таких хмарних технологій можна з упевненістю віднести хмарні платформи, які надають можливість розгортати серверну інфраструктуру у хмарах, а не безпосередньо на фізичних серверах. Використання хмарних платформ надає можливість налаштовувати сервіси швидше і масштабувати їх відповідно потребам або навантаженням. Оскільки всі сервери та мережеве обладнання розташовано у хмарі, то цілком зникає потреба у слідкуванні за станом обладнання, що у свою чергу зменшує витрати на обслуговуючий персонал, а також впливає на ефективність виконуваної роботи, оскільки вся увага приділяється безпосередньо виконанню поставлених задач.

Опис моделі

Розгортання інфраструктури у хмарному середовищі може виконуватися безпосередньо вручну, але це має наступні недоліки[1]:

- витрата великої кількості часу, якщо інфраструктуру потрібно розгорнути не один раз;
- людський фактор, який веде за собою можливі помилки.

Оскільки інфраструктура розгортається у хмарі, то є можливість автоматизувати цей процес, починаючи з моменту створення сервера у хмарі і до моменту видачі готової інфраструктури[3].

В даній статті представлено результати розробки моделі процесу автоматизації розгортання інфраструктури для системи управління контентом в хмарному середовищі (рис.1), а також один з можливих варіантів практичної реалізації одного з блоків даної моделі.

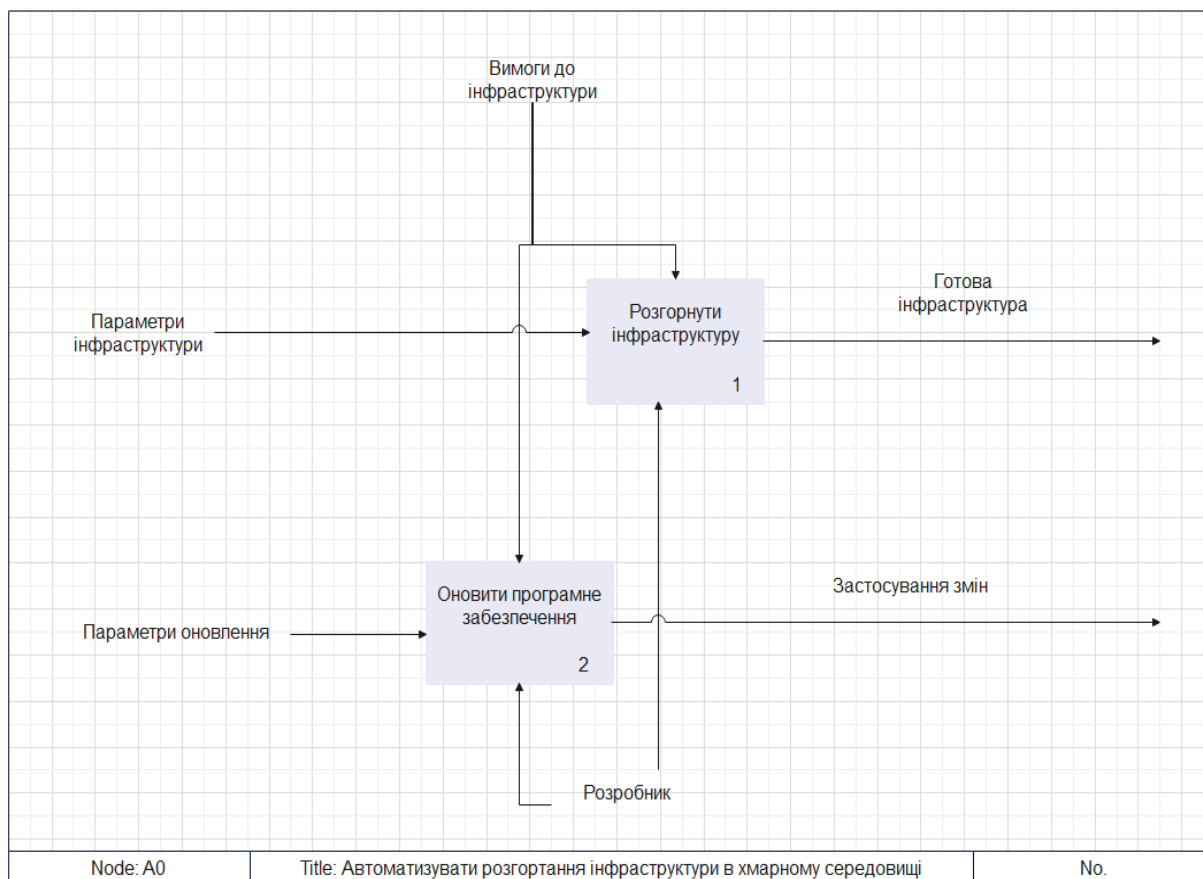


Рис. 1 IDEF0 діаграма процесу автоматизації розгортання інфраструктури у хмарі

Опис моделі виконано за стандартом IDEF0 [2]. На рисунку 1 зображено батьківську діаграму, яка складається з 2 батьківських блоків, щоб показати, яким загальним чином відбувається процес автоматичного створення інфраструктури і її оновлення. Основними блоками продемонстрованими в даній моделі є наступні.

Блок 1 (рис.1) розгортання інфраструктури. Цей блок відповідає за прийом вхідних вимог для інфраструктури, таких, як операційна система сервера, характеристики сервера, специфікація мережі, версія системи управління контентом(СУК) та бази даних (БД), специфікація резервного копіювання тощо. Під час роботи з цим блоком відбуваються процеси створення, як самої інфраструктури за вказаними параметрами, так і сервера і розміщення на ньому системи управління контентом та бази даних, і на виході ми отримаємо повністю готову інфраструктуру готову до експлуатації. Розробник відповідає за створення програм та контроль результатів виконання. Сам же процес розгортання інфраструктури відбувається автоматично без зовнішнього втручання. [6].

Блок 2 (рис.1) відповідає за оновлення програмного забезпечення. Цей блок виконується окремо від блоку 1. Він приймає вимоги до можливих оновлень в інфраструктурі та викликає відповідне програмне забезпечення, що може виконати оновлення. На вхід подаються відповідні вхідні параметри для оновлення програмного забезпечення на сервері а також версії системи управління контентом і відповідальна за оновлення система застосовує зміни в хмарі [5]. На виході ми отримаємо оновлення інфраструктури, яка була створена в блоці 1.

Розглянемо детальніше кожен з батьківських блоків.

Для батьківського блоку 1 «Розгорнути інфраструктуру» (рис.1) було розроблену наступну діаграму(див. рис. 2).

Кожен блок даної моделі розробляється та контролюється розробником, проте він не втручається безпосередньо в процес розгортання.

Блок 11 (рис.2) приймає дані для розгортання інфраструктури. Цей блок призначений для отримання всіх специфікацій, пов'язаних з інфраструктурою: операційна система сервера,

характеристики сервера, специфікація мережі, версія системи управління контентом (СУК) та бази даних (БД), специфікація резервного копіювання тощо.

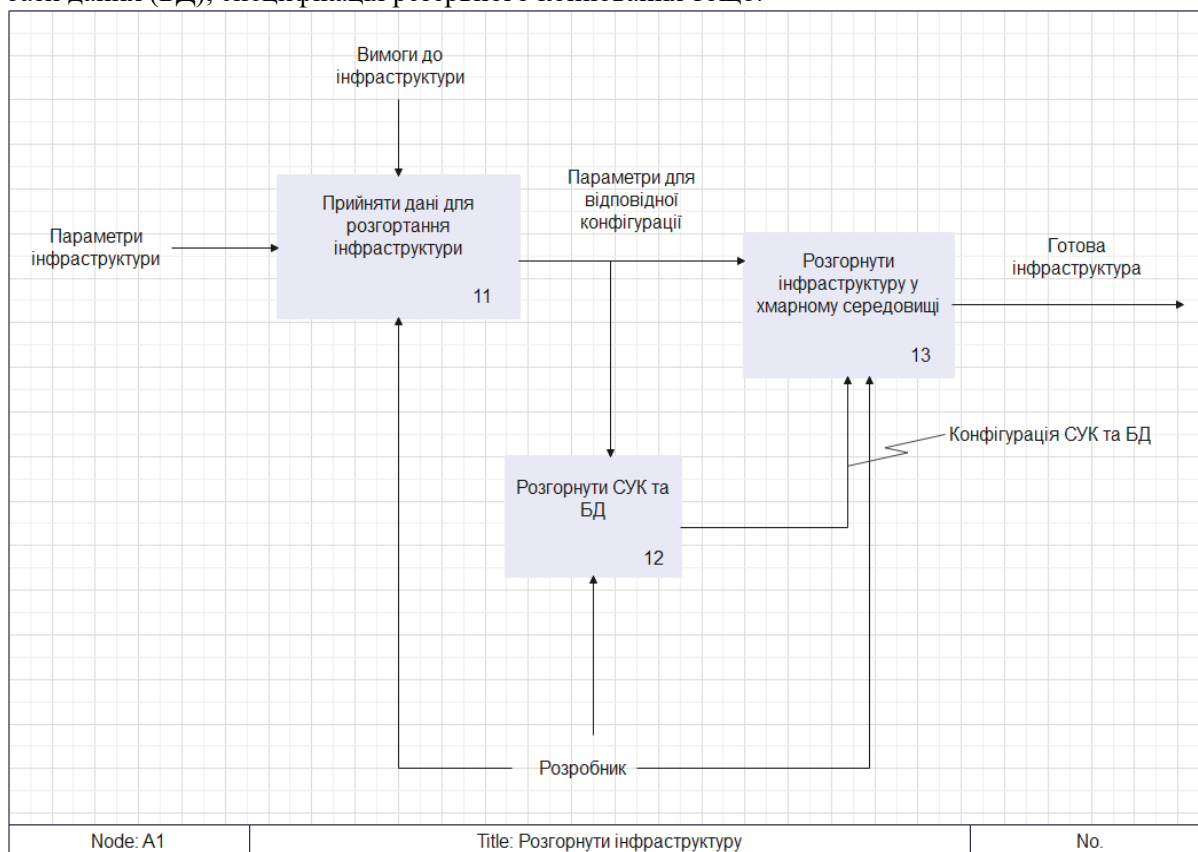


Рис. 2 IDEF0 діаграма процесу розгортання інфраструктури у хмарі

Таким чином, даний блок виступає інтерфейсом для взаємодії між розробником та компонентами, які відповідають за розгортання інфраструктури. На виході ми отримуємо перенаправлення параметрів відповідним програмним забезпеченням та активацію процесу розгортання [4].

Блок 12 (рис.2) виконує дії по розгортанню СУК та БД. На вхід блоку подаються параметри для системи управління контентом та бази даних (їхні версії), які застосовуються в конфігурації, згідно якої надалі відбудеться створення СУК та БД на сервері у хмарному середовищі.

Блок 13 (рис.2) розгортає інфраструктуру у хмарному середовищі. Цей блок є найважливішим у моделі, оскільки на даному етапі відбувається взаємодія з хмарним середовищем та розгортання всіх потрібних елементів інфраструктури згідно вказаних параметрів. Також даний блок використовує конфігурації з блоку 12 та застосовує їх у хмарному середовищі.

Другим батьківським блоком (рис.1) є «Оновити програмне забезпечення». IDEF0 діаграма процесу оновлення інфраструктури у хмарі представлена на рисунку 3.

Аналогічно з батьківським блоком 1 кожен етап створюється та контролюється розробником, проте він не втручається в сам процес оновлення. Кожен етап оновлення відбувається автоматично.

Блок 21 (рис.3) приймає дані для оновлення інфраструктури. Цей блок відповідає за прийом параметрів для програмного забезпечення, яке потрібно оновити, та передає ці дані відповідному сервісу, який буде відповідальний безпосередньо за оновлення. Таким чином даний блок виступає інтерфейсом для взаємодії між розробником та компонентами, які відповідають за розгортання інфраструктури [4].

Блок 22 (рис.3) оновлює інфраструктуру у хмарі. Даний блок є засобом для управління конфігурацією та централізованої установки застосунків на віддалених серверах (у даному

випадку програмного забезпечення на сервері у хмарі а також версії системи управління контентом) [5].

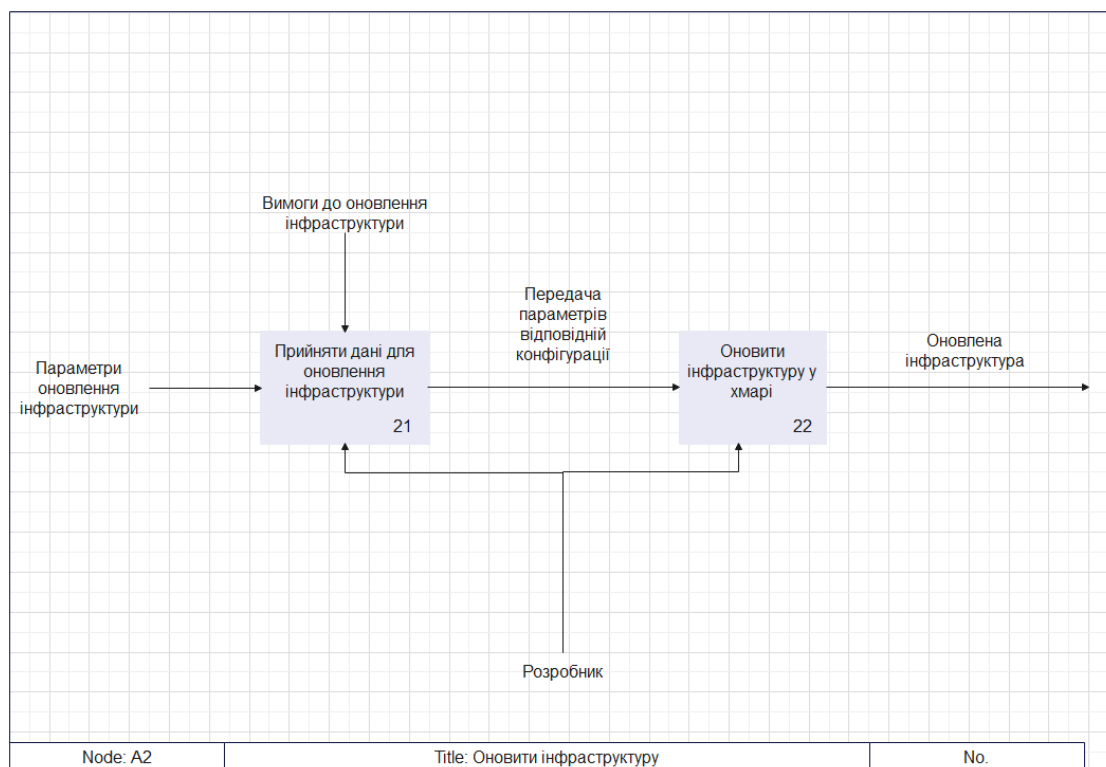


Рис. 3 IDEF0 діаграма процесу оновлення інфраструктури у хмарі

Реалізація представленої моделі

Згідно вищеописаної моделі автоматизації розгортання інфраструктури у хмарі вже було створено конфігурацію для розгортання системи управління контенту та бази даних для неї. Така конфігурація запускає 2 сервіси у віртуальному середовищі на сервері:

-1 сервіс – системі управління контентом, до якої можна звернутися за вказаним портом через програмне забезпечення для перегляду веб-сайтів;

-2 сервіс – безпосередньо база даних, яка вже з'єднана з системою управління контентом. Звернення до бази даних може відбуватися як через програмне забезпечення у браузері, так і через пристрої для забезпечення взаємодії оператора комп'ютера з операційною системою (консоль).

В якості інструментарію для практичної реалізації даної моделі, а саме для управління ізольованими контейнерами, було використано засоби Docker. Docker розподіляє ядро операційної системи на окремі віртуальні контейнери, які є окремими процесами. Це веде до більш ефективного використання ресурсів системи, швидкого розгортання програм а також масштабування та перенесення програмних продуктів в інше середовище без втрат стабільності роботи [7].

Висновки

Розроблена модель автоматизації розгортання інфраструктури для системи управління контентом у хмарному середовищі представляє собою один із можливих варіантів автоматизації створення інфраструктури у хмарі. Дана модель, розроблена у форматі IDEF0, є універсальною для великої кількості програмного забезпечення, яке може бути використано в ході практичної реалізації такої моделі. Також було реалізовано один з блоків розробленої моделі, а саме розгортання системи управління контентом та бази даних для неї за допомогою програмного забезпечення Docker.

ЛІТЕРАТУРА

1. What Is Cloud Automation and How Does It Benefit IT Teams? [Електронний ресурс] – Режим доступу: <https://thenewstack.io/what-is-cloud-automation-and-how-does-it-benefit-it-teams/>
2. IDEF0 Diagram Software [Електронний ресурс] – Режим доступу: <https://www.edrawsoft.com/idef0-flowcharts.html>
3. What is Infrastructure as Code (IaC)? [Електронний ресурс] – Режим доступу: <https://www.redhat.com/en/topics/automation/what-is-infrastructure-as-code-iac>
4. Infrastructure as Code Concepts [Електронний ресурс] – Режим доступу: https://kb.novaordis.com/index.php/Infrastructure_as_Code_Concepts#Stack_Configuration
5. Automation Servers [Електронний ресурс] – Режим доступу: <https://learn.microsoft.com/en-us/cpp/mfc/automation-servers?view=msvc-170>
6. What is IT infrastructure automation and how to implement it? [Електронний ресурс] – Режим доступу: <https://maximaconsulting.com/newsroom/it-infrastructure-automation>
7. Що таке Docker і навіщо він? [Електронний ресурс] – Режим доступу: <https://qagroup.com.ua/publications/shcho-take-docker-i-navishcho-vin/>

МІРОШНИЧЕНКО Дмитро Олександрович – студент 1 курсу магістратури групи КУ-51 кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: ha12284086@student.karazin.ua; ORCID: 0000-0003-1159-0985.

Наукові інтереси:

- *хмарні обчислення.*
- *моделювання процесів автоматизації.*

ТОЛСТОЛУЗЬКА Олена Геннадіївна – д. т. н., професор, професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: elena.tolstoluzka@karazin.ua; ORCID: 0000-0003-1241-7906.

Наукові інтереси:

- *багатопроцесорні та розподілені системи автоматизації.*
- *технології розподілених систем та паралельні обчислення.*

ПЕТРИЧЕНКО Максим Олександрович – DevOps Team Lead; IT Craft; проспект Гагаріна, 43/2, Харків, Харківська область, Україна, 61001; e-mail: maxim.petrichenko@itechcraft.com

Наукові інтереси:

- *технології віртуалізації та контейнерів.*
- *методи зберігання та обчислення великих об'ємів даних.*

РОЗРОБКА СИСТЕМИ МОНІТОРИНГУ МЕРЕЖІ ТЕЛЕФОННИХ ЛІНІЙ

Вступ

Контактні центри є сьогодні одним із найбільш поширених та ефективних способів роботи з клієнтами-споживачами бізнесу. Так, залежно від специфіки конкретного бізнесу, будь то, наприклад, банківська система або шиномонтаж, процес обробки клієнтських запитів може відрізнятися. Головною проблемою контакт-центрів сьогодні є необхідність утримувати велику кількість працівників (абонентів, «продажників», менеджерів та інших) і, як наслідок, забезпечувати контроль за ними. Найчастіше, витрати на заробітні плати працівників та пов'язані з цим додаткові витрати можуть становити до 70% від загальних місячних витрат підприємства. Крім того, чим більша та чи інша фірма, тим складнішою технологічною структурою вона володіє. Завданням сучасних технологій в галузі роботи call-центрів є максимально автоматизувати бізнес-процеси, щоб знизити їх ресурсозатрати і підвищити якість сервісу, що надається, якість зв'язку або якість обслуговування операторами.

Постановка задачі

Програмне забезпечення, що забезпечує можливість контролю за бізнес-процесами контакт-центрів априорі має багато індивідуальних особливостей структури компанії, для якого воно розроблялося. Так, зокрема, система моніторингу, що розробляється в цій роботі, повинна мати наступні функціональні можливості:

- авторизація користувачів;
- відображення поточної інформації щодо виділених ліній;
- відображення поточної інформації щодо дзвінків;
- відображення поточної інформації щодо операторів;
- моніторинг за окремими операторами;
- прослуховування дзвінків поточних та завершених з метою контролю якості;

Практична реалізація проекту

Практичну реалізацію системи моніторингу та її функціональності можна розбити на такі основні етапи:

- розробка дизайну інтерфейсу програми та її «карти»; – верстка сторінок;
- створення структури бази даних; – розробка програмних алгоритмів та розгортання програми.

При цьому необхідно мати на увазі, що система моніторингу, що розробляється, повинна володіти позначеними в задачі даної роботи функціональними можливостями.

Розробка дизайну інтерфейсу програми та її «карти»

Розроблена в даній роботі система моніторингу є односторінковим веб-застосунком (single page application - SPA). Незважаючи на те, що користувач фактично не змінює URL сторінок у браузері, вміст програми динамічно змінюється, начебто користувач працював зі звичним багатосторінковим веб-сайтом. При цьому швидкість завантаження віртуальних сторінок значно вища, тому що немає необхідності при переході між логічними блоками вебзастосунку підвантажувати нову HTML-структуру сторінки з серверної частини програми.

Після створення макетів основних сторінок на першому етапі відбувається створення дизайну допоміжних сторінок. Це пов'язано з тим, що залежно від того, наскільки повно реалізовано функціональність програми в дизайні головних екранів, визначається функціональність допоміжних сторінок.

Сторінки веб-зв'язу крім надання доступу до функціональності системи моніторингу, повинні також мати інтуїтивно-зрозумілий дизайн і не бути візуально перевантаженим контентом.

Макет сторінки відображення інформації списком наведено на рисунку 1.

Лінійний номер	Оператор	Активних викликів	Вільно	Занято	На паузі	Недоступно
Черга 8441 Mosautoshina	Усього операторів: 27	8	1	10	1	15
Черга 8241 Adamas	Усього операторів: 7	0	0	6	1	0
Черга 8242 Adamas_callcenter	Усього операторів: 8	2	0	7	1	0
Черга 8243 Adamas_direktor	Усього операторів: 0	0	0	0	0	0
Черга 8071 NaukaNet 1 Line	Усього операторів: 7	0	0	5	1	0
Черга 8080 MB-Samara	Усього операторів: 5	1	0	5	0	0
Черга 8191 Позитроника	Усього операторів: 5	0	2	3	0	0

Рис. 1 Сторінка відображення інформації списком

Мета цього виду відображати ключову інформацію щодо вибраних телефонних ліній на поточний час. Кожна сторінка списку містить лише десять ліній. Елементи кожної лінії, такі як «Всього операторів» розкривають більше інформації про операторів на даній лінії та про дзвінки на лінії.

Створення структури бази даних

Створення структури бази даних відбувалося в такий спосіб, щоб забезпечити найбільшу (прозору) простоту взаємозв'язку сутностей. База даних має такі ключові сутності:

- телефонні лінії (черги);
- оператори;
- дзвінки;

– Користувачі, до яких належать користувачі системи з ролями - адміністратор, менеджер, менеджер замовника, супервайзер, спостерігач.

Сутності представлені в базі даних окремими відносинами, що зберігають відповідно інформацію про користувачів, телефонні лінії, операторів і дзвінки на лініях.

Особливу увагу варто приділити дзвінкам, які, окрім тільки текстової інформації, мають звукові доріжки. Усі аудіофайли транслюються у режимі живого часу на зовнішній накопичувач серверного обладнання. Відношення «Дзвінки» зберігає лише назви аудіофайлів, згенеровані за принципом суміщення дати та часу початку та закінчення дзвінка, а також номери черги та оператора. Для прискорення пошуку медіа-файлів використовується хешування назв та їх кластеризація у словниках бази даних. Так, файли розбиваються на групи в залежності від номера телефонної лінії, тобто файли розташовані в сусідніх аркушах пам'яті, що сприяє швидкому пошуку та вийманню потрібних даних.

Спрощена ER-діаграма бази даних представлена рисунку 2.

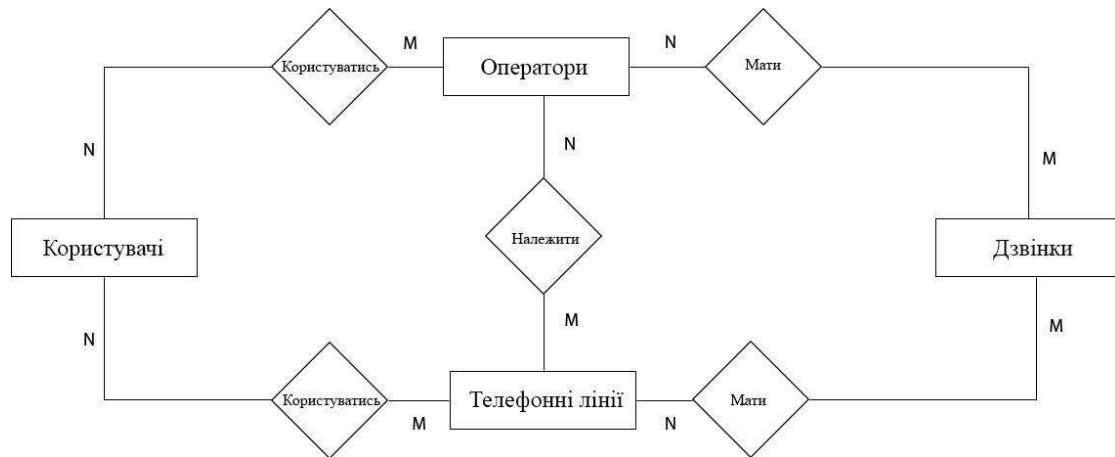


Рис. 2 – Спрощена ER-діаграма

Виходячи з ER-діаграми видно, що безліч користувачів системи можуть мати (мати доступ) до багатьох операторів та телефонних ліній. У свою чергу, кожен з операторів і кожна з телефонних ліній можуть бути переглянуті одним або декількома користувачами. Аналогічним зв'язком багатьом мають відносини «Оператори», «Телефонні лінії» і «Дзвінки». При нормалізації даних кожен зв'язок багатьом перетворена на два зв'язку одним до багатьох буферних таблиць.

Розробка програмних алгоритмів та розгортання програми

Усі алгоритми, що забезпечують працездатність системи моніторингу, становлять "двигун" розробленого програмного забезпечення. У цій роботі основним інструментом розробки програмної складової є мова програмування JavaScript. Іншими технологіями, що використовувалися при розробці системи моніторингу, стали React, Redux, Node.js (Koa), MySQL, WebSockets, Asterisk Management Interface (AMI), WebRTC, HTTP, SIP, TCP, та багато інших. Система моніторингу має багату функціональність, а тому алгоритмів, що забезпечують її працездатність, безліч. Серед них у цій роботі розглянуто алгоритми взаємодії АТС (Автоматична телефонна станція), сервера та інтерфейсу.

Основна сутність алгоритмів взаємодії сервера з автоматизованою телефонною системою компанії – це, використовуючи підключення сервера додатку до АМІ АТС отримувати інформацію про зміни, що виробляються на телефонних лініях, проводити їхню обробку, а потім зберігати оброблену інформацію до бази даних. Тут найважливішу роль відіграє позначення з боку клієнта, які саме події АМІ цікавлять систему моніторингу для проведення обробки лише невеликої кількості подій.

До подій належать такі:

- QueueCallerJoin – надходження нового дзвінка на певну лінію;
- AgentConnect - взяття поточним оператором поточного дзвінка;
- Hangup - закінчення поточного дзвінка;

Інформація надсилається з АТС у неструктурованому вигляді за постійно відкритим підключенням і потім потрапляє на спеціальні обробники подій на серверній частині системи моніторингу, що розбиває отриманий рядок на JavaScript-об'єкти:

```

case 'queuemember': { agent: event.stateinterface.length
=== 4 ?
event.stateinterface :
event.stateinterface.
split('/')[1] , queue:
event.queue, status:

```

```
event.st  
atus,  
penalty:  
event.pe  
nalty,  
paused:  
event.pa  
used,  
lastcall  
:  
event.la  
stcall }
```

Згодом дані об'єкти використовуються, по-перше, для збереження інформації до бази даних програми з метою архівування, по-друге, для надсилання інформації клієнтам вебзастосунку. З урахуванням того, що структура таблиць АМІ-евентів у базі даних підготовлена спочатку для зручного зберігання оброблених даних, для збереження інформації застосовуються прості SQL-запити. Для надсилання підготовлених даних на клієнтську частину програми використовується WebSocket-підключення, що відкривається на час користування користувача програмою.

ЛІТЕРАТУРА

1. Bähr, K. Cross-platform development with JavaScript [Електронний ресурс] / Katharina Bähr. - Режим доступу: <https://www.zuehlke.com/blog/en/cross-platform-developmentwithjavascript>, вільний
2. Node.js MongoDB Create Database [Електронний ресурс] / W3Schools. - Режим доступу: https://www.w3schools.com/nodejs/nodejs_mongodb_create_db.asp, вільний
3. Як працюють дата-центри: сьогодні та завтра [Електронний ресурс] / Habr. - Режим доступу: https://habr.com/ru/company/vps_house/blog/343282, вільний.
4. Ляшенко В.П., Кобильська О.Б., Дям'янченко О.П. Математичні моделі теплообміну з умовами імпедансного типу у багатошарових областях. *Вісник Кременчуцького національного університету імені Михайла Остроградського*. 2017. Вип. 6/2017 (106). С. 37–43.

МОРОЗ Євген Русланович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м.Харків, Україна, 61022; e-mail: ha11867786@student.karazin.ua; ORCID 0000-0002-1607-0877.

Наукові інтереси:

- *Комп'ютерні мережі, моделювання комп'ютерних мереж.*

ХРУСЛОВ Максим Михайлович – кандидат фізико-математичних наук, доцент кафедри електроніки і управляючих систем; Харківський національний університет імені В.Н.Каразіна;

майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: maksym.khruslov@karazin.ua; ORCID 0000-0001-9639-9340

Наукові інтереси:

- *Аналіз даних, Машинне навчання, Інтернет речей.*

ОРЛЕНКО О.А, МІШИН О.В.

РОЗМІТКА ПОЧАТКУ ТА КІНЦЯ ТРЕКІВ В АУДИОФАЙЛІ МЕТОДОМ КЛАСТЕРИЗАЦІЇ

The topic of the work is an overview of the algorithm for marking the beginning and end of tracks in an audio file using clustering methods. This will eliminate the need to store huge databases for the analysis of audio files by teaching methods with the teacher. In addition, you will not need to have pre-recorded fragments of voice, sounds, etc. to receive the tagged audio file.

To use this approach, it is necessary to prepare data, conduct clustering and evaluate the results. That is why the aim of the work is to study the necessary methods of data processing and clustering with their subsequent application to the markup.

The main research methods are experiments in the form of comparison of performance results with the use of a certain method of data processing and without it. As for clustering methods, in this case a theoretical comparison of the feasibility of their use.

After all the research and experiments, it was found that the use of clustering methods to determine the beginning and end shows quite good results, but not good enough accuracy. Therefore, a list of possible methods to improve accuracy was put forward.

In general, it can be established that an algorithm for marking the beginning and end of tracks in an audio file using clustering was obtained, a correctness study was performed, accuracy problems were identified and possible steps to improve accuracy were identified for further study.

Ключові слова: *artificial intelligence, python, machine learning, data processing, clusterization, spectrogram.*

В епоху неспинного розвитку інформаційних технологій все більшої популярності набуває вирішення проблем та проведення прогнозування за допомогою штучного інтелекту[1]. Аналіз аудіофайлів теж не є виключенням. Основною перевагою є то, що з'являється можливість звільнити людину від такої роботи як багаточасове прослуховування аудіотреку з метою визначення початку та кінця треків всередині. Дана проблема може бути вирішена декількома кроками.

1) Навчання з вчителем та можливо використання неймереж. Підхід є досить популярним та демонструє гарні результати, проте головною проблемою є необхідність мати базу даних на основі якої буде проведено навчання. Ми ж розглядаємо момент коли про аудіотрек відомо максимум кількість треків, тому даний підхід не підходить.

2) Навчання без вчителя та кластеризація. Є також досить популярним підходом, та використовується в основному там де неможливо використати навчання з вчителем через відсутність бази даних чи некоректність роботи. Для нашого випадку досить гарно підходить, тому буде використаний цей підхід.

Методи кластеризації які використовуються при навчанні без вчителя також є досить різними та використовуються в різних ситуаціях. Серед основних є[2]:

1) метод к-середніх (KMeans) – найбільш популярний метод кластеризації на даний момент. Основною ціллю даного методу є мінімізація сумарних квадратичних відхилень точок кластерів від центроїдів(середніх координат). Даний метод передбачає необхідність вказати кількість кластерів. Основні обчислення зводяться до вирахування відстані точок від центру кластера, та розрахунок нового розташування центру;

2) ієрархічна кластеризація. Суть даного алгоритму полягає в тому, що спочатку кожній точці присвоюється окремий кластер, а далі 2 кластера об'єднуються в один, доки не буде одного загального кластеру. Алгоритм не потребує кількості кластерів, так як вони визначаються автоматично, проте для роботи даного алгоритму необхідна велика кількість ресурсів;

3) метод кластеризації на основі щільності (DBSCAN) – популярний алгоритм кластеризації даних, що використовується в аналізі даних в якості заміни методу k-середніх. Метод не потребує вказання кількості кластерів, проте необхідно налаштувати два параметра: *eps* та *min_samples*. *eps* – це максимальна відстань між сусідніми точками, а *min_samples* мінімальна кількість точок в окружності(сусідів), що можуть представляти один кластер. Загалом ці параметри є по замовчуванню, проте для кращої роботи методу їх все ж потрібно налаштовувати;

4) метод найближчих сусідів – метод кластеризації, основним алгоритмом якого є віднесення точки, кластер для якої визначається до того кластеру, до якого відносяться більшість її

найближчих сусідів. Даний алгоритм є найпростішим та через це має не досить гарну точність. Швидкість роботи також не досить радує.

Згідно з наведеними вище методами кластерів можна зробити такі висновки, для нашого випадку:

1) метод найближчих сусідів нам не підходить, так як він не досить точний та не досить швидко працює;

2) метод ієрархічної кластеризації не підходить через необхідність великої кількості ресурсів;

3) метод кластеризації на основі щільності – є заміною яку використовують замість метод k-середніх. Визначення максимальної відстані між точками, що належать одному кластеру в нашому випадку не є сталою величиною, так як передбачається робота з різними аудіо-файлами, тому кращим варіантом буде вказати необхідну кількість кластерів.

З вище вказаного слідує, що буде використовуватися метод k-середніх.

Для використання вказаного методу необхідно підготувати дані:

1) для аналізу аудіофайлу часто використовують спектограму, саме тому першим етапом в підготовці даних є завантаження аудіофайлу, перетворення його в спектограму, та завантаження спектограми. Приклад спектограми наведено на рисунку 1.

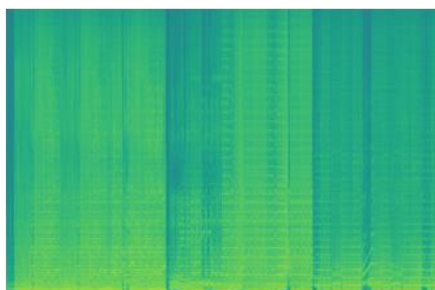


Рис. 1 спектрограма, що має бути отримана в результаті виконання вказаних кроків

2) для відображення потужності звуку в спектограмі використовують різні кольори, тому при завантаженні ми матимемо 3-вимірний масив який необхідно буде привести до 2-вимірному.

3) для цього необхідно взяти медіану кольору, та замінити нею масив, що відповідає за колір. Тепер необхідно привести масив до одновимірному. В нашому випадку для кластеризації будемо використовувати два виміри: спектограма та час.

4) Для того щоб привести 2-вимірний масив в одновимірний використаємо метод головних компонент. Він дозволить зменшити розмірність та зберегти максимально можливе значення інформації. Тепер необхідно підчистити дані від викидів, та нормалізувати їх. Це дозволить покращити процес кластеризації.

5) Для видалення викидів можна використати один з відомих методів автоматичного знаходження та видалення викидів. Наприклад IsolationForest. Далі провести нормалізацію за відомою формулою:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

де x_{norm} – нове значення, що буде отримано після нормалізації,

x – старе значення, що буде нормалізовано,

x_{min} – мінімальне значення, в всьому масиві даних,

x_{max} – максимальне значення в всьому масиві даних.

На цьому підготовка даних для кластеризації є завершена.

Використовуючи наведене вище були отримані наступні результати:

1) Прискорення – зменшення часу на отримання часових міток шляхом використання кластеризації замість прослуховування аудіотреку. Розраховується за формулою:

$$p = \frac{T_{track}}{T_{cluster}} \quad (2)$$

де p – значення прискорення,

T_{track} – тривалість треку,

$T_{cluster}$ – тривалість кластеризації.

В результаті тестування було встановлено, що прискорення коливається в межах від 55 до 145 та в середньому становить 95. Графічне представлення прискорення наведено на рисунку 2.

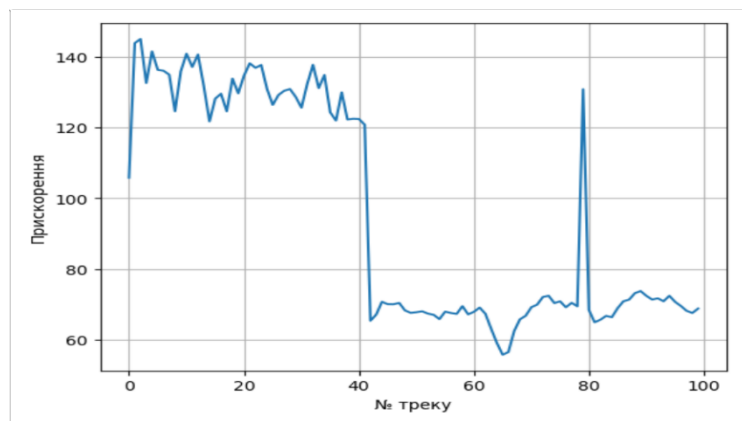


Рис. 2 Прискорення, що було отримане в результаті тестування

2) Втрати даних – відхилення отриманої часової мітки від реальної. Демонструє яка кількість інформації з треку буде втрачена в результаті похибки при розбитті. Розраховується за формулою:

$$dl = \frac{T_{deviation}}{T_{track}} \quad (3)$$

де dl – значення втрати даних,

T_{track} – тривалість треку,

$T_{deviation}$ – відхилення.

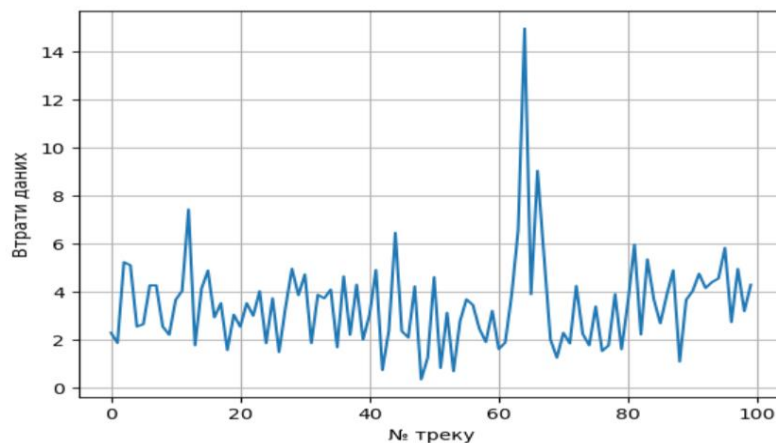


Рис. 3 Втрати даних, що була отримане в результаті тестування

В результаті тестування(рис. 3) було встановлено, що втрата даних коливається в межах 0,5 до 15 відсотків, що говорить про те, що в гіршому випадку ми втратимо близько 15 відсотків даних, тому існує необхідність покращення точності.

Можливими методами покращення точності є:

1) збільшення частоти дискретизації чи кількості точок для дискретного перетворення Фур'є. Ця дія відноситься до підготовки даних перед кластеризацією, а саме – створення спектрограми. Внаслідок збільшення вказаних характеристик припускається, що це дозволить збільшити набір даних, внаслідок зменшення часового кроку. Тобто мається на увазі те, що якщо наприклад на даний момент, кількість відліків на одиницю часу рівно 2 (F_s в параметрах створення спектрограми), то при збільшенні можливе покращення розрізнення треків, що призведе до кращої точності, проте при цьому збільшуються затрати на обчислення, що потрібно враховувати при виконання цієї дії;

2) покращення обробки викидів. В нашому випадку процес знаходження викидів, було перекладено на алгоритм Isolation Forest. Для покращення роботи, необхідно провести тонке налаштування цього алгоритму та підібрати оптимальні параметри. Так як обробка викидів впливає на правильність кластеризації, дана дія повинна покращити загальну точність роботи алгоритму.

3) обробка діапазонів кластерів. Через видалення викидів, інколи виникає ситуація, коли відстань між двома кластерами є досить великою. В даному випадку в нашій ситуації знаходиться середнє значення між кінцем першого кластеру та початком другого кластеру, та отримане значення трактується як нова межа між двома кластерами. В середньому це не має негативного впливу, та не погіршує точність, проте необхідно все ж відслідкувати роботу такого метода на більшій кількості даних. Якщо метод буде показувати не досить задовільні результати, необхідно визначити новий спосіб визначення нової межі. Ця дія не є обов'язковою так як, можлива ситуація, що завдяки наведеним вище діям, ситуація з великою відстанню між двома кластерами зникне, проте якщо це не так, то можливе рішення уже наведено.

4) Використання додаткового методу обробки даних та кластеризації для отримання додаткової інформації, щодо аудіотреку.

ЛІТЕРАТУРА

1. Russell & Norvig. Artificial Intelligence A Modern Approach:[Електронний ресурс]. 2003.
URL:<https://cs.calvin.edu/courses/cs/344/kvinden/resources/AIMA-3rd-edition.pdf>
2. A. Jain, R. Dubes, Algorithms for Clustering Data, Prentice-Hall, 1981. ISBN 0-13-022278-X
URL: https://homepages.inf.ed.ac.uk/rbf/BOOKS/JAIN/Clustering_Jain_Dubes.pdf

ОРЛЕНКО Олександр Анатолійович – бакалавр; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022; e-mail: or.olexandr1@gmail.com; ORCID: 0000-0001-6078-7915.

Наукові інтереси:

- штучний інтелект;
- кластеризація даних;
- обробка аудіофайлів;

МІШИН Олександр Вікторович – ст. викладач; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків, Україна, 61022; e-mail: o.v.mishyn@karazin.ua; ORCID: 0000-0001-7478-757X.

Наукові інтереси:

- штучний інтелект;
- логічне програмування;
- машинне навчання;
- нейронні мережі;
- комп'ютерне моделювання фізичних процесів;

ПЕТУХОВ І.І., ЛИСИЦЯ О.Ю.

CFD-МОДЕЛЮВАННЯ ТЕПЛОМАСООБМІННИХ ТА ГІДРОГАЗОДИНАМІЧНИХ ПРОЦЕСІВ У СТРУМЕНЕВИХ АПАРАТАХ

Показано переваги струминних апаратів при вирішенні задач використання надлишкового тиску однофазних середовищ та випарного охолодження рідини. Проаналізовано спрощені методи розрахунку, припущення при їхньому формуванні, ступінь обґрунтування та універсальності цих припущень. Розглянуто можливість зменшення числа припущень та уточнення деяких з них при використанні CFD-моделювання. Наведено основні рівняння та етапи CFD-моделювання тепломасообмінних та гідрогазодинамічних процесів у струминних апаратах, зокрема з використанням пакету ANSYS Fluent. З урахуванням прямокутного профілю каналу, задача попереднього моделювання може вирішуватися у двовимірній постановці. При розрахунку сопла заданої геометрії на вході задаються тиск і температура рідини. Неоднорідний у поперечному перерізі потік у камері змішування або у криволінійному каналі визначає параметри на вході в дифузور. Для однофазного потоку неоднорідність визначається профілем швидкості, для двофазного ще й профілем паровмісту. Базова система рівнянь для нерівноважного двофазного потоку з міжфазним енерго-масообміном базується на використанні рівнянь нерозривності, збереження імпульсу та енергії для кожної фази. Розрахунок здійснюється на основі нестационарного вирішувача Pressure Based Solver, моделі турбулентності Realizable k-ε. Параметри перед дифузором дозволяють визначити тиск і температуру потоку на його виході для граничного режиму, коли стрибок ущільнення розміщується на вході дифузору. При перевищенні цього тиску стрибок виходить з дифузору і витрата через нього зменшується. Сформована математична модель дозволяє розраховувати основні параметри робочого процесу та більш надійно проектувати газові ежектори системи кондиціонування повітря літака та струминні охолоджувачі рідини.

Ключові слова: струменевий апарат, CFD-моделювання, тепломасообмін, гідрогазодинаміка.

Вступ

Відмінною рисою струминних апаратів є простота конструкції та реалізація обмінних процесів між потоками, у тому числі різнофазними, при безпосередньому контакті. Ці якості визначили їх переваги для вирішення задач ефективного використання надлишкового тиску або наявної роботи одного із потоків, а також у якості тепло- масообмінних пристроїв. Прикладом використання надлишкового тиску може бути система кондиціонування повітря літака. Стиснене (первинне) повітря береться або безпосередньо від двигуна, або від допоміжної силової установки. Далі воно використовується для забезпечення заданих параметрів (температура, тиск) інших потоків, які забезпечують комфортне середовище для пасажирів, пілотів, вантажу, а також запобігання обледенінню зовнішніх поверхонь літака і двигуна [1]. Використання ежектора зменшує загальне споживання стисненого повітря і, таким чином, призводить до зниження витрати палива.

Струменеві апарати з різнофазними потоками використовуються для випарного охолодження рідин. Висока швидкість потоку і розвинена поверхня контакту фаз забезпечують малі габарити пристрою і швидкість охолодження 400...600 К/с [2]. Крім того, тут наявна робота відведеного пару корисно використовується для підвищення повного тиску рідини при їх спільному розширенні.

Однак струменеві апарати мають істотні недоліки, пов'язані з низьким ККД і вузькою зоною стійкої роботи при фіксованій геометрії, особливо при надкритичних перепадах тиску. Наприклад, для системи кондиціонування повітря літака, де досить широкий діапазон параметрів первинного повітря, і можлива їх ступінчаста зміна, необхідно використання кількох ежекторів, що вмикаються по чергові. Забезпечити ефективність і реалізувати керування такими системами неможливо без надійного математичного опису поведінки струминних апаратів для всіх режимів роботи в тому числі і перехідних.

Основна частина

Струменеві апарати системи відбору повітря використовують пружні середовища без фазових переходів. Такі апарати складаються з первинного сопла, камери всмоктування, камери змішування, каналу постійної, як правило, площі та дифузора (Рис. 1). За рахунок різниці тисків між потоками вторинний потік прискорюється, «всмоктується» основним. Передача кінетичної

енергії від одного потоку до іншого при безпосередньому контакті є загальною рисою робочого процесу струменевих пристроїв.

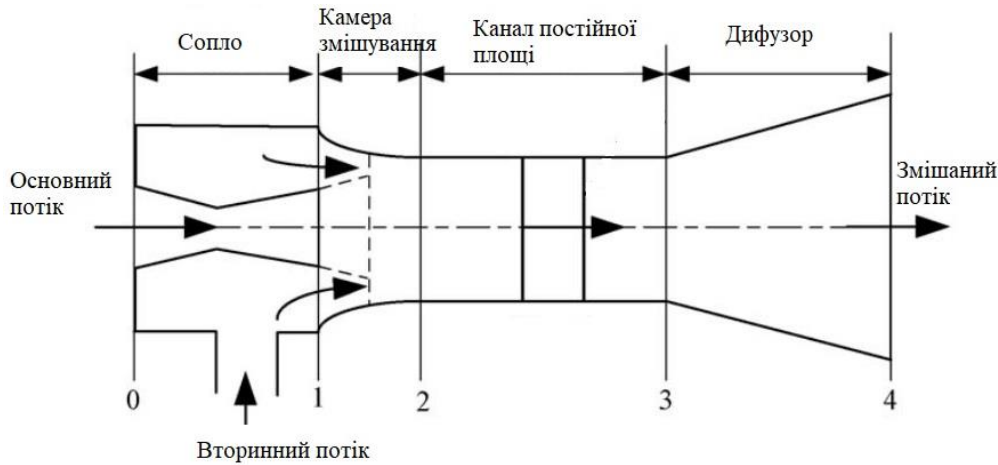


Рис. 1 Схема струменевого апарату системи відбору повітря літака

Математичний опис робочого процесу струминного апарату дуже складний внаслідок суттєвої нерівноважності процесів переносу маси, енергії та імпульсу, особливо у камері змішування та при сепарації фаз. Існуючі моделі робочого процесу у струминних апаратах зазвичай використовують спрощений підхід та суттєво обмежені припущення [3, 4]. CFD-моделювання в цьому сенсі має значні переваги тому, що здатне забезпечити більш точний та детальний опис характеристик потоку та процесів тепломасобміну. Задача додатково ускладнюється через комплексність процесів переносу у камері змішування, високу швидкість потоків, наявність перехідних процесів. Попередньо задача може вирішуватися у двовимірній постановці, оскільки ежектор зазвичай представляє собою профільований прямокутний канал з достатньо великою шириною. Доцільно проводити CFD-моделювання з використанням пакету ANSYS Fluent на основі нестационарного вирішувача Pressure Based Solver із застосуванням моделі турбулентності Realizable k-ε.

Базова система рівнянь для нерівноважного двофазного потоку базується на використанні рівнянь нерозривності, збереження імпульсу та енергії для кожної фази. В позначеннях [5] така система (1-3) для q-ї фази має наступний вигляд:

рівняння нерозривності (1):

$$\frac{\partial}{\partial t}(\alpha_q \rho_q) + \nabla \cdot (\alpha_q \rho_q \vec{v}_q) = \sum_{p=1}^n (\dot{m}_{pq} - \dot{m}_{qp}) + S_q \quad (1)$$

рівняння збереження імпульсу (2):

$$\begin{aligned} \frac{\partial}{\partial t}(\alpha_q \rho_q \vec{v}_q) + \nabla \cdot (\alpha_q \rho_q \vec{v}_q \vec{v}_q) = & -\alpha_q \nabla p + \nabla \cdot \vec{\tau}_q + \alpha_q \rho_q \vec{g} \\ & + \sum_{p=1}^n (\vec{R}_{pq} + \dot{m}_{pq} \vec{v}_{pq} - \dot{m}_{qp} \vec{v}_{qp}) + (\vec{F}_q + \vec{F}_{\text{lift},q} + \vec{F}_{\text{wl},q} + \vec{F}_{\text{vm},q} + \vec{F}_{\text{td},q}) \end{aligned} \quad (2)$$

рівняння збереження енергії (3):

$$\frac{\partial}{\partial t}(\alpha_q \rho_q h_q) + \nabla \cdot (\alpha_q \rho_q \vec{u}_q h_q) = \alpha_q \frac{dp_q}{dt} + \vec{\tau}_q : \nabla \vec{u}_q - \nabla \cdot \vec{q}_q + S_q + \sum_{p=1}^n (Q_{pq} + \dot{m}_{pq} h_{pq} - \dot{m}_{qp} h_{qp}) \quad (3)$$

Доцільним виглядає поелементне CFD-моделювання з використанням фрагментів наявних моделей розрахунку параметрів відсепарованої плівки та дифузора. При розрахунку сопла заданої геометрії на вході задаються тиск і температура рідини. На вході в дифузор параметри визначаються в залежності від характеристик неоднорідного у поперечному перерізі потоку у камері змішування або у каналі сепаратора. Для однофазного потоку неоднорідність визначається профілем швидкості, для двофазного – ще й профілем паровмісту. Параметри в перерізі перед дифузorzом дають можливість розрахувати тиск і температуру потоку на виході з дифузору для граничного режиму, коли стрибок ущільнення розміщується на вході дифузору.

Висновки

Основними результатами CFD-моделювання процесів у зазначених струминних апаратах є масова витрата в елементах, температура і граничний тиск потоку на виході дифузору. Сформована двовимірна математична модель дозволяє розрахувати ці параметри у широкому діапазоні роботи апаратів та більш надійно проектувати такі пристрої. Її використання дозволить визначати характеристики та зону сталої роботи апаратів, підвищити ефективність та розширити діапазон їхнього використання.

ЛІТЕРАТУРА

1. Martinez I. Aircraft Environmental Control. 2013. Retrieved from <http://webserver.dmt.upm.es/~isidoro/tc3/Aircraft%20ECS.pdf> (Last accessed: 09.10.2022).
2. Петухов И.И. Научные основы рабочего процесса систем быстрого охлаждения и получения шугообразного состояния жидкости. Наукові праці: Науково-методичний журнал. Т. 111. Вип. 98. Техногенна безпека. Миколаїв: Вид-во ЧДУ ім. П. Могили, 2009. – С. 83-89.
3. Fan J., Eves J., Thompson H.M., Toropov V.V., Kapur N., Copley, D. and Mincher, A., Computational fluid dynamic analysis and design optimization of jet pumps, Computers & Fluids. 2011. No. 46. – pp. 212-217. DOI: 10.1016/j.compfluid.2010.10.024 (Last accessed: 09.10.2022).
4. Sobieski W. Jet pumps – numerical modelling possibilities upon the bifurcation phenomena, Chair of Mechanics and Machine Design University of Warmia and Mazury in Olsztyn. Techn. Sc., 2010. November, 13. – pp. 240-255. DOI 10.2478/v10022-010-0023-6 (Last accessed: 09.10.2022).
5. Help Ansys Fluent. Release 2021 R1.

ПЕТУХОВ Ілля Іванович – к. т. н., доцент, доцент кафедри аерокосмічної теплотехніки; Національний аерокосмічний університет ім. М.С. Жуковського "ХАІ", вул. Чкалова 17, м. Харків, Україна, 61070. E-mail: ilya2950@gmail.com. Номер ORCID: <http://orcid.org/0000-0002-0645-7912>

Наукові інтереси:

- термодинамічні процеси у багатофазних потоках.

ЛИСИЦЯ Олексій Юрійович – к. т. н., доцент; доцент кафедри аерокосмічної теплотехніки; Національний аерокосмічний університет ім. М.С. Жуковського "ХАІ", вул. Чкалова 17, м. Харків, Україна, 61070. E-mail: a.lisitca@khai.edu. Номер ORCID: <https://orcid.org/0000-0002-5679-8459>

Наукові інтереси:

- CFD-моделювання теплогідравлічних процесів в двигунах та енергетичних установках.

CURRENT ALGORITHMS FOR IDENTIFICATION USING BIOMETRIC DATA

This paper is devoted to the current topic of actual algorithms of identification based on biometric personal data (biometrics) in the system of electronic trust services. They are becoming one of the most used tools in everyday life. This could include things like fingerprint and facial scanners in modern cell phones, as well as more formal aspects such as a personal passport or signature.

The main problem with electronic identification is that the most effective algorithms: based on fuzzy extractors, where a fuzzy extractor refers to a system (object, algorithm) that transforms biometric data into random sequences, enabling encryption techniques for biometric security. And also algorithms based on abstract minutiae cylindrical codes where the minutiae are unique for each fingerprint, which determine points of change in the structure of capillary lines (termination, splitting, breaking, etc.), the orientation of capillary lines and coordinates at these points. Although they fulfill their role, they do not always work correctly, posing a possible danger to the user.

A comprehensive analysis of the advantages and disadvantages of such algorithms requires further development and use of a kind of symbiosis of these algorithms to solve existing problems and improve the overall response. This should not only provide a higher level of protection, but also greatly simplify the mathematical complexity of data processing algorithms. And it is this that will lead to an increase in the number of correct responses and increase the efficiency of the use of biometric technology in electronic trust services systems.

Keywords: *biometry, minutiae, imprecise extractor, algorithm, code.*

Problem Statement

Now more than ever, proper identification and authentication of a person is important. And what could be more personal than a person's individual biometric features, such as voice, retina, gait, individual fingerprint and, of course, DNA. All of these things come together under the term "biometrics." Such systems require more detailed study, all-consuming improvement and further development to avoid false positives and achieve a fully correct identification and authentication process, where an error can be fatal.

Review of potentially better identification algorithms and identification of areas for their further development to solve existing problems.

Current identification algorithms

Several major trends in the development of biometrics are trying to solve the actual problems of authentication - an algorithm based on abstract-minutia cylindrical codes, as well as an algorithm that uses fuzzy extractors. These are the ones most commonly used in biometrics to ensure a high probability of correct identification.

Abstract-Minutiae Cylindrical Codes. Abstract-Minutiae cylindrical codes are based on the fine details of fingerprint descriptors, which take into account the smallest information in the fingerprint image to match them. Minutiae are features unique to each fingerprint, defining points of capillary line structure change (termination, bifurcation, break, etc.), orientation of capillary lines and coordinates at these points. Each fingerprint can contain up to 70 or more miniatures. It is through their comparison that the local features are compared.

An algorithm based on cylindrical minutiae codes uses three-dimensional data structures called cylinders, where each cylinder is oriented in the direction of the central minutiae throughout the image. Minutiae are spatial points where orientation, frequency, and energy have a higher differential of change.

The orientation image computed by STFT analysis differs from traditional orientation images computed using simple derivatives. With the exception of the kernel and delta points, any local area of the fingerprint image provides consistent texture information using STFT. This is not the case with traditional gradient-based orientation estimation.

The disadvantage of this approach is the complexity of mathematical calculations of a large number of minutiae and a significant number of cylinders that are processed but discarded for one reason or another (considered invalid).

Fuzzy extractors. A fuzzy extractor is a system (object, algorithm) that converts biometric data into random sequences that provide the ability to apply encryption methods for biometric security. They are used to encrypt and authenticate user transactions. In this case, the biometric input is treated as a key. The word "fuzzy" in the extractor's name implies that the values of the resulting sequence have a form close enough to the original and can confirm the authenticity of the identity.

Algorithm using fuzzy extractors - a method that allows to uniquely recover the secret key from inaccurately reproduced biometric data involving auxiliary data, which is open. According to the algorithm, a sequence of actions is performed:

- initialization, where a security parameter is specified that defines the length of the public and secret keys and the triggering threshold, based on which the algorithm generates a secret master key and public parameters;
- extraction, where a specific identity and the secret master key are specified, and the algorithm itself processes the data and returns the user's secret key;
- encryption, where the algorithm returns a ciphertext based on the user's secret key, his identity, and the embedded message;
- decryption, where by the secret key and the ciphertext encrypted with the identity, the algorithm returns the message if the identity data is confirmed, or stops working otherwise.

Subsequently, the stability of such an algorithm is evaluated from the point of view of force attacks such as: full brute force, collision creation, etc., and, if stability to them is obtained, the algorithm is moved to the analysis of its stability against analytical attacks.

The disadvantage of this approach is the complexity of mathematical apparatus of data processing in the presence of a significant number of errors and the inability to provide the necessary accuracy when it is necessary to correct these errors.

Conclusions

Thus, the most important and promising areas for further research are those related to the incorrect operation of fuzzy extractors and abstract-minuative codes, which address the issues of correct identification and authentication of the person. The application of identification based on the synthesis of the "strengths" of both algorithms will level out their individual drawbacks.

REFERENCES

1. Podhaiko V.O., Rassomakhin S.G. *Analysis of identification algorithms in electronic trust services systems*. // Coll. of Scient. Papers of the First International Scientific and Technical Conference "Communication Systems and Technologies, Informatization and Cyber Security: Current Issues and Development Trends", Kyiv, November 25-26, 2021. – P. 239-240.
2. A. Menezes, P. van Oorschot and S. Vanstone, *Handbook of applied cryptography*. – Boca Raton, FL: CRC Press, 1996. – P. 306-312.
3. Yen-Lung Lai, Jung-Yeon Hwang, Zhe Jin, Soohyong Kim, Sangrae Cho and Andrew Beng Jin Teoh. *A Symmetric Keyring Encryption Scheme for Biometric Cryptosystems*. – arXiv:1807.02251v1 [cs.CV], 6 Jul 2018. – P. 1-15.
4. Wajih Ullah Baig, Umar Munir, Waqas Ellahi, Adeel Ejaz, Kashif Sardar. *Minutiae Texture Cylinder Codes for fingerprint matching*.// Information Sciences. V.502, October 2019. – P. 492-509.

ПОДГАЙКО Владислав Олегович – аспірант ХНУ ім. В.Н. Каразіна, м. Харків, майдан Свободи, 6, vladosla2@ukr.net ORCID: 0000-0003-1394-3588.

Наукові інтереси:

– ідентифікація та автентифікація у різноманітних системах за допомогою біометричних даних.

РАССОМАХИН Сергій Геннадійович – д.т.н., доцент, завідувач кафедри БІСТ, м. Харків, майдан Свободи, 6, bist@karazin.ua, ORCID: 0000-0003-1394-3588.

Наукові інтереси:

– ідентифікація та автентифікація у різноманітних системах за допомогою біометричних даних.

ПОМОГАЄВ К.О, ТОЛСТОЛУЗЬКА О.Г, АРТЮХ О.А.

КОМП'ЮТЕРНА МОДЕЛЬ 3D-ВІЗУАЛІЗАЦІЇ ПРИМІЩЕНЬ ІЗ ЗАСТОСУВАННЯМ МЕТОДІВ DATA MINING

The topic of the work is a computer model of 3d visualization of premises using data mining methods. The goal of the work is to reduce the time of developing a 3D visualization model while simultaneously expanding the options of 3D models. It is proposed to create an algorithm using data mining tools to develop a 3D visualization model of educational premises. The algorithm will allow you to automate the process of developing a 3D model and minimize the need for manual creation. In the course of this work, the concepts of modeling and 3D modeling are considered. The most important existing programs for 3D modeling were selected and analyzed. Based on this, the program for creating a 3D model of educational premises was selected. The main methods of data mining are considered and analyzed. Received photos of the premises and developed requirements for the 3D model. A modeling algorithm using Data mining was developed. Based on this algorithm, the process of developing a 3D model is described step by step. The model was tested and adjusted. The paper presents the results and images of the 3D model. As a result of the study, a modeling algorithm using Data mining tools was developed. The algorithm makes it possible to increase the efficiency of the development of a model of 3D visualization of educational premises. The 3D model has a convenient user interface, the ability to view from different angles and provides a complete idea of the modeling objects and their detailing.

Ключові слова: Data Mining, методи, технології, 3D-візуалізація, 3D-модель.

Аналіз методів Data Mining

Data mining (Аналіз даних) – це важливий аналітичний процес, призначений для дослідження даних. Подібно до реального процесу видобутку алмазів або золота з землі, найважливішим завданням Data mining є вилучення нетривіальних знань із великих обсягів даних. У роботі були розглянуті чотири методи Data mining:

- Регресія (прогностична)
- Аналіз правил асоціації (опис)
- Класифікація (прогностична)
- Кластеризація (опис)

Регресія — це найпряміша, найпростіша версія того, що називається «силою передбачення». При використанні регресивного аналізу можна передбачити значення певної (безперервної) ознаки на основі значень інших ознак у даних, припускаючи лінійну або нелінійну модель залежності.

Мета регресійного аналізу полягає у визначенні значень параметрів для функції, завдяки яким функція найкраще відповідає набору даних спостереження, які надаються. Наступне рівняння виражає ці відносини в символах. Це показує, що регресія є процес оцінки значення безперервної цілі (y) як функція (F) одного або кількох предикторів ($x_1, x_2, \dots, x_n$), набору параметрів ($\theta_1, \theta_2, \dots, \theta_n$) і міри помилки (e).

Аналіз правил асоціації - це важлива галузь досліджень аналізу даних, який в основному використовується для виявлення відповідних посилань між елементами в наборі даних. Апріорний алгоритм і Алгоритм FP-Tree є класичними алгоритмами в аналізі правил асоціації. Обидва базуються на тому, що кількість даних у базі даних транзакцій не зміниться, і кожен елемент даних має однакове значення. Метою аналізу правила асоціації є виявлення прихованого зв'язку між різними наборами елементів даних у базі даних. Загалом, враховуючи базу даних транзакцій, проблема аналізу правил асоціації – це процес пошуку правила асоціації через визначену користувачем мінімальну підтримку і мінімальну впевненість.

Класифікація — це ще одне важливе завдання, яке потрібно виконати перед тим, як занурюватися в жорстку фазу моделювання аналізу даних. Класифікація — це функція аналізу даних, яка призначає елементи в колекції цільовим категоріям або класам. Припустимо, що є набір записів: кожен запис містить набір атрибутів, де одним із атрибутів є певний клас. Мета полягає в тому, щоб знайти модель для класу, яка зможе передбачити невидимі або невідомі записи (із зовнішніх подібних джерел даних) точно так, якби мітка класу була помічена або відома, враховуючи всі значення інших атрибутів.

Кластеризація є важливим методом, метою якого є визначення груп об'єктів, щоб об'єкти в одному кластері були схожі один на одного, а об'єкти в різних групах – ні. Проблема кластеризації в цьому сенсі зводиться до наступного:

Кластеризація групує екземпляри даних у підмножини таким чином, що подібні екземпляри групуються разом, а різні екземпляри належать до різних груп. Таким чином, екземпляри організовуються в ефективне представлення, яке характеризує вибірккову сукупність.

Аналіз методів 3D-візуалізації

Кожен 3D-об'єкт, який відображається в 3D-анімації, складається з 3D-сітки, яку можна створити за допомогою різноманітних технік. Простіше кажучи, процес створення сітчастих об'єктів у 3D програмному забезпеченні називається 3D-моделюванням. У роботі були розглянуті популярні техніки 3D-моделювання для візуалізації продукту: полігональне моделювання, 3D-скульптинг, 3D-візуалізація одягу, сплайнове моделювання, процедурне моделювання та 3D-сканування.

Полігональне моделювання є найбільш використовуваним методом, особливо для відеоігор і анімаційних студій. За його допомогою можна створювати 3D-об'єкти з менших компонентів, які називаються «трикутниками» або «багатокутниками». З багатьох причин полігональне моделювання зробити набагато легше порівняно з іншими. Воно робиться шляхом створення контуру об'єкта, заповнення простору між краями багатокутників та остаточного застосування текстури до моделі.

3D-скульптинг, також відомий як цифровий скульптинг, дозволяє художнику маніпулювати об'єктами з тривимірних форм на комп'ютері, ніби вони зроблені з оцифрованої глини. 3D-Скульптинг вимагає багаторівневого процесу, у якому цифровий скульптор починає використовувати інструменти для маніпулювання геометрією скульптури. Сформувавши основну форму та силует, художник додаватиме дедалі більше деталей у процесі поділу шарів, поки 3D-скульптура не стане більш реалістичною.

3D-візуалізація одягу дозволяє художникам створювати багато дизайнів відповідно до специфікацій модельєра. Кілька додатків, розроблених для цієї техніки, дозволяють художнику визначити, чи добре виглядатиме дизайн для всіх розмірів. Для багатьох професіоналів індустрії моди ця технологія корисна для кращого бачення дизайну одягу.

Сплайнове моделювання, також відоме як моделювання NURBS, створює 3D-об'єкти на основі кривих, схожих на кілки для встановлення намету, водночас створюючи поверхні навколо них. Сплайнове моделювання більш точне до деталей і ідеальне для фільмів, архітектури та дизайну продуктів.

Процурне моделювання — це загальний термін, який описує кілька передових методів створення моделей із наборів правил. Ці правила інтегровані в алгоритм, параметри якого можна налаштувати.

3D-Сканування — це процес створення високоточних 3D-моделей об'єктів реального світу. На відміну від створення 3D-моделей з нуля, 3D-сканування фіксує точний вигляд фізичного об'єкта за допомогою 3D-лазерного сканування та перетворює його на 3D-об'єкт. 3D-сканер може фіксувати не лише форму, але й текстуру та колір об'єкта, що спрощує отримання точної інформації.

Аналіз технологій для 3D-візуалізації

Найкраще програмне забезпечення для 3D-моделювання розкриває творчі здібності, допомагаючи втілити в життя привабливі та творчі проекти. Можна виокремити 4 найпопулярніших середовища для 3D-моделювання: Blender, Maya, 3D Max, Cinema 4D.

Blender є безкоштовним і доступним інструментом, щоб розпочати 3D-моделювання, цифрове ліплення або загалом вивчати світ 3D. Як пакет для 3D-моделювання Blender надає інструменти, необхідні для створення 3D-моделей професійного рівня. Однак у нього є сильні та слабкі сторони на різних етапах конвеєра — він хороший для ліплення, але має проблеми з наддрібними деталями.

Перевагою Maya є її професійний та обширний набір інструментів для анімації. Однак вона не рекомендується як перший пакет програмного забезпечення для 3D для більшості людей, оскільки він не найпростіший у вивченні, не найдешевший і навіть не найпотужніший. Однак

якщо метою є влаштуватися на роботу 3D-аніматора, технічного директора з монтажу, технічного художника для ігор або будь-яку іншу роль у сфері цифрових розваг, знання Maya є безцінним навиком.

Перевагою Cinema 4D є те, що вона дозволяє швидко створювати красиві та привабливі візуальні ефекти, зокрема анімаційну графіку, і полегшує ітерацію. У цьому ключі все покращується. Бути в змозі матеріалізувати щось із уяви легше в Cinema 4D, ніж у більшості порівнянного програмного забезпечення.

3ds Max — це практичне програмне забезпечення для вивчення та відпрацювання інструментів для 3D-моделювання. Саме високотехнічні інструменти для моделювання та їх велика кількість є головною перевагою програми. Однією з найкращих функцій є функціональність плагіна, яка покращує його функції та робить його глобально використовуваним додатком.

Процес розробки комп'ютерної моделі

Тож метою роботи є зменшення часу розробки моделі 3D-візуалізації з одночасним розширенням варіантів 3D-моделей. Для реалізації цього завдання була обрана програма Blender. Ідея реалізації проєкту полягає в тому, що створюється 3D-модель будівлі, в якій відсутні певні її елементи. Після цього пишеться алгоритм засобами мови програмування Python та методу класифікації, за допомогою якого можна швидко створити невідсутні елементи. При чому, таких елементів буде створено декілька типів, і цьому посприятиме реалізація методу класифікації. В результаті користувач зможе за кілька секунд створювати необхідні елементи в будівлі та швидко переглядати їх варіанти. Розглянемо реалізацію 3D-моделі в нотатції IDEF0.

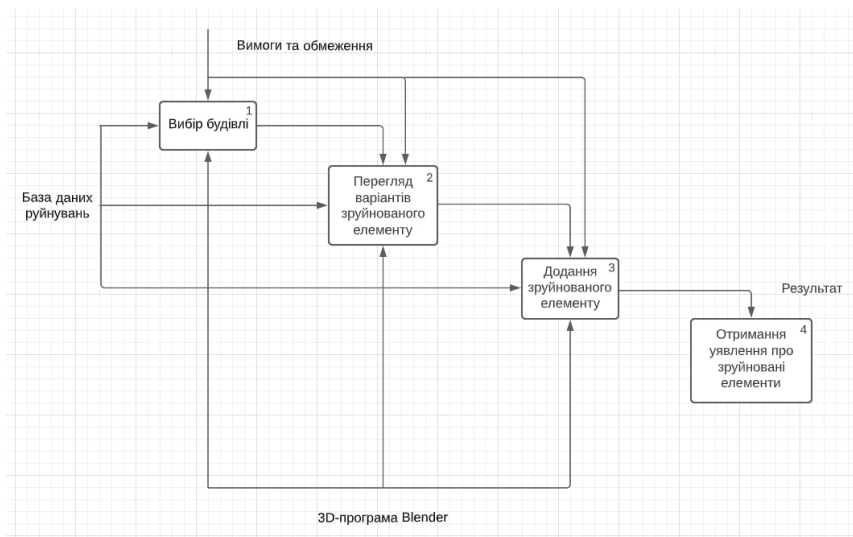


Рис. 1 3D-модель в нотатції IDEF0

Вхідними даними розроблюваної моделі є база даних руйнувань. Тобто є певна кількість будівель, в яких зруйновані деякі їх елементи, наприклад вікна або двері. Ці елементи в результаті мають бути додані. Знизу схеми можна побачити інструментарію, який виконує дії, вказані в блоках. Це обрана програма для 3D-моделювання, а саме Blender. Вся розробка і дії відбуваються у цій програмі. Зверху схеми можна побачити вимоги та обмеження.

Згідно поставленої мети 3D-модель має наступні вимоги та обмеження:

1. У 3D-будівлі повинні бути відсутні деякі елементи.
2. 3D-будівля повинна мати відповідну та адекватну кількість використаних інструментів та полігонів.
3. 3D-модель повинна мати від двох та більше варіантів зруйнованих та невідсутніх елементів для кожної будівлі.

Повертаючись до схеми IDEF0, першим блоком є вибір 3D-будівлі. Користувач обирає необхідну йому 3D-будівлю із бази даних руйнувань. Також є можливість додати нову 3D-будівлю у базу даних, якщо користувач не знайде потрібного варіанта. Другим блоком є перегляд

варіантів зруйнованого 3D-елементу. Тобто користувач переглядає варіанти 3D-елементу та обирає той, який більше підходить. У третьому блоці користувач додає обраний варіант зруйнованого 3D-елементу, натискаючи на відповідну кнопку. В результаті користувач отримує уявлення про зруйновані 3D-елементи будівлі та про те, як ця будівля виглядає з доданим 3D-елементом.

У роботі був розглянут процес реалізації 3D-моделі на прикладі першої будівлі. Першим етапом було створення 3D-моделі будівлі. В цій моделі відсутні вікна, як показано на рисунку 2.

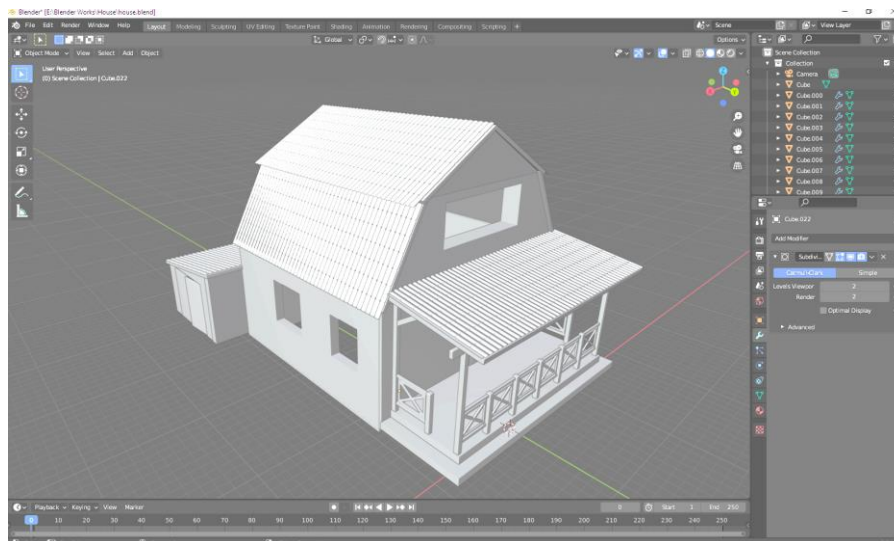


Рис. 2 3D-будівля без вікон

Оскільки в цій будівлі відсутні вікна, саме їх потрібно було створити. Однак, способом їх розробки було не звичайне 3D-моделювання, а створення алгоритму на мові Python. Для створення 3D-вікон, були розроблені спеціальні панелі в програмі Blender, простіше кажучи, кнопки. Далі були розроблені відповідні класи та функції для реалізації алгоритму. Після цього був створений новий елемент addon. Туди був поміщений код програми та створені відповідні класи. Взагалі аддон – це будь-яка стороння програма або скрипт, який додається до програми, щоб надати їй додаткові функції та можливості. Далі цей addon був збережений як python файл у папці на жорсткому диску. Тепер його можна розповсюджувати, щоб інші користувачі змогли його завантажити та використати. Далі цей addon був збережений як python файл у папці на жорсткому диску. Тепер його можна розповсюджувати, щоб інші користувачі змогли його завантажити та використати. В результаті, можна побачити, розроблені типи вікон, які з'являються після натискання відповідних кнопок. На рисунку 2 можна побачити, що були створені вікна після натискання кнопки Add Window 3.

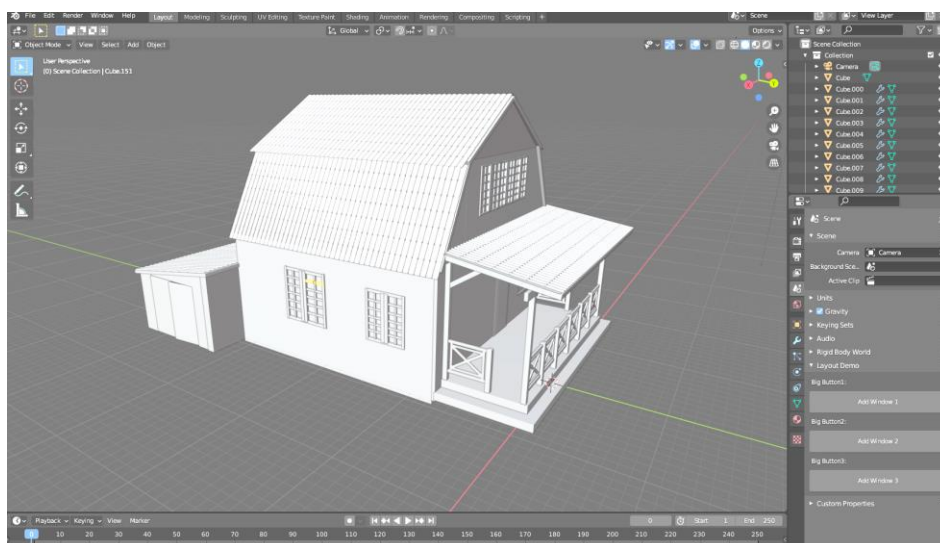


Рис. 3 Створення вікон

В результаті, була розроблена комп'ютерна модель 3D-візуалізації із застосуванням методу класифікації. Модель дозволяє дуже швидко, за кілька секунд, створювати 3D-елементи, які відсутні у зруйнованих будівлях. Цей процес відбувається в рази швидше ніж створення цих елементів вручну. Поставлена мета про зменшення часу розробки моделі 3D-візуалізації з одночасним розширенням варіантів 3D-моделей була виконана. Щоб скористатися розробленим аддоном користувачу потрібно завантажити програму Blender, а також завантажити створений аддон за наступним посиланням: https://drive.google.com/drive/folders/19nEfwXA0-mPCGKUvFAYRlvVe0QoMEIMh?usp=share_link і активувати аддон у програмі. Користувач має інтерактивний доступ до моделі, може швидко переглядати варіанти відсутніх 3D-елементів та обирати необхідний для себе варіант. В майбутньому планується розробити та реалізувати зв'язок бази даних руйнувань із розробленими 3D-моделями.

ЛІТЕРАТУРА

1. King Roger. 3D Animation for the Raw Beginner Using Autodesk Maya. 2nd Edition. — Taylor & Francis Group, LLC, 2019. — 450 p. — ISBN 9780815388791.
2. Danan Thilakanathan. 3D Modeling For Beginners: Learn everything you need to know about 3D Modeling! 1st Edition - ISBN-13: 978-1530799626.
3. John M. Blain. The Complete Guide to Blender Graphics: A Beginner's Guide. 1st Edition - ISBN-13: 978-1466517035
4. Pradeep Mamgain. MAXON Cinema 4D R20: Modeling Essentials. Kindle Edition
5. Brian L. Smith. 3ds Max Design Architectural Visualization: For Intermediate Users. 1st Edition - ISBN-13: 978-0240821078
6. <https://docs.blender.org/manual/en/latest/> (Last accessed: 10.11.2022).
7. Jiawei Han. Data Mining Concepts and Techniques. Third Edition — 2012. <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf> (Last accessed: 11.11.2022).

ПОМОГАЄВ Костянтин Олександрович – студент факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: xa11867765@student.karazin.ua; ORCID: 0000-0002-5010-2369.

Наукові інтереси:

– *Розробка 3D-моделей місцевості та предметів, анімація, створення художніх фільмів, web-design.*

ТОЛСТОЛУЗЬКА Олена Геннадіївна – д. т. н., с.н.с.; професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: elena.tolstoluzka@karazin.ua; ORCID: 0000-0003-1241-7906.

Наукові інтереси:

– *Науково-методологічні основи технології розробки мультипаралельного програмного забезпечення інформаційних та управляючих систем.*

АРТЮХ Олексій Анатолійович – ст. викладач кафедри ТПС факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: oleksiy.artiuh@student.karazin.ua; ORCID: 0000-0002-2038-4131.

Наукові інтереси:

– *Мережеві технології, моделювання інформаційних процесів в мережах.*

ПОПОВА А.О., ТОЛСТОЛУЗЬКА О.Г.

АНАЛІЗ ПІДХОДІВ ДО ОБРОБКИ ДАНИХ, ЩО ПРИХОДЯТЬ ВІД РЕЧЕЙ. СИСТЕМИ ПОБУДОВИ ІНТЕРНЕТУ РЕЧЕЙ

Theme of work. Approaches to building systems based on the Internet of things. Methods of data transfer from things to processing platform. **Purpose of work.** Introduction and study of systems construction using Internet technologies of things, architectures of such systems.

Methodsof research. Analysis, parse, explanation. **Results.** The result of the work is an overview of the architecture of the systems built with the help of the Internet things, analysis of the platforms of data processing with things, analysis of methods of data transfer from things to platforms. **Conclusions.** In the work were considered the concept of the Internet of things, the architecture of systems built by the Internet of things, their analysis. System data is often used to enable physical objects, such as sensors or devices, to communicate with each other or to send them to a platform for further processing. Platforms are of different types and are used depending on the user's needs. Sensors can communicate via Wi-Fi or GSM. In the case of the latter, IoT-type SIM cards are used, also called M2M SIM Card (machine-to-machine).

This approach to automation of data collection from different sensors via the Internet of things helps to accelerate the process of data exchange and allows devices to interact with each other without human intervention. It covers a wide range of tasks in different fields because the system is set up once and then works independently.

This architecture is needed to reduce human resource requirements and automate the interaction between devices and the data processing system. This architecture can scale with more devices without processing the processing platform itself and without causing significant costs.

Ключові слова: Інтернет речей, модель, датчики, система.

Вступ

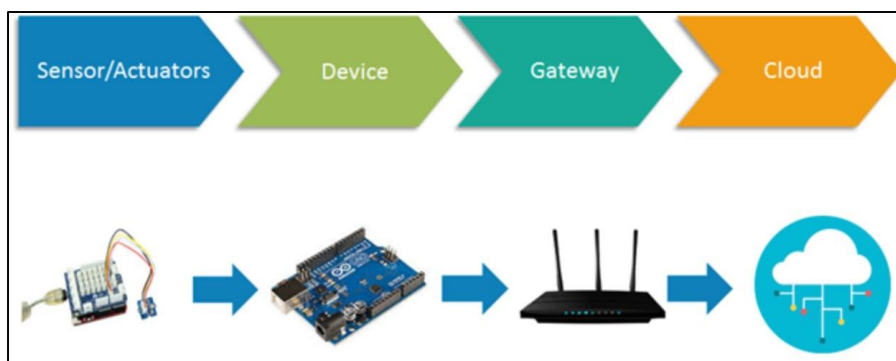
В нашому житті важливе місце зараз займають системи відстеження. Можна відстежувати місцезнаходження своїх навушників, своєї собаки, своєї машини або громадського транспорту. За допомогою спеціальних пристроїв можна слідкувати за відсотком вологості чи температурою у теплиці. Існує багато прикладів застосування обробки даних, що приходять від речей які відправляють зібрані дані до системи обробки чи до інших речей з метою подальшого їх використання. Для побудови таких систем досить часто застосовується технологія Internet of Things. Завдяки їй та фізичним об'єктам, підключеним до мережі Інтернет для обміну даними існує можливість побудови складних систем. Концепція IoT може значно поліпшити багато сфер нашого життя і допомогти створити більш комфортний, розумний і безпечний світ. Приклади IoT варіюються від речей, таких як розумні годинники до розумних будинків, які можуть, наприклад, контролювати і автоматично змінювати ступінь освітлення і опалення.

Архітектура системи інтернету речей

Інтернет речей – це поняття з'єднання будь-якого пристрою з мережею Інтернет та з іншими підключеними пристроями. Це гігантська мережа зв'язаних речей і людей що збирають і діляться даними. З'єднання різних об'єктів і додавання до них датчиків додає рівень цифрового інтелекту до пристроїв, дозволяючи їм спілкуватися в реальному часі без залучення людини. Інтернет речей робить навколишній світ розумнішим, об'єднуючи цифрові та фізичні всесвіти [1, 2].

Можна розділити стек технології IoT на чотири технологічних рівня: – кінцеві пристрої;

- програмне забезпечення;
- комунікації;
- платформа аналізу.



1 Модель архітектури IoT системи

Пристрої – це об'єкти, які фактично утворюють «речі» в Інтернеті речей. Вони відіграють роль інтерфейсу між реальним і цифровим світами передаючи зібрані дані до обробки. Сенсори або пристрої збирають дані зі свого середовища. Ці дані можуть бути такими ж простими, як і читання температури або як складні, як повна подача відео. Часто використовуються сенсори-пристрої, тому що кілька сенсорів можуть бути об'єднані разом або датчики можуть бути частиною пристрою, який робить більше, ніж просто відчуття. Наприклад, телефон є пристроєм, який має кілька датчиків (камера, акселерометр, GPS і т.д.), але телефон не просто датчик, оскільки він також може виконувати багато дій [3, 4].

Компонентом завдяки якому підключені пристрої можна назвати «розумними» є програмне забезпечення. Воно відповідає за хмарний зв'язок, збір даних, інтеграцію пристроїв між іншими пристроями і аналіз даних в реальному часі. Він також забезпечує взаємодію з системою IoT.

Датчики або пристрої можна підключати до хмари різними методами, включаючи: стільниковий зв'язок, супутниковий зв'язок, Wi-Fi, Bluetooth, глобальні мережі малого споживання (LPWAN), підключення через шлюз/маршрутизатор або пряме підключення до Інтернету через Ethernet. (не хвилюйтеся, ми пояснимо, що все це означає, у нашому розділі підключення). Кожна опція має компроміс між енергоспоживанням, діапазоном і пропускну здатністю. Вибір найкращого варіанта підключення залежить від конкретної програми IoT, але всі вони виконують одне й те саме завдання: передають дані в хмару. Рівень комунікацій (зв'язку) включає в себе як рішення фізичного зв'язку (стільниковий і супутниковий, LAN), так і спеціальні протоколи, що використовуються в різних IoT середовищах (ZigBee, Thread, Z-Wave, MQTT, LwM2M). Для фізичного зв'язку зазвичай використовується SIM-карта IoT, також вона називається M2M SIM Card (machine-to-machine). Вона є модулем ідентифікації абонента, який використовується в IoT пристроях для ідентифікації так як вони намагаються підключитися до 2G, 3G, 4G-LTE, Cat-M, NB-IoT [4, 5].

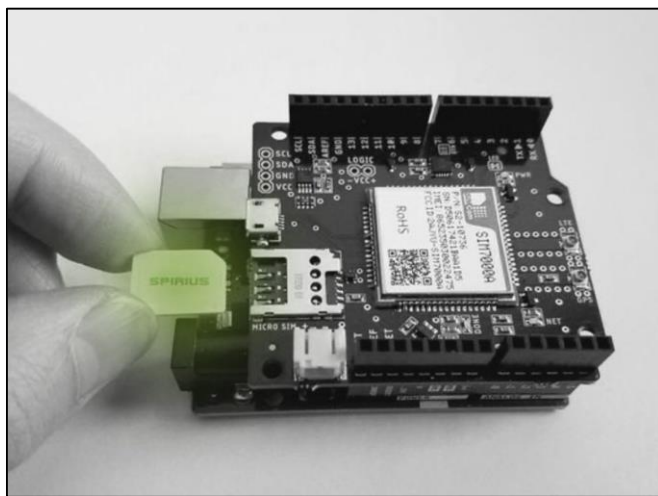


Рис. 2 Приклад встановлення IoT SIM-карти у модуль

Платформа IoT – місце, де всі ці дані збираються, аналізуються і передаються користувачеві зручним способом. Пристрої здатні «відчувати» те, що відбувається навколо них і повідомляти про це користувачеві через обраний канал зв'язку. Платформи можуть встановлюватися локально або в хмарі.

Аналіз підходів до обробки даних, що приходять від речей

Аналіз даних що приходять від речей відбувається у платформах. Дані можна обробляти як на платформі локально так і відправляти їх до хмарних сервісів.

Існує п'ять типів платформ IoT:

- платформи керування підключеннями; – платформи керування пристроями;
- хмарні платформи;
- платформи підтримки програм;
- розширені аналітичні платформи.

Платформи керування підключенням IoT зосереджені навколо мережевого компонента систем IoT. Вони надають можливість користувачам використовувати готове програмне забезпечення, обладнання для підключення пристроїв та керування даними, необхідні для підтримки пристроїв онлайн. Їхні мережі зазвичай покладаються на існуючі послуги оператора та Wi-Fi.

Платформи керування пристроями IoT спеціалізуються на важкій роботі з пристроями IoT. Вони гарантують, що все підключено, а також постійно інформують про стан пристроїв. Також вони оновлюють мікропрограму всередині пристрою, сповіщають про зміни та звітують про показники. Така платформа автоматизує виконання рутинних завдань, пов'язаних з пристроями, незалежно від їх кількості.

Хмарні платформи IoT надають інфраструктуру, необхідну для створення цілісної системи IoT. Вони є центральним місцем для існування та роботи всіх внутрішніх процесів і даних. Однією з найбільших переваг хмарних платформ є їхня масштабованість; незалежно від того, з якого розміру системи починається робота; хмарна платформа може розвиватися разом з системою IoT.

Платформи підтримки програм IoT – це універсальний підхід, що пропонує користувачам усе необхідне для запуску системи Інтернету речей. Вони надають пристрої, програмне забезпечення, розробку та розгортання систем IoT. Це універсальний магазин для швидкого запуску системи, позбавляючи необхідності самостійно керувати розробниками, конфігурацією мережі та апаратним забезпеченням.

Розширені аналітичні платформи IoT є чудовим рішенням для систем IoT, керованих даними. Користувачі, які шукають складні системи Інтернету речей, які використовують машинне навчання, штучний інтелект, статистичне моделювання та масовий збір даних, можуть використовувати цю платформу для інтерпретації та дій на основі зібраних даних. Системи Інтернету речей, які в основному працюють над отриманням даних, а не виконанням завдань, отримають найбільшу користь від цих платформ [6, 7].

Висновок

У роботі були розглянуті поняття Інтернету речей, архітектури систем побудованих за допомогою Інтернету речей, проведений їх аналіз. Дані системи часто застосовуються для того щоб фізичним об'єктам, таким як датчики чи пристрої обмінюватись даними між собою чи відправляти їх до платформи для подальшої обробки. Платформи бувають різних типів і використовуються в залежності від потреб користувача. Датчики мають змогу обмінюватись даними за допомогою Wi-Fi мережі або за допомогою GSM. В випадку з останнім використовуються SIM-карти типу IoT, також вона називається M2M SIM Card (machine-to-machine). Даний підхід до автоматизації збору даних з різних датчиків за допомогою Інтернету речей допомагає прискорити процес обміну даними та надає змогу пристроям взаємодіяти між собою без втручання людини. Він покриває великий спектр задач у різних галузях через те, що система налаштовується один раз і далі працює самостійно. Ця архітектура потрібна для того, щоб зменшити потреби у людських ресурсах і автоматизувати взаємодію між пристроями та

системою обробки даних. Така архітектура може масштабуватись із збільшенням кількості пристроїв без переробки самої платформи обробки і не спричиняти великих витрат

ЛІТЕРАТУРА

1. What is Internet of Things (IoT): [Електронний ресурс], URL: <https://www.ibm.com/blogs/internet-of-things/what-is-the-iot/>
2. What is the IoT? Everything you need to know about the Internet of Things right now: [Електронний ресурс], URL: <https://www.zdnet.com/article/what-is-the-internet-of-things-everything-you-need-to-know-about-the-iot-right-now/>
3. What is Internet of Things: [Електронний ресурс], URL: <https://www.makeuseof.com/tag/what-is-internet-of-things>
4. Что такое IoT и что о нем следует знать: [Електронний ресурс], URL: <https://habr.com/ru/company/otus/blog/549550/>
5. What is an IoT SIM Card and How is it Used in IoT devices: [Електронний ресурс], URL: <https://www.twilio.com/blog/what-is-an-iot-sim-card>
6. The 5 Types of IoT Platforms: [Електронний ресурс], URL: <https://blog.velosiot.com/the-5-types-of-iot-platforms>
7. How an IoT System Actually Works: [Електронний ресурс], URL: <https://www.leverage.com/iot-ebook/how-iot-systems-work>

ПОПОВА Анастасія Олександрівна – студентка кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: xa11867792@student.karazin.ua; ORCID: 0000-0003-1145-5695.

Наукові інтереси:

- моделювання складних комп'ютерних систем,
- технології розподіленого і паралельного програмування

ТОЛСТОЛУЗЬКА Олена Геннадіївна – д. т. н., с. н. с.; професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: elena.tolstoluzka@karazin.ua; ORCID: 0000-0003-1241-7906.

Наукові інтереси:

- технології розподіленого і паралельного програмування.

RADOUTSKYI K. E., RADOUTSKA A. K.

ADVANTAGES AND DISADVANTAGES OF AGILE PROJECT MANAGEMENT METHODOLOGY

The Agile project management methodology breaks a project down into several phases and uses these stages to guide the development. Finally, this condition of continuous development mirrors the design backwards and forwards process. The only difference is that with Agile, we remain in this state throughout the course of the program, but with design, we do so only during the project's life. Instead of focusing on creating a flawless product, the methodology emphasizes continual product improvement. The purpose of this work was to determine whether the popular methodology is effective, improves the speed of software development and delivery. In this work, the methodology was applied in practice, due to which positive and negative aspects were identified. Since there is no clear formulation of the principles for implementing Agile and the gradation of determining how successfully it is used by the team, it is not possible to say for sure how much the team applied the principles, including the fact that the result depends on how each member understood the basis of Agile principles. As a result, we can say that this methodology should be individually adjusted for each team, including an emphasis on understanding goals and motivation. Goal emphasis changing, a lack of documentation, and difficult information transmission are some of Agile's drawbacks, but it also speeds up development, makes it more transparent, encourages performance enhancement, and makes cooperation simple.

Keywords: project management, software development process, agile methodology, software delivery process, managing software development, estimating software delivery.

The Agile methodology is used for project management and divides a project into many parts and uses these phases to control the development. Continuous improvement at every stage and ongoing collaboration with stakeholders are really important here. Teams cycle through a process of planning, carrying out, and assessing once the job starts. In 2001 the group of individuals created manifest for this methodology, which became the base for Agile. Idea is embraced by the Agile Manifesto, as is the key statements that software should be built in a continuous loop where client needs are inputted and "good enough" software is outputted [1,2].

It refers to next principles:

1. Individuals and interactions over processes and tools
2. Working software over comprehensive documentation
3. Customer collaboration over contract negotiation
4. Responding to change over following a plan

This state of ongoing improvement finally resembles the backwards and forwards process of design. The only distinction is that in Agile, we remain in this state for the life of the program, whereas with design, we remain in this condition for the duration of the project.

Agile implementation allows a department to deliver products to clients as quickly as possible. Because methodology requires a lot of feedback between customers and employees, the approach introduces a great system of cooperation between customers and other departments. In a flexible system, management encourages employees to think creatively to solve problems and create simple solutions to challenges in project development. In addition, Agile departments can better recognize and address issues rapidly due to the fragmented structure of production and correction. Departments may swiftly address problems before the next production cycle by incorporating input from customers and other team members. The period between comprehending an improvement and putting it into practice may be shortened since management encourages people to act autonomously using this methodology. It assist a team in producing product without interrupting production for significant fixes or process changes as an Agile department concentrates on production and maintain previous version all the time, even if it has issues, that will be solved in next deployment. Additionally, the Agile departmental values start a production cycle with less preparation work needed before they can start selling product, reducing time between the planning and profit stages [3].

But besides Agile advantages there are some unresolved issues in implementing Agile as main iterative approach to project management and software development. Firstly, while switching from one approach of management to another across departments, it can take time to fully adapt to the new responsibilities and style of project progression within the system. Teamleads may struggle to adapt to

leadership styles at first, but consistent practice and training can help all employees to new tactics. Agile management encourages employees to work independently toward goals, but consider dividing departments into groups for the first few goals of a project. In this way, each member can consult each other about questions and communication difficulties.

Another problem is that because the Agile methodology focuses on multiple goals at a time, certain goals may not receive as much focus throughout the entire project. Variable goals within a department can lead to a lack of specific goals for a team, which may lead to unknown deadlines and less accurate costs. It may be challenging to direct all members of a team toward a single objective due to the fact that Agile methodology requires focus shifting in accordance with which aspect of a project requires the most attention. It may become preferable for the majority of department members to concentrate on a single objective in later stages of the project.

In addition, the department may not be able to predict profits prior to production beginning because Agile production relies on continuous improvement and customer feedback. Because Agile production prioritizes delivering acceptable products to customers as quickly as possible, potential issues may become more difficult to anticipate over time.

In conclusion, the Agile focuses on producing a continuously improving product rather than a perfect one. Similarly to other different methodologies, it has some pros and cons, which should be considered. Agile disadvantages are goal focus shifting, lack of documentation and complicated information transfer, but at the same time, Agile makes development faster, easier and more transparent, increase performance improvement and ease of collaboration.

LITERATURE

1. Diana Larsen, Esther Derby, and Ken Schwaber, Agile Retrospectives: Making Good Teams Great, Pragmatic Bookshelf; 1st edition, 2006. 178 p.
2. Jim Highsmith, Agile project management: creating innovative products, Addison-Wesley Professional, 2009, 432 p.
3. Mark K. Leighton and Steven J. Ostermiller, Agile Project Management for Dummies, For Dummies; 3rd edition, 496 p.

РАДОУЦЬКИЙ Костянтин Євгенович – senior lecturer of the Department of Systems and Technologies Modeling; Kharkiv National University named after V.N. Karazina, Maidan Svobody, 4, Kharkiv-22, Ukraine, 61022; e-mail: radoutskyi@karazin.ua; ORCID: 0000-0003-3610-535X.

Scientific interests:

– *Development of methods and algorithms for compact storage of information.*

РАДОУЦЬКА Анна Костянтинівна – student; student of the Department of Artificial Intelligence, Kharkiv National University of Radioelectronics, Kharkiv, Prospect Nauki, 14b, 61166; e-mail: anna.radoutska@nure.ua.

Scientific interests:

– *Development of mobile applications for iOS*

РАДОУЦЬКИЙ Костянтин Євгенович – старший викладач кафедри моделювання систем і технологій; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: radoutskyi@karazin.ua; ORCID: 0000-0003-3610-535X.

Наукові інтереси:

– *Розробка методів та алгоритмів компактного зберігання інформації.*

РАДОУЦЬКА Анна Костянтинівна – студентка кафедри штучного інтелекту, Харківський Національний Університет Радіоелектроніки, м. Харків, Проспект Науки, 14б, 61166; e-mail: anna.radoutska@nure.ua.

Наукові інтереси:

– *Розроблення мобільних додатків для iOS*

СКОБ Ю.О.

МОДЕЛЮВАННЯ НАСЛІДКІВ ВПЛИВУ ТЕПЛОВОГО ВИПРОМІНЮВАННЯ НА ДОВКІЛЛЯ

Актуальність теми роботи визначається потребами вирішення засобами математичного моделювання актуальної науково-прикладної проблеми зниження ризиків порушення безпечного функціонування техногенних об'єктів на основі аналізу руху газової суміші у зоні викидів газових домішок і прогнозу стану повітря техногенних об'єктів в умовах формування температурного збурення повітря на довкілля і пошуку раціональних конструкцій захисних засобів для усунення або пом'якшення негативних наслідків цього впливу. Метою роботи є системне вдосконалення та підвищення ефективності математичних моделей, методології та програмного засобу комп'ютерної реалізації системної моделі підтримки прийняття рішень для аналізу руху багатокомпонентної газової суміші у розташуванні техногенних об'єктів під впливом теплового випромінювання. Теоретичною основою роботи є методи математичного моделювання, обчислювальні методи математичної фізики, пробіт-аналізу, теорії оцінки ризиків і прийняття рішень. Для побудови програмного засобу комп'ютерної реалізації системної моделі використовувалися принципи теорії комп'ютерних систем моделювання. За результатами моделювання отримано розподіли умовної ймовірності летального ураження людини внаслідок впливу теплового випромінювання від високотемпературного газового струменя у заданих контрольних точках для двох варіантів розрахункової схеми: без захисту робочого місця персоналу та за присутності непроникної для випромінювання суцільної захисної перепони. Виявлений характер розподілу ймовірності летального ураження людини дає змогу стверджувати про необхідність урахування впливу непроникних для теплових потоків перепон під час оцінювання безпекового стану на техногенному об'єкті, де має місце викид в атмосферу високотемпературних газів.

Ключові слова: диференціальні рівняння змішаного типу з частинними похідними, методи наскрізного розрахунку, газові суміші, пробіт-аналіз, ймовірність ураження, пробіт-функція.

Актуальність дослідження

Часто техногенні аварії супроводжуються горінням газоповітряної суміші з утворенням високотемпературних продуктів хімічної реакції, які є джерелом теплового випромінювання у навколишній простір. Висока температура характеризує збурення нормального стану повітря. Експозиція людини тепловому потоку певної щільності утворює дозу теплового випромінювання і може призводити до тяжких наслідків. Тому визначення ризику для такого роду аварії підприємства є важливим та актуальним інженерно-практичним завданням.

Математична модель впливу теплового випромінювання

Оцінювання наслідків аварії такого типу включає визначення ймовірності ураження персоналу тепловою радіацією на основі моделювання руху продуктів згоряння газу в атмосфері. Отримані поля температур дозволяють визначити щільності теплового потоку від джерела теплової радіації, теплову дозу та умовну ймовірність ураження людини.

Аналіз та прогноз наслідків впливу техногенної аварії на персонал засновані на математичному моделюванні руху газової суміші у приземному шарі атмосфери [1].

Максимальну щільність теплового потоку поглиненого випромінювання $q_{\max}$ можна визначити таким чином

$$q_{\max} = C_0 \left[\left(\frac{T_1}{100} \right)^4 - \left(\frac{T_2}{100} \right)^4 \right] \frac{\Delta F_1}{\pi r^2}, \quad (1)$$

де C_0 – коефіцієнт випромінювання полум'я, кВт/м²/К⁴;

T_1 – середня температура випромінюючої поверхні, К;

T_2 – середня температура поверхні приймача випромінювання, К;

r – відстань від поверхні випромінювання до приймача, м;

ΔF_1 – площа поверхні випромінювання, м².

Ймовірність ураження P залежить від пробіт-функції Pr – верхньої межі певного інтегралу нормального закону розподілу з математичним очікуванням 5 та дисперсією 1

$$P = \frac{1}{\sqrt{2\pi}} \int_0^{Pr} e^{-\frac{1}{2}(\tau-5)^2} d\tau, \quad (2)$$

де τ – час експозиції тепловому потоку, с.

Пробіт-функція для летального наслідку для персоналу техногенного об'єкту під впливом теплового випромінювання у загальному випадку визначається за формулою

$$P_r = -14 + 2,56 \ln(q^{1,33} \tau), \quad (3)$$

де q – поточна щільність теплового потоку, кВт/м².

Під час прийняття рішення про безпековий стан на техногенному об'єкті підвищеної небезпеки, яким є підприємство, де використовується займистий газ, експерти з безпеки зазвичай для отримання чисельної оцінки умовної ймовірності ураження (2) персоналу, що перебуває під впливом вражаючого фактора, для конкретного значення пробіт-функції за результатами математичного моделювання газового потоку можуть скористатися автоматизованою технологією пробіт-аналізу [2].

Апробація математичної моделі

У роботі розглядається струменеве витікання високотемпературних (3450 К) продуктів згоряння водню із отвору, який виник внаслідок часткового руйнування обладнання зберігання (рис. 1). Пляма отвору витікання прийняла форму кола радіусом $R = 0,5$ м з координатами центру $X_c = 7,5$ м, $Z_c = 2,5$ м у розрахунковому просторі з габаритами $15,0 \times 9,0 \times 15,0$ м у декартовій системі координат XYZ і варіантом за кількістю комірок уздовж координатних осей $15 \times 9 \times 15$. На відстані $Z_w = 1$ м від центру отвору витікання встановлювалася захисна суцільна, непроникна для теплових і газових потоків перепона кубічної форми з габаритами $1,0 \times 1,0 \times 1,0$ м. Швидкість повітря дорівнювала 0 м/с. Визначення стану параметрів повітря і безпеки здійснювалось у контрольних точках P_i на дистанції 5, 6, 7, 8 та 9 м від епіцентру витіку.

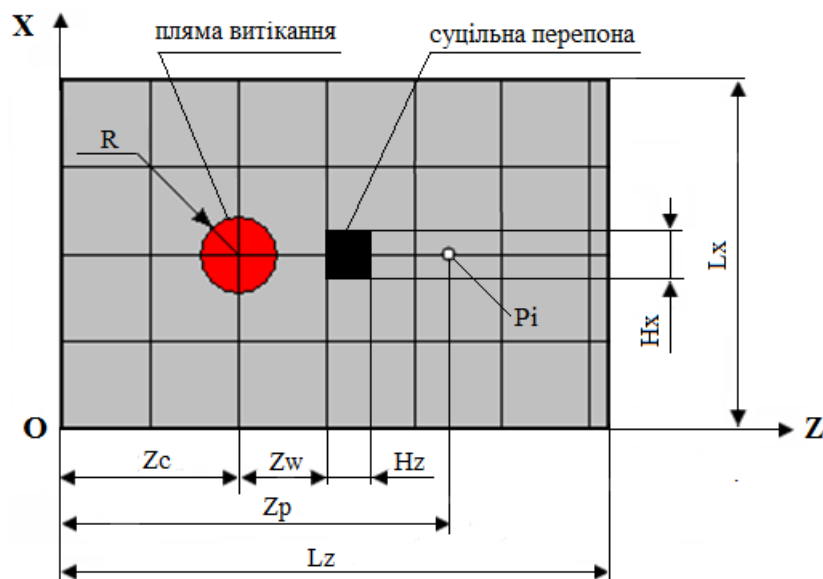


Рис. 1 Мапа розташування об'єктів на майданчику

Аналіз отриманих результатів

Високотемпературні продукти згоряння водню із заданою інтенсивністю викидаються з отвору в атмосферу, змішуються з повітрям і стають джерелом теплового потоку, який впливає довкілля. Вважається, що викид гарячого газу починається з моменту часу $\tau_1 = 0$ с і примусово припиняється після $\tau_2 = 5$ с, і хмара, втрачаючи максимальну концентрацію через розсіювання, починає поступово зникати (рис. 2). Розподіли умовної ймовірності летального ураження людини внаслідок впливу теплового випромінювання від високотемпературного газового струменя у контрольних точках P_i для варіанту без перепони представлені на рис. 3а. Видно, що з віддаленням від епіцентру теплового випромінювання максимальна умовна ймовірність летального наслідку поступово плавно зменшується.

Установлення перепони кардинально змінює безпековий стан у контрольних точках (рис. 36). Точка P_1 із найнебезпечнішої у варіанті без перепони стає навпаки найбільш безпечною внаслідок максимальної захищеності від теплових потоків завдяки перепоні. З відстанню від перепони її корисна дія зменшується, відкриваючи доступ більшій кількості джерел теплового випромінювання, але й інтенсивність випромінювання вщухає, як і у першому варіанті без перепони (рис. 3).

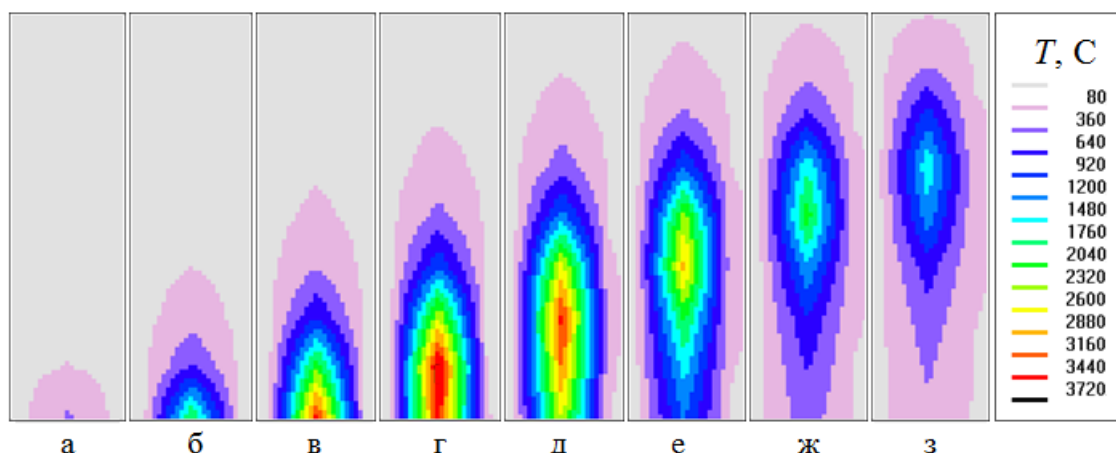


Рис. 2 Поля температури у площині YOZ: а-з – у моменти часу 1, 3, 5, 7, 9, 11, 13, 15 с, відповідно

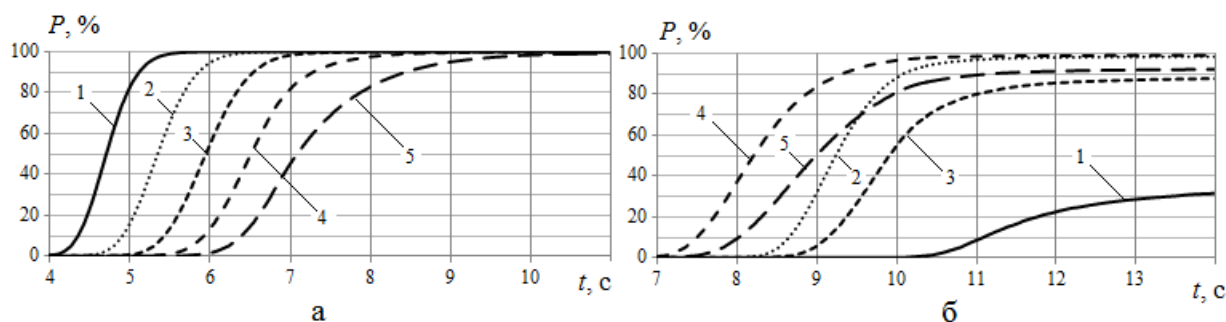


Рис. 3 Зміна ймовірності ураження без захисту (а) та з ним (б): 1-5 – відповідні контрольні точки P_i

Висновки

Виявлений характер розподілу ймовірності ураження дає змогу стверджувати, що необхідно урахувати вплив непроникних для теплових потоків перепон для оцінки безпекового стану на об'єкті, де має місце викид високотемпературних газів

ЛІТЕРАТУРА

1. Skob, Y.A., Ugryumov, M.L., Granovskiy, E.A. Numerical assessment of hydrogen explosion consequences in a mine tunnel. *International Journal of Hydrogen Energy*. 2021. Vol. 46. Issue 23. P. 12361–12371. DOI: <https://doi.org/10.1016/j.ijhydene.2020.09.067> (Last accessed: 30.10.2022).
2. Skob Y., Ugryumov M. and Granovskiy E. Numerical Evaluation of Probability of Harmful Impact Caused by Toxic Spill Emergencies. *Environmental and Climate Technologies*. 2019. Vol. 23. Issue 3. P. 1–14. DOI: <https://doi.org/10.2478/rtuct-2019-0075> (Last accessed: 30.10.2022)

СКОБ Юрій Олексійович – д. т. н., доцент; професор кафедри математичного моделювання та штучного інтелекту; Харківський національний аерокосмічний університет імені М. Є. Жуковського «Харківський авіаційний інститут», вул. Чкалова, 17, Харків, Україна, 61070; e-mail: y.skob@khai.edu; yuriy.skob@gmail.com; ORCID: 0000-0003-3224-1709.

Наукові інтереси:

– математичні моделі в механіці газових сумішей.

СУХОВЕРСЬКИЙ Д.О., ДОЛЯ Г.М.

МОДЕЛЬ РОБОТИ ВОЛОКОННО-ОПТИЧНОЇ ЛІНІЇ ЗВ'ЯЗКУ ІЗ АКУСТООПТИЧНИМ МОДУЛЯТОРОМ СВІТЛА

The topic of the work is a computer process simulation of a light beam passing through an acousto-optic modulator using the two-dimensional Fourier transform.

The work purpose is to get acquainted with Fourier analysis in optics, the principles of operation of an acousto-optic modulator, to study the angle influence of a light beam incidence on a Bragg diffraction grating in an acousto-optic modulator on the useful signal level change, and to calculate the bit error ratio (BER) change in the fiber-optic communication line (FOLC) based on the useful signal level change.

As research methods, computer process simulation of a light beam passing through an acousto-optical light modulator was carried out. Also, an analysis of the deviation influence of the incidence on the bit error rate was performed based on the simulation results.

The report results are a computer model of the acousto-optic modulator, based on which the deviation influence of the incidence angle of the beam from the Bragg angle on the bit error rate in the fiber-optic communication line was studied.

The report concludes that compliance with the angle of incidence of the beam in the VOLZ is critically important. Also, the simulation gave us accurate data on the useful signal change when the beam incidence angle is deviated. Finally, based on this data, it is possible to calculate the BER error in fiber-optic communication lines that contain an acousto-optic modulator.

Ключові слова: Акустооптичний модулятор, Дифракція Брегга, Швидке перетворення Фур'є, Коефіцієнт бітових помилок (BER).

Сучасний світ – світ інформатики. Інформаційні процеси у складних розподілених системах неможливі без високоякісних волоконно-оптичних ліній зв'язку (ВОЛЗ), ці лінії зв'язку містять елементи, якість яких впливає на якість інформації, що передається, якість інформації вимірюється коефіцієнтом бітових помилок (англ. аббр. BER – “Bit error ratio”).

Для розрахунку BER потрібен лише один параметр - Q-фактор, що визначається наступним виразом:

$$Q = \frac{|\mu_1 - \mu_0|}{\sigma_1 + \sigma_0}, \quad (1)$$

де μ_1 – рівень корисного сигналу, μ_0 – рівень шуму, σ_1, σ_0 - середньоквадратичні відхилення корисного сигналу та шуму відповідно.

Рівень корисного сигналу може деградувати від старіння лінії зв'язку або її елементів (наприклад, акустооптичного модулятора). В акустооптичному модуляторі світла за допомогою п'єзоелектричного перетворювача створюється біжуча ультразвукова хвиля в оптично прозорому матеріалі. Завдяки появі ділянок стиснення та розтягування, що відрізняються показником заломлення, в середовищі формується дифракційна ґратка вздовж напрямку розповсюдження акустичної хвилі.

При падінні світлової хвилі на дифракційну решітку кожен максимум ультразвукової хвилі відбиває світлову хвилю. Таким чином формується багато відбитих світлових хвиль з оптичною різницею ходу Δ :

$$\Delta = 2\Lambda \sin \theta, \quad (2)$$

де Λ – довжина звукової хвилі, θ - кут падіння променя.

Якщо різниця ходу дорівнює довжині світлової хвилі λ , відбиті промені за рахунок інтерференції посилюють одне одного. Отже для першого максимуму інтенсивності відбитого променя справедлива умова Вульфа-Брегга:

$$2\Lambda \sin \theta = \lambda, \quad (3)$$

де Λ – довжина звукової хвилі, θ - кут падіння променя, λ – довжина світлової хвилі.

Потужність сигналу на виході модулятора критично залежить від точності дотримання кута Брегга. Щоб встановити кількісне співвідношення впливу стабільності кутового положення вхідного променя по відношенню до акустичної решітки було розроблено імітаційну модель.

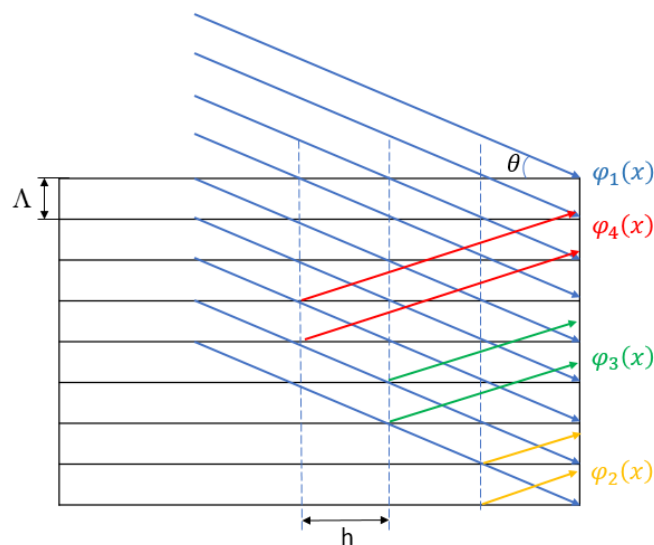


Рис.1. Модель проходження світлового променя через акустичну дифракційну решітку

На кожному із слоїв решітки відбувається відбиття променю, фаза відбитих променів різна та при інтерференції промені можуть посилювати одне одного, що відбувається саме при падінні променю під кутом Бреґга, при відхиленні кута падіння через різницю фаз інтерферовані хвилі послаблюють одна одну. В даній моделі ми обчислюємо відбиті промені від кожного слою решітки та обраховуємо їх двумірне перетворювання Фур'є. Потім отримані інтенсивності для кожного слою складаються за формулою:

$$I = |e_1 + e_2|^2, \quad (4)$$

де I – інтенсивність результуючого променю, e_1, e_2 – інтенсивності різних слоїв.

Моделювання було проведене в середовищі Matlab.

Початкові дані:

- 1) Довжина світлової хвилі $\lambda = 850$ нм;
- 2) Довжина звукової хвилі $\Lambda = 12.5$ нм;
- 3) Кут падіння відповідає Бреґгівському;
- 4) Показник заломлення $n = 1.333$;
- 5) Товщина $d = 1$ мм.

По результатам моделювання був побудований такий графік:

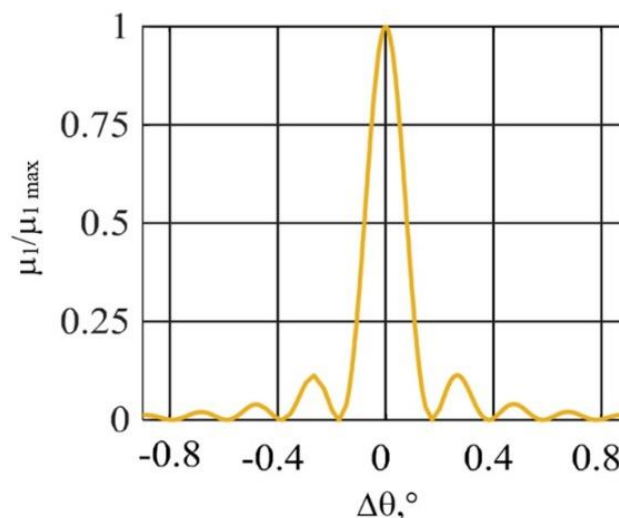


Рис. 2. Залежність дифракційної ефективності акустооптичного модулятора від відхилення кута падіння від кута Бреґга (по осі X - відхилення кута падіння променю у градусах, по осі Y - рівень корисного сигналу по відношенню до рівню сигналу).

Для дослідження критичності дотримання кута Бреґга був побудований графік залежності коефіцієнту бітових помилок в діапазоні 10 кутових секунд:

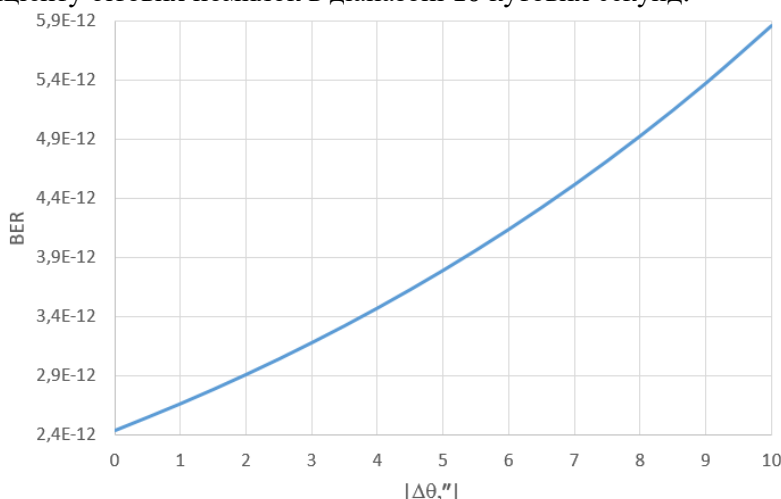


Рис. 4. Залежність BER у ВОЛЗ від відхилення кута падіння від кута Бреґга (по осі X – відхилення кута падіння у кутових секундах, по осі Y – коефіцієнт бітових помилок).

Отже, за результатами моделювання, ми побачили що відхилення кута падіння хоча б на одну секунду збільшує імовірність помилки з $2,44089E-12$ до $2,66738E-12$, або на 9,279%. Що є значним впливом, при цьому функція швидко зростає, тому дотримання точності кута Бреґга є критично важливим. Отримані закономірності та числові оцінки можуть бути корисні розробникам волоконно-оптичних ліній зв'язку для складних розподілених інформаційних систем.

ЛІТЕРАТУРА

1. Roger L. Freeman Fiber-Optic Systems for Telecommunications Wiley , 2 Aug. 2002 - 416
2. Korpel, A. Acousto-Optics, Second Edition; Technology & Engineering, CRC Press, 1996 - 360.

СУХОВЕРСЬКИЙ Дмитро Олександрович – e-mail: adelfis17@gmail.com; ORCID: 0000-0003-4256-3293.

Наукові інтереси:

– комп'ютерне моделювання оптичних процесів.

ДОЛЯ Григорій Миколайович – д. т. н., професор; професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: gdolya@karazin.ua; ORCID: 0000-0001-8888-5189.

Наукові інтереси:

– оптоінформатика, лазерна локація, спеклометрія.

ТЕЛЕЖЕНКО Д.О., ТОЛСТОЛУЗЬКА О.Г.

ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ АРХІТЕКТУРИ ВІРТУАЛЬНИХ РОЗПОДІЛЕНИХ СИСТЕМ

Тема роботи. Методи та моделі синтезу архітектури віртуальних розподілених комп'ютерних систем.

Мета роботи. Ознайомлення та вивчення найбільш розповсюджених методів та моделей синтезу архітектури віртуальних розподілених комп'ютерних систем.

Методи дослідження. Аналіз, пояснення, класифікація.

Результати. Результатом виконання роботи являється огляд і постановка задачі дослідження архітектури віртуальних розподілених систем.

Висновки. У роботі були розглянуті постановка задачі дослідження архітектури віртуальних розподілених систем, проведений аналіз та розглянута постановка задачі. Дана архітектурна проблема покриває великий спектр задач у різних галузях в віртуальних розподілених системах. Ця архітектура потрібна для того, щоб зменшити потреби у ресурсах клієнта та розмір надійної обчислювальної бази, що встановлює фізичну ізоляцію між ними та службою віртуальних машин і створює єдину точку адміністрування. Така розподілена архітектура віртуальної машини може забезпечити суттєво кращу цілісність та керованість, ніж монолітна архітектура, так як вона архітектура повинна добре масштабуватися із збільшенням кількості клієнтів і не спричиняти великих накладних витрат.

Ключові слова: Віртуальні машини, архітектура, система.

Вступ

Сучасні віртуальні машини, такі як Java та Inferno, виникли як мережеві обчислювальні платформи. Хоча ці віртуальні машини забезпечують абстракції вищого рівня та більш складні послуги, ніж їх попередники двадцять років тому, їх архітектура по суті залишається незмінною. Сучасні віртуальні машини все ще є монолітними, тобто вони складаються з тісно пов'язаних між собою сервісних компонентів, які таким чином копіюються на всі комп'ютери в організації. Ця груба копія послуг утворює одне з найслабших місць у сучасних мережевих системах, оскільки створює загальноновизнані та широко розголошені проблеми безпеки, керованості та продуктивності.

Віртуальні машини так само як і фізичні машини можуть піддаватись впливу небезпечного (вірусного) програмного забезпечення. Щоб убезпечити свої віртуальні машини, треба захищати своє середовище так само, як захищається фізична приватна машина. Треба подумати про те, щоб вимкнути непотрібні функції, звести до мінімуму використання консолі віртуальної машини та дотримуватись інших найкращих практик. Важливим аспектом розподілених віртуальних систем є також як продуктивність так і керованість. Так як некоректно налаштована система, може буде занадто складна у керованості, що згодом призведе до проблем також і у масштабуванні, що у свою чергу погано впливає на продуктивність таких систем. Для того, щоб мінімізувати такі потенційні проблеми необхідно розробити архітектуру для системи, що у свою чергу буде робити систему "гнучкою" для змін.

Аналіз проблематики та актуальність.

Побудова розподіленої архітектури включає багато різних компонентів і багато різних стратегій, від клієнт-серверних обчислень до керування великими даними та вмістом. Найбільш розповсюдженими методами синтезу архітектури віртуальних розподілених комп'ютерних систем є наступні:

- **Розділення сервісів базуючись на функції**, що виконують дані сервіси, тобто ефективно компонентизувати свою програму. Ідеальним компонентом є програмний блок, який можна розробляти та керувати незалежно від інших блоків програми. Метою побудови розподіленої системи є розробка програми, яка працює навіть якщо кожна частина зберігає свою незалежність, вона повинна залишатися синхронізованою з цілим. Ми хочемо мати можливість ізолювати погано працюючих "учасників" або змінювати послідовність, не зачіпаючи випадково інших учасників.

- **Чітко визначити межі обслуговування.** Треба звертати увагу на частини, які складають єдине ціле. Наскільки добре архітектор системи розділе свою програму на її складові служби, на стільки добре або погано це вплине на синхронізацію процесів і міжфункціональний зв'язок.
- **Визначити, як розподілені служби будуть спілкуватися.** У мікросервісі складові служби є компонентами поза процесом. Вони спілкуються за допомогою запиту веб-служби або віддаленого виклику процедури. Треба звести до мінімуму зв'язок між службами, однак це обмеження неможливе, якщо вашим службам потрібно здійснювати кілька зовнішніх викликів, як-от платіжний шлюз, наприклад.

Ретельна архітектура програмного забезпечення починається з концепції та розвивається з неї. Навіщо витрачати час на розробку чудового програмного забезпечення лише для того, щоб воно мало продуктивність через не дуже хорошу віртуальну машину чи операційну систему? Треба думати про функції, а не про всю програму. Для безперервної інтеграції (CI) необхідно, щоб функції були повністю мобільними та зручними для автоматизації.

Однак аналіз концепцій та принципів побудови віртуальних розподілених комп'ютерних систем показує, що головним їх недоліком є відсутність урахування особливостей паралельних обчислювальних систем.

Взагалі, існує багато підходів до процесу архітектурного проектування, які залежать від професійного досвіду, а також від майстерності та інтуїції розробників. Але треба відмітити декілька етапів, загальних для всіх процесів архітектурного проектування.

Структурування системи. Програмна система структурується у вигляді сукупності відносно незалежних підсистем. А також визначається взаємодія між системами.

Моделювання керування. Розробляється базова модель керування взаємовідношеннями між частинами системи.

Необхідно зауважити, що архітектуру великих систем неможливо описати за допомогою одного типу моделі. Тобто при розробці окремих частин системи можемо використовувати різні архітектурні моделі. В такому випадку архітектура системи може бути занадто складною, тому важливо розробнику підібрати найбільш відповідну модель, і в подальшому модифікувати її відпо-відно до вимог системи яка розробляється. Архітектура системи впливає на продуктивність, надійність, зручність супроводження та інші характеристики системи. Тому моделі архітектури, вибраної для даної системи, можуть залежати від нефункціональних системних вимог.

Постановка задачі.

Для того, щоб вирішити описану проблему, необхідно розробити та впровадити нову системну архітектуру для мережевих обчислень на базі розподілених віртуальних машин. Ця архітектура потрібна для того, щоб зменшити потреби у ресурсах клієнта та розмір надійної обчислювальної бази, що встановлює фізичну ізоляцію між ними та службою віртуальних машин і створює єдину точку адміністрування. Така розподілена архітектура віртуальної машини може забезпечити суттєво кращу цілісність та керованість, ніж монолітна архітектура, так як вона архітектура повинна добре масштабуватися із збільшенням кількості клієнтів і не спричиняти великих накладних витрат.[1]

Таким чином виникає задача вдосконалення існуючих методів та засобів синтезу архітектури віртуальних розподілених комп'ютерних систем.

Можна стверджувати, що запровадження нової загальної архітектури віртуальних розподілених комп'ютерних систем повинна привести до полегшення роботи розподілених систем, та покращити масштабування у разі потреби.[2]

Викладене вище обумовлює актуальність рішення науково-прикладної задачі створення методів та моделей синтезу архітектури віртуальних розподілених комп'ютерних систем з метою підвищення ефективності роботи розподілених систем в цілому, а також у подальшому його масштабуванні.[3]

Для досягнення даної мети як рішення поставленого наукового завдання в цілому, були сформульовані ряд задач. До їхнього числа належать.

1. Аналіз концепцій і принципів побудови віртуальних розподілених комп'ютерних систем з урахуванням паралельних обчислювальних систем.

2. Розробка моделі синтезу архітектури віртуальних розподілених комп'ютерних систем.
3. Розробка методів масштабування віртуальних розподілених комп'ютерних систем.
4. Практичне підтвердження працездатності й вірогідності розроблених моделей, методів синтезу архітектури віртуальних розподілених комп'ютерних систем. [5]

Об'єкт дослідження – процеси синтезу архітектури в віртуальних розподілених комп'ютерних системах.

Предмет дослідження – моделі й методи синтезу архітектури в віртуальних розподілених комп'ютерних системах.

Методи дослідження. Теоретичні основи роботи базуються на принципах і методах системного аналізу, математичного й імітаційного моделювання. Як математична основа для розробки методів та моделей синтезу архітектури віртуальних розподілених комп'ютерних систем і перспективних розподілених обчислювальних систем, що задовольняє заданим вимогам й обмеженням, використовуються теорія множин, теорія графів, комбінаторні методи дискретної математики й алгебра просторово-часових структур семантико-числової специфікації.

Висновки

У роботі були розглянуті постановка задачі дослідження архітектури віртуальних розподілених систем, проведений аналіз та розглянута постановка задачі. Дана архітектурна проблема покриває великий спектр задач у різних галузях у віртуальних розподілених системах. Ця архітектура потрібна для того, щоб зменшити потреби у ресурсах клієнта та розмір надійної обчислювальної бази, що встановлює фізичну ізоляцію між ними та службою віртуальних машин і створює єдину точку адміністрування. Така розподілена архітектура віртуальної машини може забезпечити суттєво кращу цілісність та керованість, ніж монолітна архітектура, так як вона архітектура повинна добре масштабуватися із збільшенням кількості клієнтів і не спричиняти великих накладних витрат.

ЛІТЕРАТУРА

1. Arora, Sanjeev; Barak, Boaz (2009), Computational Complexity – A Modern Approach, Cambridge, ISBN 978-0-521-42426-4.
2. Морозов О.О. Ситуаційні центри – основа управління системами великої розмірності /О.О. Морозов // Математичні машини та системи, 1997, № 2. – С. 7–10.
3. Ghosh, Sukumar (2007), Distributed Systems – An Algorithmic Approach, Chapman & Hall/CRC, ISBN 978-1-58488-564-1.
4. Roosta, Seyed H. (2000). «Parallel processing and parallel algorithms: theory and computation». Springer, p. 114. ISBN 0-387-98716-9.
5. Almasi, G.S. and A. Gottlieb (1989). Highly Parallel Computing. Benjamin-Cummings publishers, Redwood City, CA.
6. Curriculum Guidelines for Undergraduate Degree Programs in Computer Engineering. Association for Computing Machinery. 2004. с. 60. Архів оригіналу за 12 червня 2019.
7. Дэвид М. Харрис, Сара Л. Харрис. Цифровая схемотехника и архитектура компьютера. / пер. с англ. Imagination Technologies. – М.: ДМК Пресс, 2017. – 792 с.: ил.

ТЕЛЕЖЕНКО Денис Олександрович – аспірант; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: denisque75@gmail.com; ORCID: 0000-0002-8377-8517.

Наукові інтереси:

– технології розподіленого і паралельного програмування.

ТОЛСТОЛУЗЬКА Олена Геннадіївна – д. т. н., с. н. с.; професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: elena.tolstoluzka@karazin.ua; ORCID: 0000-0003-1241- 7906.

Наукові інтереси:

– технології розподіленого і паралельного програмування.

ТОВСТОКОРЕНКО О.Ю.

ПРОБЛЕМИ ПІДТРИМКИ ПРОЦЕСІВ КОНФІГУРУВАННЯ СЦЕНАРІЇВ АВТОМАТИЗАЦІЇ У РОЗРОБЦІ ТА СУПРОВОДІ СИСТЕМ «РОЗУМНИЙ БУДИНОК»

Вступ

Сьогодні однією з найбільш поширених областей домену «Інтернету речей» безумовно є «Розумний будинок», який об'єднує в собі різні типи будівель та відповідно задач, пов'язаних з кожним з типів. Але незалежно від типу будівлі, усі подібні системи повинні мати функціональність, що забезпечує можливість налаштування автоматизації. В більшості програмних систем управління «Розумним будинком» (СУРБ) функціональність створення та редагування скриптів автоматизації надає можливості застосування методів та моделей штучного інтелекту. Хоча вже існує досить велика кількість публікацій за даною темою, все ж такі питання як проблеми конфігурування сценаріїв автоматизації у розробці та супроводі СУРБ лишаються ще недостатньо опрацьованими, і представлена робота має на меті аналіз процесу створення сценаріїв автоматизації різних SHMS, формування набору можливих варіантів для частково автоматизації зазначеного процесу.

Актуальність і стан досліджень у предметній галузі

Наприклад, у [1] пропонується синтез аналізу штучного інтелекту та контролю Інтернету речей (IoT) для використання індивідуальному навчальному середовищі. Як результат застосування запропонованого підходу, пристрої IoT можна налаштувати, щоб визначити, чи може та чи інша модель аналізу AI ефективно підвищити рівень концентрації користувача у навчальному середовищі. У [2] зібрано інформацію з кількох дослідницьких статей щодо Інтернету речей, штучного інтелекту та досліджено зв'язок між цими темами з метою всебічного представлення та узагальнення відповідної літератури в цих областях.

Поширення підключених пристроїв IoT створює велику кількість задач аналізу даних для підходів на основі штучного інтелекту. Час відгуку, необхідний таким пристроям, вимагає обробки даних IoT на периферії, але на межі зазвичай бракує ресурсів для вивчення моделей штучного інтелекту. У [3] запропоновано архітектуру, яка надає часткове рішення зазначеної проблеми. На відміну від існуючих опитувань, [4] детально розглядає технології на основі штучного інтелекту та Інтернету речей у трьох основних аспектах: розумне обслуговування, розумна стійкість і розумна безпека. Потім автори додатково висвітлюють тенденцію до майбутніх розумних бібліотек. У [5], спираючись на літературу про штучний інтелект, IoT та IoNT, автори описують поточний стан досліджень IoNT та їх наслідки, а також пропонують, як штучний інтелект можна використовувати у програмах IoNT.

Запропонований підхід

В результаті аналізу деяких існуючих СУРБ [6] з відкритим програмним кодом, було побудовано загальну схему процесу створення сценарію автоматизації (СА), який у вигляді діаграми IDEF0 представлено на рис. 1.

Першим кроком в зазначеному процесі є підготовка пристроїв системи та налаштування їх програмних модулів (у випадку необхідності), результатом кроку є конфігураційний файл інтеграції у форматі ".xml", який визначає назву пристрою, метадані вихідної інформації для сенсорів та метадані вхідної інформації для актуаторів.

Другим кроком є визначення набору пристроїв з конфігураційного файлу інтеграції, які будуть використовуватися в СА.

На третьому кроці, на основі списку пристроїв з конфігураційного файлу інтеграції визначаються події тригери, у випадку виявлення яких, СУРБ буде розпочинати виконання сценарію автоматизації.

На четвертому кроці визначаються умови, які необхідно перевірити у випадку спрацьовування тригера та перед виконанням безпосередньо дій скрипту.

На останньому, п'ятому кроці, визначається перелік дій, які необхідно виконати.

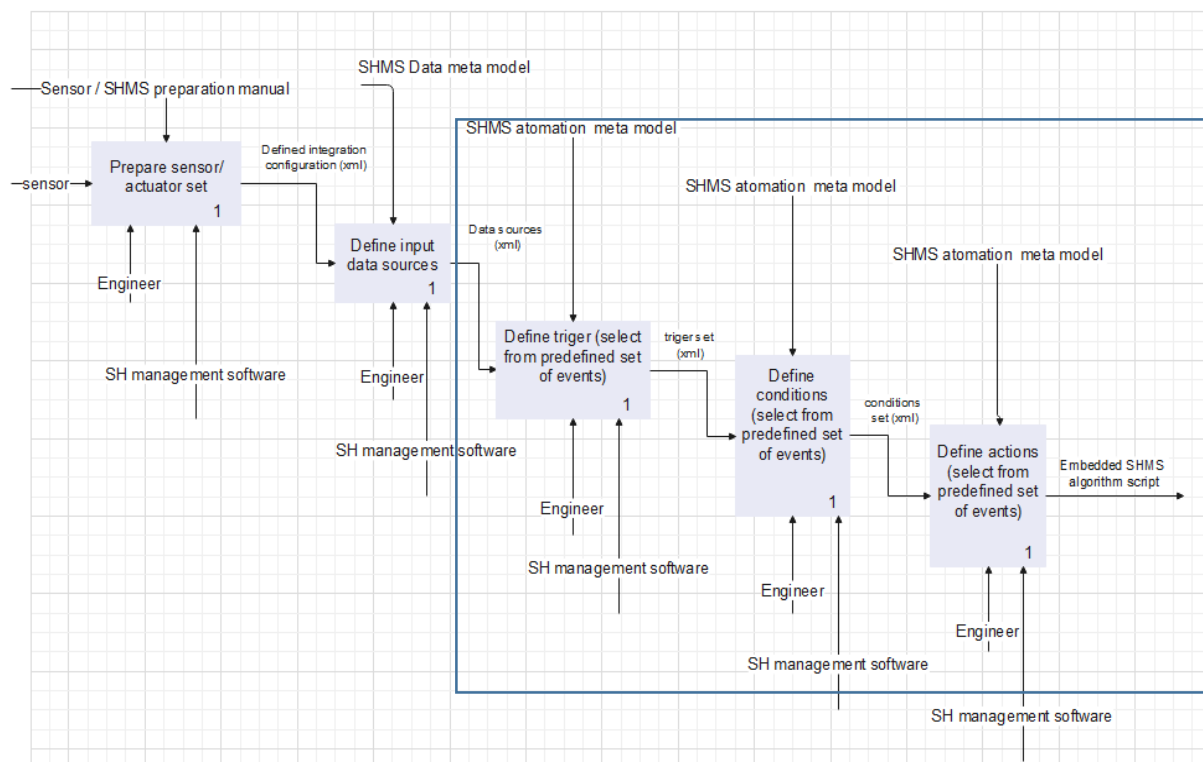


Рис. 1 - Загальна схема процесу створення сценарію автоматизації SHMS

На кроках з третього по п'ятий, тригери, умови та дії мають обмежений перелік варіантів в контексті налаштованої системи, та є можливість вибору одного чи декількох елементів зі списку. Зазначені набори формуються, в більшій мірі, на основі можливостей компонентів, тому, при наявності підходів до накопичення варіантів налаштувань [7], даний процес може бути частково автоматизовано за рахунок застосування методів штучного інтелекту, наприклад Case-based reasoning (CBR).

CBR — це підхід до вирішення нових задач, заснований на досвіді, шляхом адаптації раніше успішних рішень до подібних задач. У CBR повторне використання знань із уже розв'язаних задач або кейсів базується на передумові, що чим схожими є дві проблеми, тим більш схожими будуть їхні рішення [8].

Під час першого кроку, пошуку, нова задача зіставляється з задачами попередніх випадків за допомогою обчислення функції подібності, наприклад латентно-семантичний аналіз, і знаходиться найбільш подібна задача та її збережене рішення. Якщо запропоноване рішення не відповідає необхідним вимогам нової задачі, відбувається наступний крок, адаптація, і створюється нове рішення. Отримане рішення та нова задача разом утворюють новий випадок, який включається в базу інцидентів під час етапу навчання. Таким чином, алгоритм CBR адаптується, оскільки можливості системи розширюються шляхом накопичення нового досвіду.

Висновки

В роботі представлено загальну схему процесу створення сценарію автоматизації SHMS. На схемі виділено проблемну область, яка може бути повністю або частково автоматизованою за рахунок використання методів штучного інтелекту. Як один із варіантів, було запропоновано CBR для часткової автоматизації зазначених кроків.

У якості подальших досліджень планується проаналізувати альтернативні до CBR методи штучного інтелекту для вирішення описаної задачі. На основі аналізу методів планується

розробка підходу та програмного прототипу до підтримки процесу створення та підтримки сценаріїв автоматизації для систем управління «Розумним будинком».

ЛІТЕРАТУРА

1. C. -Y. Liu, H. -Y. Yu and I. -C. Tsai, "Applying AI Analysis-based IoT System Control to the Individualized Learning Field," *2019 IEEE International Conference on Engineering, Technology and Education (TALE)*, 2019, pp. 1-7, doi: 10.1109/TALE48000.2019.9225908.
2. Kuzlu, M., Fair, C. & Guler, O. Role of Artificial Intelligence in the Internet of Things (IoT) cybersecurity. *Discov Internet Things* **1**, 7 (2021). <https://doi.org/10.1007/s43926-020-00001-4>
3. Calo, Seraphin B., Maroun Touma, Dinesh C. Verma and Alan Cullen. "Edge computing architecture for applying AI to IoT." *2017 IEEE International Conference on Big Data (Big Data)* (2017): 3012-3016.
4. Bi, Siguo, Cong Wang, Jilong Zhang, Wutao Huang, Bochun Wu, Yi Gong, and Wei Ni. 2022. "A Survey on Artificial Intelligence Aided Internet-of-Things Technologies in Emerging Smart Libraries" *Sensors* **22**, no. 8: 2991. <https://doi.org/10.3390/s22082991>
5. Mishra, Anoop, Tripathi, Abhishek and Khazanchi, Deepak. "Chapter 1. A framework for applying artificial intelligence (AI) with Internet of nanothings (IoNT)". *Machine Learning for Sustainable Development*, edited by Kamal Kant Hiran, Deepak Khazanchi, Ajay Kumar Vyas and Sanjeevikumar Padmanaban, Berlin, Boston: De Gruyter, 2021, pp. 1-16. <https://doi.org/10.1515/9783110702514-001>.
6. Товстокоренко О.Ю., Ткаченко М.В., Використання концепції лінійок програмних продуктів в контексті процесу проектування систем «Розумний дім». «Наукові підсумки 2019 року», XXXVII Міжнародна науково-практична інтернет-конференція. – м. Вінниця, 09 грудня 2019 року. – Ч.14, с. 84.
7. Гамзаєв Р. О. Застосування методів доменного моделювання для підтримки варіабельності програмного забезпечення в розробці систем «Розумний будинок» / Р. О. Гамзаєв, М. В. Ткачук, О. Т. Товстокоренко // Інформаційні системи та технології: матеріали статей 9-ї Міжнародної науково-технічної конференції, Харків, 17-20 листопада 2020 року / наук. ред. А. Д. Тевяшев, Л. Б. Петришин, В. Г. Кобзєв. – ХНУРЕ. – Х.: Друкарня Мадрид, 2020
8. Juan A. Recio-Garcia, Derek Bridge, A Case-Based Template Recommender System for Building Case-Based Systems 2008, Department of Software Engineering and Artificial Intelligence, Universidad Complutense de Madrid, Spain,-18p.

ТОВСТОКОРЕНКО Олег Юрійович – доктор філософії (PhD) з комп'ютерних наук, старший викладач кафедри моделювання систем і технологій факультету комп'ютерних наук, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, м. Харків, 61022; e-mail: oleg.tovstokorenko@karazin.ua; ORCID: 0000-0003-2664-1650.

Наукові інтереси:

– аналіз та моделювання розподілених програмних систем, моделі та технології розробки лінійок програмних продуктів, знання-орієнтовані методи розробки та інженерії якості програмного забезпечення.

УДК 65.0(075.8)

СЕКЦІЯ 2,7

ТОЛСТОЛУЗЬКИЙ Є.Д., МІРОШНИК М.А.

МОДЕЛЬ МЕРЕЖЕВОГО ПЛАНУВАННЯ ІЗ ЗАСТОСУВАННЯМ МЕТОДІВ МУЛЬТИПАРАЛЕЛЬНОЇ ОБРОБКИ ІНФОРМАЦІЇ

Theme of the research: network planning model using multiparallel information processing methods.

The main purpose of the research - the purpose of the work is to describe the methods of multiparallel information processing and how to use them for network planning of IT projects. The use of multiparallel information processing methods will reduce the time for calculating network graphs, will allow analysis of the duration of work based on the ability to dynamically set the required amount of resources (operators performing tasks).

The methods of the research - to build and calculate a network graph, the PERT methodology, multiparallel information processing methods, C++ and JS languages are used.

Results - network work graph model was created, using the PERT methodology, the risks of performing work were calculated, the critical path and options for performing work with a different number of employees, from one to the maximum possible number, were calculated.

Conclusions - the use of multi-parallel information processing methods is possible and effective for network planning calculations and building network graphs, however, it is an insufficient apparatus for the sole calculation of the necessary indicators at this design stage.

Key words: multiparallel, network planning, PERT, network graphs.

Вступ

Процес проектування є невід'ємною частиною успішного та тривалого проекту. Інформаційні технології дозволяють спростити і автоматизувати цей процес.

Метою будь-якого проекту є створення унікального продукту. Робота над цим проектом включає обчислювальну та комунікаційну підтримку, інформаційні ресурси та персонал, відповідальний за підтримку динамічних інформаційних моделей.

Чим довший проект, тим більші ризики виникають під час реалізації. Вони впливають на дохід, час розробки тощо. Ризики проекту включають: кадровий потенціал, кількість людей, якість технічної підтримки, людські помилки тощо. Проаналізувавши ці критерії, можна сказати, що автоматизація будь-якого процесу знижує ризик людської помилки та скорочує час планування.

У даній роботі розглядається можливість автоматизації процесу побудови мережесх графіків за допомогою підходу мультипаралельної обробки, який дозволяє перетворювати алгоритми лінійної дії в паралельні алгоритми. Метою даного дослідження є скорочення часу мережевого планування для виконання ІТ-проектів за рахунок використання методів мультипаралельної обробки.

Постановка задачі

Необхідно створити модель, здатну автоматично побудувати мережеву діаграму на основі даних, отриманих на етапі структурного планування, що є елементом мережевого планування. Тобто з набору готових даних, проектів та їх характеристик побудувати мережеву діаграму проекту та проаналізувати часові витрати та ризик виконання проекту в заданий термін. Початковий набір даних може виглядати так:

i	j	t_{i-j}^o	t_{i-j}^{mv}	t_{i-j}^p	t_{i-j}^{cp}	σ_{i-j}^2
0	1	14	20	25	19.8	3.36
1	2	4	6	10	6.3	1
2	3	7	10	15	10.3	1.7
0	4	3	4	5	4	0.1
4	5	5	7	9	7	0.4
4	6	15	22	28	21.8	4.6
5	7	10	16	20	15.6	2.7
7	8	8	11	14	11	1
8	9	5	13	18	12.5	4.6
3	9	5	12	18	11.8	4.6

Рис. 1. Приклад таблиці параметрів мережевого графа

Таблиця структурована відповідно до методології PERT (Project Evaluation and Review Technique). Цей метод було обрано, оскільки він зосереджений на аналізі часу та ідентифікації ризиків проекту. Методи мультипаралельної обробки було обрано, оскільки вони призначені для розпаралелювання процесів і дозволяють автоматично обчислювати та будувати схеми паралельних графів (мережових графів) різної глибини від однієї гілки до максимального розпаралелювання.[1-3].

На даний час в літературі описано п'ять методів паралельної обробки інформації:[4]

- метод суміщення незалежних операторів (СО);
- метод конвеєрної обробки (КО);
- метод декомпозиційної обробки (ДО);
- кодово-матричний метод (КММ);
- метод суміші алгоритмів (СА).

Метод суміщення незалежних операторів (робіт) полягає в одночасному початку виконання у кожний з дискретних моментів часу деякої кількості робіт, для яких виконуються такі умови:

- а) ці роботи не пов'язані послідовними зв'язками;
- б) для кожного з таких робіт до моменту часу, що розглядається, попередня робота є завершеною.

Метод конвеєрної обробки полягає у виконанні для алгоритму наступних перетворень:

- а) виконанні фрагмента “конвеєра” (гілки) робіт від початку до кінця одну роботу за іншою;
- б) послідовна реалізація виконання гілок, при відсутності здатності виконувати їх паралельно;

Метод декомпозиційної обробки (ДО) забезпечує можливості подальшого (у порівнянні з конвеєрною обробкою) збільшення швидкості виконання проекту за рахунок часткового часові поєднання процесів конвеєрної обробки ряду суміжних наборів робіт; одночасно з цим зростає час вирішення робіт у кожній незалежній гілці.

Метод суміші полягає у створенні на проектному рівні суміші робіт, використанні такої суміші для досягнення 100% завантаження виконавця робіт та забезпечення за рахунок цього потенційно можливого підвищення ефективності реалізації аналізованої сукупності робіт. Суміш робіт формально розглядається і виконується як єдина робота.

Модель

Використовуючи метод суміщення незалежних операцій дозволило отримати вигравш в часі як показано на слайді порівняно з використанням традиційного підходу. Нижче приведені послідовна та паралельна граф схеми для порівняння часу виконання проекту.

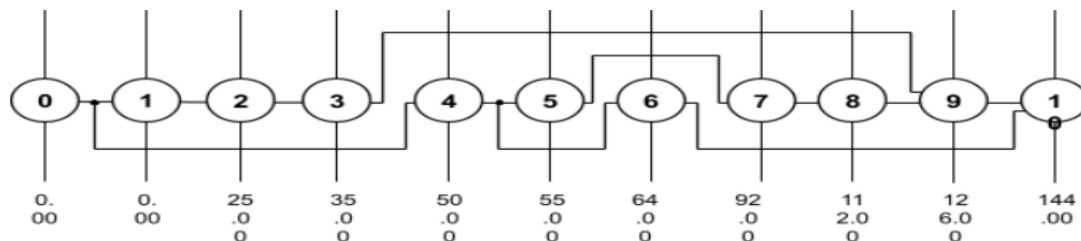


Рис. 2. Мережевий граф послідовної реалізації проекту

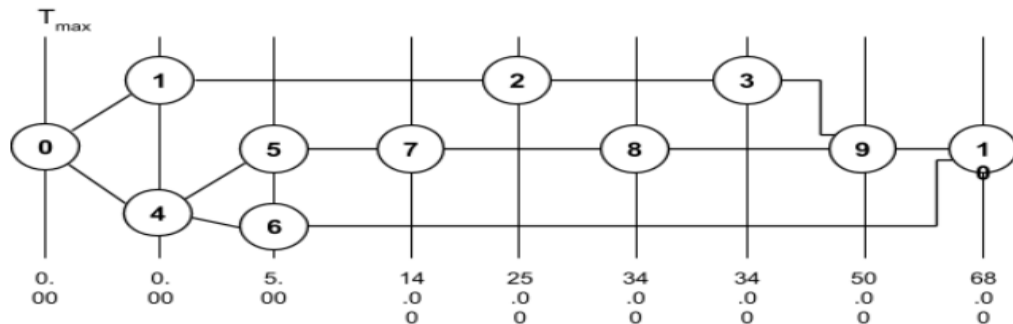


Рис. 3. Мережевий граф паралельної реалізації проекту

Можемо побачити, що послідовний час виконання проекту складає 144, коли паралельний час виконання складає 68. Маючи ці дані, лице яке приймає рішення (бізнес-аналітик, проєкт-менеджер, клієнт) обирає необхідні конфігурації проекту. Тобто, чи можна виділити більше бюджету на роботу трьох працівників, та скоротити час виконання проекту, або є обмеження у бюджеті і робота одного працівника над проектом є раціональною.

Висновки

На основі проведеного аналізу літератури та досліджень в дипломі бакалавра, можна зробити висновок, що використання раціональної сукупності методів паралельної обробки інформації надасть можливість підвищити якість розробки плану робіт ІТ проекту.

ЛІТЕРАТУРА

1. Разу М.Л.(2006).Управління проектом. Основи проектного управління.
2. ProjectManager. Project Management guide[Електронний ресурс]/ ProjectManager – 2020 – Режим доступу до ресурсу: <https://www.projectmanager.com/project-management> .
3. Інтуїт. Мережеве планування [Електронний ресурс]/ Інтуїт – 2013 – Режим доступу до ресурсу: <https://intuit.ru/studies/courses/496/352/lecture/8389> .
4. Поляков Г.А., Шматков С.И., Толстолужская Е.Г., Толстолужский Д.А. (2012). Синтез и анализ параллельных процессов в адаптивных времяпараметризованных вычислительных системах. – Х.: ХНУ имени В.Н. Каразина, 2012.-672с;

ТОЛСТОЛУЗЬКИЙ Євген Дмитрович – студент факультету комп'ютерних наук Харківського національного університету імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: evventol@gmail.com ; ORCID: 0000-0002-2039-0267.

Наукові інтереси:

- застосування методів теорії ухвалення рішень при управлінні складними телекомунікаційними системами, паралельні обчислення.

МІРОШНИК Марина Анатоліївна– д.т.н., професор; професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна; м.Харків, майдан Свободи, 6, 61022; e-mail: m.miroshnyk@karazin.ua, ORCID 0000-0002-2231-2529

Наукові інтереси:

- комп'ютерні системи та компоненти, розподілені обчислення, Інтернет речей.

УРСУ Р.Є., ЛАБЕНКО Д.П.

МОДЕЛЬ УПРАВЛІННЯ ЯКІСТЮ ПРОДУКЦІЇ

Today, one of the most important economic parameters that determine the market strategy of product positioning is the price / quality ratio. At the same time, the lower limit of the sales price of products is determined by its cost price and can be determined on the basis of functional and cost analysis. To evaluate the "quality" parameter, an approach based on "quality management" is considered.

The object of research is the process of Product Quality Management.

The subject of the research is Methods and tools of Product Quality Management.

The goal is to create a product quality management model.

The paper considers modern technologies for creating an automated control system. The optimal technologies for creating this automated control system that are most acceptable for the subject area considered in the paper are analyzed and selected.

The result is an ask that works with the characteristics of the model, which are formed on the basis of the enterprise and improve the quality of products.

Scope of application – automation of processes in enterprises, optimization of work at the enterprise, and so on.

Ключові слова: модель, SCADA, АСК, якість, ПЛК, підприємство, HMI, желатин, варка, котел.

Методи розробки моделі

Розглянемо найбільш вживані методи розробки моделі.

Можна виділити 5 основних методів розробки:

- За допомогою платформи Rapid SCADA;
- За допомогою додатка Адміністратор(ScadaAdmin);
- За допомогою додатків які підтримує Scada;
- З використанням среди програмування CoDeSys;
- За допомогою протоколу Modbus.

Структура SCADA-систем

Будь-яка SCADA-система складається із трьох компонентів: віддалений термінал (RTU), диспетчерський пункт управління (MTU) і комунікаційна система (CS) (рис.1).

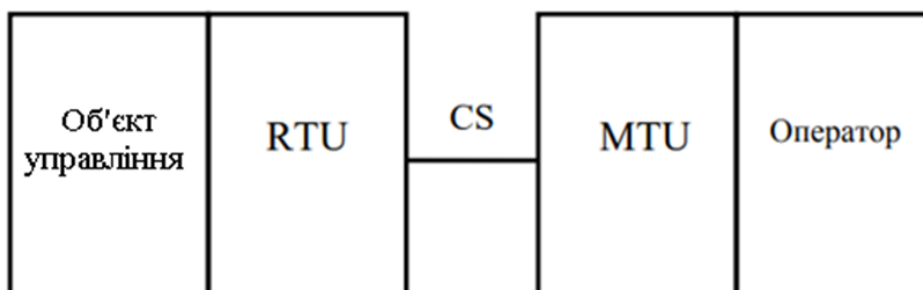


Рис. 1 Основні структурні компоненти SCADA-системи

Віддалений термінал підключається безпосередньо до контролюваного об'єкту і здійснює управління в режимі реального часу. Таким терміналом може служити як примітивний датчик, який здійснює знімання інформації з об'єкта, так і спеціалізований багатопроцесорний відмовостійкий обчислювальний комплекс, який здійснює обробку інформації та управління в режимі реального часу.

Диспетчерський пункт управління здійснює обробку даних і управління високого рівня, як правило, в режимі квазіреальності часу. Він забезпечує людино-машинний інтерфейс. MTU може бути як одиночним комп'ютером з додатковими пристроями підключення до каналів зв'язку, так і великою обчислювальною системою або локальною мережею робочих станцій і серверів.

Комунікаційна система необхідна для передачі даних з RTU на MTU і назад. Як комунікаційні системи можуть використовуватися такі канали передачі даних: виділені лінії,

радіомережі, аналогові телефонні лінії, ISDN мережі, стільникові мережі GSM (GPRS). Найчастіше пристрої підключаються до декількох мереж для забезпечення надійності передачі даних.

Аналіз предметної області

Процес виробництва желатину полягає в екстракції білка з сировини, що містить колаген. Сировиною для виробництва желатину є білкові відходи від шкур тварин. Сировина утворюється на шкіряних заводах, також є відходами шкіряного виробництва, проте воно є дуже цінним матеріалом, оскільки воно містить в собі білок. Коли воно потрапляє на желатиновий завод воно проходить кілька етапів виробництва.

Першим етапом виробництва є варіння. Процес варіння включає в себе поетапний забір білка з сировини. Сировина транспортується в варильні котли. У виручених котлах поетапно зливається до 8 бульйонів. Процес варіння відбувається в котлах, додається тепла вода, після наростання концентрації білка в бульйоні, бульйони зливаються і йдуть на наступний етап виробництва. Сировина залишається в котлі і відбувається кілька заварювань (зливів). Температура варіння на перших котлах нижча, близько - 50-60 градусів. В останніх котлах температура для того, щоб забрати білок досягає вже 90 градусів. Більш якісний желатин виходить в перших бульйонах, де температура нижче.

Другим етапом виробництва є фільтрація. З варильного відділення бульйони надходять на відділення фільтрації. На відділення фільтрації знаходяться кілька наливних фільтрів в яких використовуються дизель ГУР. Далі желатин надходить на ультрафільтрації. Головне завдання цього етапу довести концентрацію бульйону, від 3-4%, до 20%. Бульйон з такою концентрацією далі надходить на наступний етап-сушильне відділення.

Сушка желатину відбувається в тунельних сушильних апаратах. Цей сушильний апарат розбитий на кілька зон, на кожній зоні - різна температура. У підсумку після цього виходить желатин з вологістю близько 14%. Після сушильного відділення желатин пневмотранспортом подається на дроблення. Це останній етап. Желатин дробиться на декількох дробарках. Потім подається на розсіювач по-фракціям, і в залежності від бажання споживачів, желатин розфасовують в мішки.

На кожному етапі виробництва відбирають проби і роблять лабораторні аналізи. Саме готовий продукт обирають після етапу сушки. Потім за показниками в'язкості, міцності, прозорості холодцю, роблять висновки і формують готовий продукт. Залежно від вимог споживача, формують партії і відправляють готову продукцію.

Архітектура моделі

Головна сторінка складається з наступних компонентів(рис.2):

- Головні теги (найменування, номери котлів, тощо) ;
- Поле для коментаря на E-mail;
- Значення температури, пару.

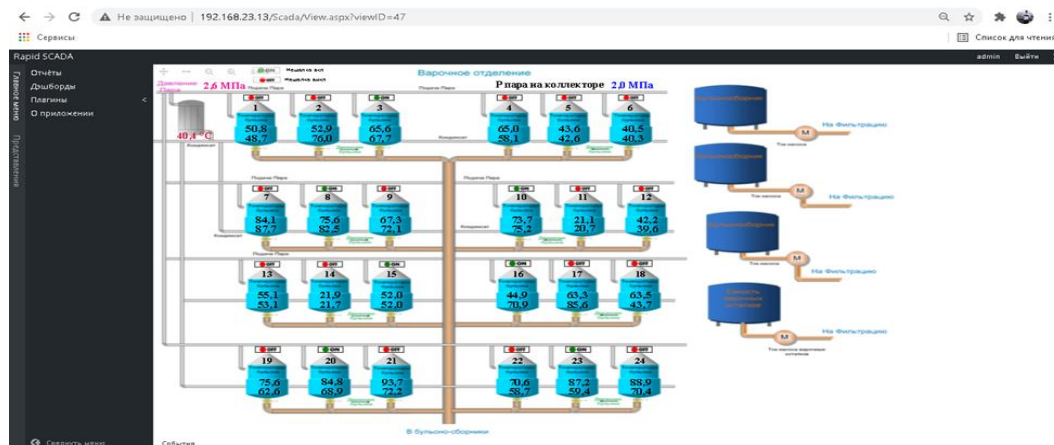


Рис. 2 Головна сторінка

Характеристика роботи моделі

У розробленій моделі представлений один з етапів виробництва желатину - варильне відділення. На цьому етапі сировину вариться в чанах, і віддає весь свій білок в бульйон. На цьому етапі важливо розуміти температуру варіння, так само тиск пара, яке нагнітає цю температуру, оскільки це впливає на якість желатину. Якісний желатин виходить, коли температура низька - 50-60 градусів. Особа, яка приймає рішення (ОПР) може моніторити дані в реальному часі і робити висновки, давати вказівки для поліпшення якості продукції. Так само ЛПР може робити висновки, виходячи з графіків, аналізувати значення в різний час, робити висновки, що пішло не так, коли продукт вийшов неякісним. Коли температура, або тиск досягне критичних значень, ЛПР при прийде повідомлення, що щось не так. Так само він зможе переглянути події, які відбувалися протягом робочої зміни, оскільки желатинове виробництво - це безперервне виробництво. Людина - не робот і не зможе контролювати все відразу, в цьому йому і допоможе модель. Так само співробітникам це буде корисно вони зможуть також моніторити дані про температуру, чи працює взагалі котел, чи немає якихось збоїв та інше. Так само вони можуть залишати коментарі, що сталося, наприклад, вихід з ладу котла, нестача температури, збій електрики. Ця схема допомагає скоротити людські помилки, що і дозволить поліпшити якість продукції.

ЛІТЕРАТУРА

1. Дружинин Г.В. Методы оценки и прогнозирования качества. - М.: Радио и связь, 1982. 160 с.
2. Богатырев А.А., Филиппов Ю.Д. Стандартизация статистических методов управления качеством. - М.: Изд-во стандартов, 1989. 120 с.
3. Статистические методы повышения качества / Под ред. Хитова К.И. - М.: Финансы и статистика, 1990. 304 с.
4. А.Ф.Джафаров Производство желатина. -М.:Агропромиздат, 1990.287с.
5. Д.И.Вирник,А.П.Власов,Д.З.Таланцев,З.В.Хохлова Технология клея и желатина,1963.472с.
6. Сучасні концепції побудови АСУ ТП на основі SCADA- систем. Використання деяких SCADA-систем в газовій галузі Scada системи АСУ ТП,2016.URL: <https://polarize.ru/uk/graphics/sovremennye-konceptii-postroeniya-asu-tp-na-osnove-scada-sistem/> (Last accessed: 12.11.2022).
7. Що таке скаду система в АСУ ТП. Основні функції SCADA-систем. Загальна структура SCADA,2018.URL: <https://autovideoparser.ru/uk/chto-takoe-skada-sistema-v-asu-tp-osnovnye-funkcii-scada-sistem-obshchaya/> (Last accessed: 14.11.2022).

УРСУ Роман Єдуардович – бакалавр; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: xa12284078@student.karazin.ua; ORCID: 0000-0003-4281-4267.

Наукові інтереси:

– АСУ ТП.

ЛАБЕНКО Дмитро Петрович – к. т. н., доцент; доцент кафедри ТПС; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: d.labenko@karazin.ua; ORCID: .

Наукові інтереси:

– веб-розробка, програмна інженерія, АСУ ТП.

ХРУСЛОВ М. М., КІПРЕНКО Є. В.

РОЗРОБКА ТА РЕАЛІЗАЦІЯ АЛГОРИТМІВ АНАЛІЗУ ТА ОБРОБКИ ДАНИХ НАТУРНИХ ТА ЧИСЕЛЬНИХ ЕКСПЕРИМЕНТІВ З ВИКОРИСТАННЯМ СУЧАСНИХ БАЗ ДАНИХ

Ключові слова: алгоритм, аналіз даних, обробка даних, база даних, натурні та чисельні експерименти, програмне забезпечення.

Тема роботи

В епоху глобальної цифровізації та автоматизації дані є ключовим елементом програмних систем. Отримані різними шляхами, при правильних обробці та аналізі вони стають надзвичайно корисними та допомагають краще розуміти предметну область та приймати рішення. Сам по собі збір даних не є складним процесом (хоча все ж таки це потребує правильного підходу та дотримання певних правил). Найбільш ресурсно затратними та важливими частинами при роботі з даними є їх обробка та аналіз. Бо саме за результатами цих процесів дані набувають цінності та можуть представляти користь. Тож для деяких галузей правильні та ефективні обробка та аналіз даних є критично важливими. Однією з таких галузей є експериментальні дослідження, даними в яких є результати проведення натурних та чисельних експериментів. Одним із варіантів для пришвидшення та підвищення ефективності обробки та аналізу цих результатів може бути застосування нового підходу — залучення до процесу сучасних баз даних, а саме NoSQL рішень. Це може надати наступні переваги для вирішення проблеми обробки та аналізу даних в цій галузі:

1. Зберігання, обробка та аналіз отриманих результатів будуть проводитись в єдиному місці.
2. Дані не обов'язково мають бути представлені у вигляді таблиць, що додає гнучкості для їх зберігання.
3. Значна частина операцій при аналізі та обробці даних може бути зроблена вбудованими інструментами самої бази даних.
4. Можлива паралельна обробка та аналіз даних кількома користувачами.
5. Можливе зберігання результатів на віддаленому сервері без потреби мати локальну копію.

Для більшої зручності всього процесу окрім бази даних треба залучення також і програмного додатку, який буде надавати інтерфейс користувача з можливістю завантаження, вивантаження та візуалізації даних, а також базовими процедурами і функціями. Використання інтеграції програмного додатку та бази даних надасть можливість розширювати вибір операцій над даними, а також зменшить вхідний поріг для користувачів-початківців.

Мета роботи

Для вирішення задачі підвищення ефективності та спрощення процесу обробки та аналізу результатів натурних та чисельних експериментів має бути спроектовано та створене таке програмне рішення, яке використовуючи сучасну базу даних надасть користувачеві простий та зрозумілий інтерфейс для проведення операцій над завантаженими даними, що в свою чергу має привести до підвищення якості та швидкості обробки та аналізу. Крім того, наявність бази даних надаватиме більшої надійності при зберіганні даних, можливість централізованого доступу до них, а також прямі операції над даними засобами самої бази даних.

Підсумовуючи вище сказане, метою роботи є аналіз, дослідження та використання технологій проектування програмного забезпечення при створенні програмного рішення з використанням сучасної NoSQL бази даних для ефективної та швидкої обробки результатів натурних та чисельних експериментів.

Методи дослідження

Відповідно до поставленої задачі було проведено дослідження мов програмування високого рівня та відповідних середовищ програмування, що забезпечують швидкість

проведення математичних операцій, надають можливість створення зручного інтерфейсу користувача та програмні додатки яких можуть автономно працювати на будь-якому ПК незалежно від операційної системи. Після вибору мови програмування та програмного середовища було здійснено аргументований вибір NoSQL бази даних для інтеграції у відповідний програмний додаток. Також було обрано необхідні технології для власне реалізації інтеграції між програмним додатком та обраною базою даних.

Результати

Результатом проведеної роботи є написання додатку (програми) на мові програмування Java версії 17 та інтеграція в нього NoSQL бази даних MongoDB. Даний додаток дозволяє проводити обробку та аналіз результатів натурних та чисельних експериментів, надаючи користувачеві інтерфейс, який було розроблено за допомогою мікро фреймворку для створення web-додатків Spark та бібліотеки мови програмування JavaScript під назвою React.

Таким чином даний програмний додаток може бути запущено локальному ПК користувача та/або на віддаленому сервері з будь-якою операційною системою.

Надалі планується розвиток даного програмного забезпечення шляхом розширення його функціоналу та підтримки форматів вхідних даних.

Висновки

Під час дослідження було розроблено та реалізовано програмний додаток для обробки та аналізу результатів натурних та чисельних експериментів із залученням сучасної NoSQL бази даних MongoDB. Додаток дозволяє швидко та ефективно оброблювати, аналізувати та зберігати дані, отримані в результаті вище згаданих експериментів, надаючи сучасний інтерфейс користувача та вирішуючи низку вже згаданих проблем. До повного списку основних використаних для реалізації технологій належать: Java, Spark, React та MongoDB.

ЛІТЕРАТУРА

1. Lafore R. Data Structures and Algorithms in Java / Robert Lafore., 2017. – (2nd Edition). – (978-0672324536).
2. Sadalage P. NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence / P. Sadalage, M. Fowler., 2012. – 192 с. – (1st Edition). – (978-0321826626).
3. Bradshaw S. MongoDB: The Definitive Guide: Powerful and Scalable Data Storage / Shannon Bradshaw.. – 514 с. – (3rd Edition). – (978-1491954461).
4. MongoDB Fundamentals: A hands-on guide to using MongoDB and Atlas in the real world / A. Phaltankar, J. Ahsan, M. Harrison, L. Nedov., 2020. – 748 с. – (978-1839210648).
5. Schildt H. Java: The Complete Reference, Twelfth Edition / Herbert Schildt., 2021. – 1280 с. – (978-1260463415).
6. Bloch J. Effective Java / Joshua Bloch., 2017. – 416 с. – (3rd Edition). – (978-0134685991).
7. Java Platform, Standard Edition & Java Development Kit Version 17 API Specification [Електронний ресурс] – Режим доступу до ресурсу: <https://docs.oracle.com/en/java/javase/17/docs/api/index.html>.
8. MongoDB Documentation [Електронний ресурс] – Режим доступу до ресурсу: <https://www.mongodb.com/docs/>.

КІПРЕНКО Євгеній Валентинович – студент кафедри моделювання систем і технологій; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, Харків-22, Україна, 61022; e-mail: xa11867742@student.karazin.ua; zhenyakiiprenko@gmail.com; ORCID: 0000-0002-8004-2299.

ХРУСЛОВ Максим Михайлович — кандидат фізико-математичних наук, доцент кафедри електроніки і управляючих систем; Харківський національний університет імені В.Н. Каразіна; майдан Свободи, 4, місто Харків, 61022, Україна; e-mail: maksym.khruslov@karazin.ua; <https://orcid.org/0000-0001-9639-9340>.

Наукові інтереси:

– Аналіз даних, Машинне навчання, Інтернет речей.

ХРУСЛОВ М.М., ЗОЛОТУХІНА О.В., ВЕКЛИЧ Д.А.

ПРОГРАМНЕ РІШЕННЯ ОПТИМІЗАЦІЇ УПРАВЛІННЯ ПЕРСОНАЛОМ НА ПРИКЛАДІ РЕСТОРАННОГО БІЗНЕСУ

Тема роботи

Питання ефективного менеджменту в ресторанному бізнесі на тлі величезної конкуренції є одним з принципових. Завдяки розвитку технологій та постійному оновленню науково-дослідницьких робіт у менеджменті, також зростає попит на змінення звичайних парадигм управління на нові, та вдосконалення можливостей їх автоматизації і реалізації за допомогою програмних рішень. Тож, для оптимізації управління персоналом у ресторанному бізнесі має бути застосоване програмне рішення. З основних проблем які можна виділити під час організації управління персоналом за допомогою програмних продуктів слід зауважити наступні:

1. Відсутність єдиного, масштабного формату для організації розкладу змін робочого персоналу на підприємствах громадського харчування.
2. Не своєчасне сповіщення персоналу про створення і зміни в робочому розкладі.
3. Відсутність інтегрованого з системою складання розкладів засобу підрахунку продуктивності персоналу в ресторані.

Для продуктивної роботи менеджер має використовувати програмний продукт, який завдяки автоматизації, буде допомагати йому створювати розклади змін для персоналу з урахуванням їх заохочень. Більш того програмне рішення має своєчасно інформувати менеджера про заохочення персоналу щодо робочого розкладу, в свою чергу робітники мають отримувати миттєву інформацію щодо змін у робочому розкладі з боку менеджера. Додатково, у програмний продукт має бути інтегровано систему для оцінки плідності роботи співробітників, з можливістю перегляду оптимальності роботи персоналу, як з боку менеджера так і з боку самих працівників.

Мета роботи

Для досягнення цілі в оптимізації управління персоналом має бути розроблене таке програмне рішення в результаті інтегрування якого має з'явитись можливість підвищення ефективності роботи співробітників. А саме завдяки автоматизації процесів та менших витрат фізичної роботи менеджера, що у свою чергу дозволить йому більше сфокусуватись на інших цінних задачах для підвищення мотивації та кваліфікації працівників. Більш того завдяки цьому програмному продукту у менеджера та працівників має з'явитися можливість в реальному часі переглядати показники ефективності своїх колег або підлеглих. На основі цієї інформації у менеджера буде можливість подавати звіти про продуктивну або не продуктивну працю робітників до вищої ланки менеджменту на основі КРІ (ключові показники ефективності) розрахунків. Це в свою чергу допоможе відрізнити ефективні впровадження з боку менеджера від не ефективних.

Тож метою роботи є аналіз, дослідження та використання технологій проектування програмного забезпечення при створенні програмного рішення здатного оптимізувати процес управління персоналом на прикладі ресторанного бізнесу, з використанням оцінки ефективності персоналу за КРІ показниками.

Методи дослідження

Основним методом дослідження є розрахунок ефективності роботи працівника на основі визначених показників роботи (середня сума чеку, кількість проданих додаткових сервісів, рівень професійної обізнаності, кількість позитивних відгуків). Менеджер має присвоїти вагу показникам та розподілити за ступенем їх важливості. Показники мають бути чітко прописані, а працівник проінформований про ступінь важливості кожного. Після цього задається чисельний коефіцієнт до відсотка фактично виконаної роботи працівником по відношенню до плану. На базі цих даних і розраховується ефективність роботи працівника.

За допомогою технологій проектування програмного забезпечення, а саме побудови компонентних програмних рішень (КПР) було створено алгоритм побудови програмного рішення здатного оптимізувати процес управління персоналом на прикладі ресторанного бізнесу, з використанням оцінки ефективності персоналу за КРІ показниками. Алгоритм

побудови програмного рішення запропонований з використання наступних технологій Javascript, React, NodeJS, PostgreSQL та Telegram Bot API.

Розрахунок та аналіз результатів

Для розрахунку ефективності роботи працівника треба провести оцінювання результатів роботи і чинників, що визначають ступінь його досягнення; аналіз динаміки результативності праці за певний період, а також динаміки стану чинників, що впливають на досягнення результатів [5]. Показники розподіляють за значимістю та кожному привласнюють вагове значення. Сума ваг за всіма показниками роботи працівника дорівнює одиниці (100%). Приклад розподілення таких показників ефективності наведено у Таблиці 1.

Табл.1 Приклад показників ефективності та їх вагомості

Показник	Вагомість	План	Факт	Індекс KPI, %
Сума продаж	V_1	P_1	F_1	K_1
Середня сума чеку	V_2	P_2	F_2	K_2
Кількість проданих додаткових сервісів	V_3	P_3	F_3	K_3
Рівень професійної обізнаності	V_4	P_4	F_4	K_4
Кількість позитивних відгуків	V_5	P_5	F_5	K_5

Показник індекс KPI (K_i) є відношення виконаної роботи (F_i) до показників плану (P_i), розраховуються за наступною формулою:

$$K_i = \left(\frac{F}{P}\right)_i \cdot 100\%, \quad (1)$$

Додатково мають бути розподілені кількісні коефіцієнти виконаного плану. Приклад розподілення таких коефіцієнтів наведено у Таблиці 2.

Табл.2 Приклад коефіцієнтів виконаного плану працівника

Виконаний план, %	Коефіцієнт виконаного плану
< 80	0
80 - 89	0,85
90 - 99	0,9
100 - 109	1
109 - 119	1,05
> 119	1,2

На основі цих даних розраховується ефективності працівника за наступною формулою:

$$E = \sum_{i=1}^n (K \cdot V)_i \cdot N, \quad (2)$$

де, K – KPI показник;

V – вагомість KPI показника;

N – коефіцієнт по виконаному плану, $N \neq const$;

n – кількість KPI показників.

Висновки

В результаті дослідження було запропоноване алгоритм побудови програмного рішення для підвищення ефективності управління персоналом на основі ресторанного бізнесу. Цей програмний продукт має дозволити проводити миттєві розрахунки ефективності роботи персоналу за обраний термін часу. Був виконаний аналіз та вивчення існуючих аналогів та методів оцінки ефективності персоналу за багатьма критеріями, з яких були відокремлені найважливіші параметри для оцінки ефективності. Для програмної реалізації обрані наступні технології: React, NodeJS, Telegram Bot API, PostgreSQL. На основі отриманих даних було створено алгоритм побудови програмного рішення, з урахуванням специфіки ресторанного бізнесу, що має підвищити ефективність менеджменту персоналу ресторанного бізнесу.

Ключові слова: програмне рішення, ресторанний бізнес, оптимізація управління, управління персоналом, автоматизація, KPI показники, робочі розклади, показники ефективності.

ЛІТЕРАТУРА

1. Оцінка персоналу [Електронний ресурс]. – Режим доступу: http://4ua.co.ua/management/qa3bc79a5c43a89521316d27_0.html. Дата звернення: 16.09.2022.
2. Автоматизація бізнес-процесів [Електронний ресурс]. – Режим доступу: <https://www.terrasoft.ua/page/business-process-automation/>. Дата звернення: 10.09.2022.
3. Крилова Л. В., Шумілова В. В. Розробка системи оцінки професійного рівня обслуговуючого персоналу [Електронний ресурс]. – Режим доступу: http://www.rusnauka.com/2_ANR_2010/Economics/5_53754.doc.htm. Дата звернення: 10.09.2022.
4. Мокряк Е.В., Бассано А. Ю. Формування системи стимулювання персоналу на основі ключових показників ефективності [Електронний ресурс]. – Режим доступу: <http://www.economy-confer.com.ua/full-article/3106/>. Дата звернення: 10.09.2022.
5. Смоквіна Г.А., Альхаммаді Талеб Абдулла Мохаммед Алі Інструменти оцінки кадрового складника системи фінансово-економічної безпеки підприємства [Електронний ресурс]. – Режим доступу: <http://www.ei-journal.in.ua/index.php/journal/article/view/438/425>. Дата звернення: 10.09.2022.
6. A Web-Based System for Evaluating SME Employees [Електронний ресурс]. – Режим доступу: <https://su-plus.strathmore.edu/bitstream/handle/11071/12845/employees.pdf>
7. ХАРАКТЕРИСТИКА МЕТОДІВ МЕНЕДЖМЕНТУ В РЕСТОРАННОМУ БІЗНЕСІ [Електронний ресурс]. – Режим доступу: <http://restaurant-hotel.knukim.edu.ua/article/view/157174/158460>. Дата звернення: 11.09.2022.
8. ДСТУ ISO 9001: 2001. Системи управління якістю. [Чинний від 2001-06-27]. Київ, 2001. 24с.
9. Діаграми компонентів: [Електронний ресурс]. // URL: https://uk.wikipedia.org/wiki/Діаграма_компонентів. Дата звернення: 20.08.2022.
10. React - JavaScript-бібліотека для створення клієнтських інтерфейсів. [Електронний ресурс] <https://reactjs.org/docs/getting-started.html>. Дата звернення : 26.09.2022.
11. Онлайн-довідка з PostgreSQL. [Електронний ресурс] <https://www.postgresql.org/> . Дата звернення : 25.09.2022.
12. Software quality attributes [Електронний ресурс]. – Режим доступу: <https://link.springer.com/chapter>. Дата звернення: 28.10.2022.
13. Software testing: [Електронний ресурс]. // URL: <http://worldcolleges.info/sites/default/files/software-testing-testing-across-the-entire-software-development-life-cycle> Дата звернення: 15.09.2022.
14. Разработка и сопровождение проблемно-ориентированных ПС / проф. Н.В. Ткачук // каф. МСТ, ХНУ імені В.Н. Каразіна. 2019-2020.
15. Онлайн-довідка з NodeJS: [Електронний ресурс]. // URL: <https://nodejs.org/uk/docs/>. Дата звернення: 15.09.2022.
16. Онлайн-довідка з ExpressJS: [Електронний ресурс]. // URL: <http://expressjs.com/en/api.html>. Дата звернення: 15.09.2022.

ХРУСЛОВ Максим Михайлович – к. ф.-м.наук, доцент кафедри електроніки і управляючих систем ХНУ імені В.Н.Каразіна; e-mail: maksym.khruslov@karazin.ua ; <https://orcid.org/0000-0001-9639-9340>

Наукові інтереси:

– Аналіз даних, Машинне навчання, Інтернет речей.

ЗОЛОТУХІНА Ольга Вікторівна – ст. викл. кафедри моделювання систем і технологій; ХНУ імені В.Н.Каразіна; e-mail: ovzolutukhina@karazin.ua; <https://orcid.org/0000-0002-9354-7018>

Наукові інтереси:

– Аналіз даних, Машинне навчання, Інтернет речей, Економіка ІТ-галузі, Менеджмент.

ВЕКЛИЧ Денис Андрійович - студент кафедри моделювання систем і технологій; ХНУ імені В.Н.Каразіна; e-mail: deniska.veklitch@gmail.com; <https://orcid.org/0000-0002-0073-099X>.

Наукові інтереси: – Аналіз даних, Машинне навчання, Інтернет речей.

МАШИННЕ НАВЧАННЯ В ЗАДАЧАХ АНАЛІЗУ І ВДОСКОНАЛЕННЯ ВЕБ-САЙТІВ

Мета роботи: дослідження задач удосконалення роботи веб-сайтів за допомогою методів машинного навчання.

Методи дослідження: у роботі розглянутий приклад гібридної моделі для оцінювання веб-сайтів авіакомпаній. Запропонована гібридна модель поєднує метод ентропійної ваги та реляційний аналіз Грея для визначення та оцінки ефективності веб-сайтів авіакомпаній. У якості ознак для моделі було обрано показники продуктивності, оптимізація дизайну, час завантаження, час відгуку, розмітку та пошкоджені посилання тощо.

Задача оцінки якості веб-сайту з точки зору машинного навчання є задачею регресії. І для її розв'язання застосовуються моделі лінійної регресії та машина опорних векторів для регресії.

Результати: проведений огляд задач, які виникають при вдосконаленні або розробці веб-сайтів і які можна розв'язувати за допомогою методів машинного навчання. Розглянутий експеримент для порівняння алгоритмів машинного навчання при створенні моделі для оцінки ефективності веб-сайту.

Висновки: у роботі проаналізовані задачі аналізу і вдосконалення веб-сайтів, для яких доречним є впровадження запропонованих моделей та методів машинного навчання. Наведений приклад застосування моделей регресії для оцінки якості веб-сайтів авіакомпаній. У майбутніх дослідженнях доцільно розробити модель оцінки веб-сайтів із використанням інших моделей машинного навчання, таких як лінійні моделі з регуляризацією або ядрова регресія, для підвищення їх точності.

Ключові слова: лінійна регресія, машинна регресія опорних векторів, машинне навчання, веб-сайт.

Сучасній людині важко уявити своє життя без комп'ютеру, персональних пристроїв, смартфонів. Кожного дня ми користуємося засобами зв'язку (месенджерами, аудіо- або відео конференцій), пошуковими системами, соціальними мережами, веб-додатками тощо. Перед розробниками постає задача удосконалення роботи застосунків різного типу зокрема із застосуванням машинного навчання.

Машинне навчання або Machine Learning – різновид штучного інтелекту в основі якого лежить вивчення комп'ютерних алгоритмів, які автоматично навчаються в міру накопичення досвіду.

Із впровадженням алгоритмів машинного навчання до веб-застосунків, веб-сайтів можна вирішити такі задачі як:

- прискорення роботи веб-сайтів;
- прогнозування позиції сайту;
- створення системи рекомендацій для інтернет-магазину;
- оптимізація пошуку на сайті;
- оцінка ефективності веб-сайтів та ін.

Розглянемо ці задачі детальніше.

Прискорення роботи веб-сайтів

Багато людей постійно намагаються підвищити продуктивність веб-додатків: пробують збільшити швидкість на повільних телефонах, кешують всю можливу інформацію тощо.

Але таке уявлення про продуктивність веб-застосунків занадто обмеженим. З позиції користувача всі ці дії – лише малесенька частини загального поняття продуктивності. Задачу прискорення роботи веб-застосунків може вирішувати машинне навчання у багатьох випадках на різних сайтах.

Наприклад, візьмемо сайт, на якому одні користувачі можуть продавати речі, а інші – купити їх собі. Коли користувач реєструє новий товар для продажу, він обирає категорію, потім потрібний рекламний пакет, заповнює подробиці, переглядає оголошення, а вже потім його публікує.

Зупинимось на першому кроці – вибір категорії. По-перше, є багато категорій, і користувач може не зрозуміти відразу, яку з них обрати. По-друге, навіть коли стає зрозуміло, до якої категорії/підкатегорії належить товар, процес все одно займає близько 12 секунд.

І тут виникає задача, що якщо заголовок, опис та ціну товару передати правильно навченій моделі машинного навчання, то вона визначить, до якої категорії належить товар. Отже замість того, щоб витрачати час на вибір категорії, користувач міг би витратити 12 додаткових секунд на реєстрацію інших товарів.

Прогнозування позицій веб-сайту

Ця задача належить до задач машинного навчання. Для побудови моделей використовують дані інформаційних сайтів.

Наприклад, є набір даних, який містить інформацію про 9751 URL. В якості характеристик використовують: час відповіді сервера, назва (title), довжина назви, опис (description), довжина опису, вміст заголовків першого рівня (H1), довжина H1, розмір контенту, співвідношення Text/HTML, кількість слів, довжина контенту.

Також важливим є тип пошукового запиту, чи є він низькочастотним чи ні. Коефіцієнти входження ключового слова до запиту важливі лише у випадку, якщо це середньочастотний або високочастотний запит, а сама сторінка має високий показник відповіді сервера.

Використовуючи алгоритми машинне навчання, можна побудувати модель прогнозу ймовірності позиції у списку, який надає Google за запитом, і проаналізувати як на цей прогноз вплине зміна коефіцієнта входження ключового слова в title, description та H1.

Система рекомендацій інтернет-магазину

На ринку інтернет-торгівлі склалася така ситуація: зріс загальний грошовий потік і збільшилася кількість продавців. Це призвело до того, що частка кожного магазину зменшилася, а конкуренція між стає дедалі напруженішою. Один із способів збільшити середній розмір купівлі (а отже, і прибуток) – пропонувати покупцям додаткові товари, які можуть їх зацікавити.

Для підбору рекомендацій у реальному часі на сайті необхідні вихідні дані. Якщо магазину ще невідомі уподобання користувача, можна підбирати рекомендації просто на основі обраних ним пропозицій. При цьому система повинна постійно навчатися та накопичувати дані про те, що подобається клієнтам. І саме за цією інформацією можна буде проводити аналіз та підбирати актуальні рекомендації за допомогою системи, заснованій на моделях машинного навчання. Крім того, такій системі можна передавати інформацію про інших користувачів, а також час від часу перенавчати.

Оптимізація пошуку на сайті

Коли людина відвідує веб-сайт, вона свідомо і несвідомо висловлює свій намір здійснити або запланувати покупку, надаючи деяку кількість інформації про свої інтереси. Наприклад, може бути введений термін у полі пошуку, перейти за посиланням на опис продукту або додати чи видалити елементи з кошика покупок.

Розуміння пошукового запиту та подальше надання відвідувачеві релевантних продуктів має вирішальне значення для задоволеності клієнтів. Наприклад, більшість пошукових систем не розуміють мови, а просто обробляють слова. Результат: покупцеві, який шукає сукню-сорочку, буде представлено цей тип одягу, але його також можуть зацікавити результати, що показують класичні сорочки – зовсім інший одяг. Ця нездатність зрозуміти наміри клієнтів призведе до незадовільного досвіду клієнта.

В вирішенні цієї проблеми може допомогти машинне навчання. Адже машинне навчання дозволяє застосувати алгоритми глибоко розуміння мови. Це означає, що пошукова система постійно удосконалюватиметься, зіставляючи пошуковий запит з поведінкою на сайті і відповідним чином оптимізуючи пошук. При цьому пошукова система зможе пропонувати правильний продукт кожному відвідувачу, навіть якщо можливих комбінацій відвідувач/продукт є незліченна кількість [2].

Оцінка ефективності веб-сайтів

Дуже важливим в процесі роботи веб-сайтів, особливо комерційних, є оцінка їх ефективності. Під оцінкою ефективності розуміють правило, за яким обрані показники ефективності порівнюють між собою або з деякою нормою, якщо вона існує або її можна встановити. Серед критеріїв оцінки ефективності веб-сайтів є часто вживаними такі типи критеріїв:

- технічні критерії швидкодії, оптимізації мережевого трафіку та вимог до апаратних ресурсів;
- технічні критерії надійності та безпеки й ефективності технічної підтримки;
- естетичні та художні критерії;
- психологічні критерії;
- системні критерії глобальних середовищ (відвідуваність сайту, рейтинг сайту та ін.).

Існує ряд досліджень, в яких пропонується підхід до оцінки ефективності веб-сайтів за допомогою машинного навчання.

У роботі [1] розглянута гібридна модель для оцінювання веб-сайтів авіакомпаній. Запропонована гібридну модель поєднує метод ентропійної ваги та реляційний аналіз Грея для визначення та оцінки ефективності веб-сайтів авіакомпаній на вибірці з одинадцяти веб-сайтів авіакомпаній. У якості ознак для моделі було обрано показники продуктивності, оптимізація дизайну, час завантаження, час відгуку, розмітку та пошкоджені посилання тощо.

Дана задача з точки зору машинного навчання є задачею регресії. І для її розв'язання застосовують моделі лінійної регресії та машину опорних векторів для регресії.

Модель лінійної регресії – це статистичний підхід до побудови лінійної моделі, яка передбачає значення метрики (цільової функції), виходячи із значень інших метрик (параметрів). Лінійна регресія використовує метод найменших середніх квадратів для коригування параметрів моделі.

Регресійна модель опорних векторів – алгоритм створює опорні вектори в області ознак великої розмірності. Потім будується гіперплощина з максимальним запасом. Функція ядра використовується для перетворення даних, що збільшує розмірність даних. Це збільшення стимулює те, що дані можуть бути розділені гіперплощиною з набагато вищою ймовірністю, і встановити мінімальну міру помилки ймовірності прогнозування.

Вагові коефіцієнти показників ефективності розраховували за допомогою методів регресії, а потім оцінювали продуктивність веб-сайтів.

Існують критерії, за якими моделі можна оцінити та порівняти з точки зору кращої якості (правильності). Серед них найчастіше використовують:

- середню абсолютну похибку – середнє абсолютне відхилення прогнозу від фактичного значення (виражається в продуктивності веб-сайту);
- середню відносну похибку - обчислюється як середнє прогнозу, який бачить у чисельнику похибку в абсолютному значенні серед прогнозованих значень і відповідних реальних значень, а знаменник — реальне значення (виражається у відсотках);
- кореляцію: забезпечує відсоткове значення кореляції між прогнозованими та фактичними значеннями в діапазоні від 0 до 100, де 100 представляє ідеальний прогноз даних моделлю (виражається у відсотках).

Авторами дослідження були проведені експерименти для порівняння алгоритмів машинного навчання для створення моделі для оцінки ефективності веб-сайту. Для оцінки продуктивності моделей були використані метрики на основі кореляції, середньої абсолютної помилки та середньої відносної помилки. Табл. 1 показує результати порівняння моделей. Кореляція в моделі лінійної регресії показує гарний прогноз – 71,5 % порівняно з кореляційною опорною векторною машиною 65,2 %. Лінійна регресія забезпечує найкращий результат із мінімальною середньою абсолютною похибкою 5,897 +/- 4,624 і мінімальною середньою відотною похибкою 9,64% +/- 7,66%.

Табл. 1 Результати порівняння моделей

Модель	Кореляція (%)	Середня абсолютна похибка	Середня відносна похибка
Множинна лінійна регресія	71,5%	5,897 +/- 4,624	9,64% +/- 7,66%
Машина опорних векторів	65,2%	6,993 +/- 5,277	11,72% +/- 9,75%

Висновки: для оцінки якості веб-сайтів авіакомпанії були проведені експерименти за двома алгоритмами, а саме лінійною регресією та машиною опорних векторів для регресії. У майбутніх дослідженнях можна створити модель оцінки веб-сайтів із використанням інших моделей машинного навчання регресії, таких як лінійні моделі з регуляризациєю або ядрової регресії, для подальшого підвищення точності моделей.

Технології машинного навчання вже стали частиною повсякденного життя, при цьому кількість стартапів та продуктів на основі машинного навчання активно зростає. Будучи причиною технологічних революцій у деяких сферах економіки, ML може бути драйвером у масштабах бізнесу та держав. Сьогодні саме час замислитися про інтеграцію машинного навчання до бізнес-процесів, щоб не втратити конкурентоспроможність.

ЛІТЕРАТУРА

1. Mohammad G. Performance Evaluation of Websites Using Machine Learning. *Palestine Ahilya University*. 2020. No. 7. URL: https://www.researchgate.net/publication/345975296_Performance_Evaluation_of_Websites_Using_Machine_Learning. (Last accessed: 20.10.2022)

2. URL: <https://www.accenture.com/nl-en/blogs/insights/optimizing-on-site-search-using-machine-learning>

3. URL: <https://nachasi.com/tech/2019/01/31/yak-pratsyuye-machine-learning/>

ЧЕПАК Анастасія Альбертівна – студентка; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: anastasiia.chepak@student.karazin.ua; ORCID: 0000-0002-8223-194X.

Наукові інтереси:

– використання машинного навчання в сучасних технологіях.

СТРИЛЕЦЬ Вікторія Євгенівна – доцент кафедри теоретичної та прикладної системотехніки, Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 4, м. Харків, Україна, 61022; e-mail: viktorija.strilets@karazin.ua; ORCID: 0000-0002-2475-1496.

Наукові інтереси:

- моделювання складних процесів;
- моделі і методи машинного навчання;
- методи штучного інтелекту.

CHUB O.I., NOVOZHYLOVA M.V.

SOFTWARE CARPENTRY ENVIRONMENT AS A MODEL OF MODERN EDUCATIONAL AND RESEARCH COMMUNITY

Introduction. The modern level and speed of development of information technologies, the aspiration of modern society to informational unification necessitate the creation and implementation of a new model of the educational and scientific community interaction to ensure free, operational and effective access to the information resources being required for the purpose of implementing innovative means of studying and utilizing information processing tools. The educational and scientific community is looking for new approaches that can ensure the necessary level of education efficiency under the condition of severe limitations of resources, where time is the first of them.

Objective. Therefore, an urgent problem today is the formulation of the principles of construction, modelling, and systematization of the properties of the information environment as an interactive educational and scientific platform. In this work we consider the problem using the example of Software Carpentry.

Research methods. System analysis, set-theoretical analysis, multi-criteria optimization, expert information processing methods.

Results. To solve the formulated problem, we conducted an analysis of the paradigms of informational educational environments, classifying the most widespread and available in the Ukrainian educational and research community: Coursera, edX, Prometheus, Udemy. Another interesting direction is the so-called microlearning. At the same time, the use of the above types of educational online resources for the purposes of organizing and conducting scientific research, as well as selecting and implementing new tools for processing information about objects of scientific research is difficult or impossible. As one of the most successful implementations in this direction the ideology of Software Carpentry is considered. Authors are participating in the project implemented by University of St. Andrews (Scotland) on translation the «Python programming language» course from English to Ukrainian on a volunteer basis.

Conclusions. We propose generalized set-theoretical model of such an information environment for educational and research community interaction.

Keywords: modeling, online learning, open science, effectivity of education, resourcing

Introduction

The modern level and speed of development of information technologies, the aspiration of modern society to informational unification necessitate the creation and implementation of a new model of the educational and scientific community interaction to ensure free, operational and effective access to the information resources being required for the purpose of implementing innovative means of studying and utilizing information processing tools. This applies not only to non-formal education, or the implementation of the lifelong learning – LLL concept [1], so-called continuous education. The last three years, when the whole world was fighting the Corona Virus Disease 2019 (COVID-19) epidemic [2], and now the war in Ukraine – have changed the way of life of Ukrainians, the way of communication, thinking, vision of the future. Pandemic and war present drastic challenges to the education system including technical, cognitive, managerial, and behavioral ones. Under such external conditions, classical methods of teaching and scientific research cannot be fully implemented, thus the educational and scientific community is looking for new approaches that can ensure the necessary level of education efficiency under the condition of severe limitations of resources, where time is the first of them.

Objectives

Therefore, an urgent problem today is the formulation of the principles of construction, modelling, and systematization of the properties of the information environment as an interactive educational and scientific platform. In this work we consider the problem using the example of Software Carpentry [3].

Results

To solve the formulated problem, we conducted system analysis of the paradigms of informational educational environments, among which the following are currently the most widespread and available in the Ukrainian educational and research community: Coursera [4], edX [5], Prometheus [6], Udemy [7].

Coursera was launched in 2012 with a mission to provide universal access to world-class learning. It is now one of the largest online learning platforms in the world, with 113 million registered

learners as of September 30, 2022. Coursera partners with over 275 leading university and industry partners to offer a broad catalogue of content and credentials, including guided projects, courses, specializations, certificates, and bachelor's and master's degrees. Institutions around the world use Coursera to upskill and reskill their employees, citizens, and students in many high-demand areas, including data science, technology, and business.

Another platform edX is an American massive open online course provider created by Harvard and Massachusetts Institute of Technology. It hosts online university-level courses in a wide range of disciplines to a worldwide student body, including some courses at no charge. It also conducts research into learning based on how people use its platform. edX runs on the free Open edX open-source software platform. 2U is the parent company, with edX operating as its global online learning platform and primary brand for products and services. edX has 3600+ courses, 42 million users and 160+ partners.

Prometheus is the leading online education platform in Ukraine with over 700 000 online registered students and over 90 online courses. Company is working in close cooperation with top Ukrainian universities, leading professors, international organization, and government.

Prometheus obtained a license for Ukrainian translation and dubbing massive open online courses for the teachers by the Massachusetts Institute of Technology, Columbia Teachers College, and Queensland Universities. The certificate of completion is officially recognized by the Ministry of Education as an official teachers' training.

Udemy Inc. is a global destination for teaching and learning online. It was founded in May 2010 by Eren Bali, Gagan Biyani, and Oktay Caglar. As of July 2022, the platform has more than 54 million students, 204,000 courses, and 71,000 instructors teaching courses in over 75 languages. There have been over 741 million course enrolments.

Students take courses primarily to improve job-related skills. Some courses generate credit toward technical certification. Udemy has made a special effort to attract corporate trainers seeking to create coursework for employees of their company.

These are large repositories of a variety of well-prepared popular training courses that have meaningful content, are designed for several weeks and even months of training, support the classic distance learning paradigm and generally require significant resources from the user to master them.

Another interesting direction is the so-called microlearning [8]. Today, the traditional distance learning format, according to various estimates of scientists, is losing its effectiveness, the development of a learning paradigm, within the framework of which it is possible to dynamically update various, short-term training courses, is attracting considerable attention.

At the same time, the use of the above types of educational online resources for the purposes of organizing and conducting scientific research, as well as selecting and implementing new tools for processing information about objects of scientific research is difficult or impossible.

As one of the most successful implementations in this direction the ideology of Software Carpentry can be considered. Software Carpentry is a platform that teaches «basic laboratory skills for research computing» [3]. The mission of this project is to unite specialists in nature protection, geographers, ecologists, zoologists and other scientists. The courses of Software Carpentry are master classes that teach how to manipulate data and information and analyze it in an effective, reproducible and collaborative way. Software Carpentry teaches how to use complex modern information technologies (R, RStudio and GitHub), analyze scenarios, create convincing visualizations and learn joint work processes, without requiring previous experience.

Software Carpentry Foundation is a volunteer non-profit organization dedicated to teaching researchers basic computing skills since early.

Since 1998, Software Carpentry teaches researchers to create purpose-built tools, whether it be a Unix shell script to automate repetitive tasks or software code in programming languages such as Python, R, or MATLAB. These enable researchers to build programs that can be read, re-used and validated, greatly enhancing the sharing and reproducibility of their research.

All Carpentry workshops are hands-on two-day training events during which the attendees gain practical skills and understanding how particular software development tools and methodologies can benefit their own work. The core curriculum taught at Software Carpentry workshops typically includes:

1. automating tasks using the Unix shell;
2. structured programming in Python or R;
3. version control using Git or Mercurial.

The list is not exhaustive – all training materials are freely available under the Creative Commons – Attribution License from the Software Carpentry's lesson repository. Materials can be reused in any way you wish, without asking for special permission, provided that the original source is cited.

It should be noted that the authors have their own successful experience of learning and cooperating with this platform. The authors are participating in the project implemented by University of St. Andrews (Scotland) on translation the «Python programming language» course from English to Ukrainian on a volunteer basis. At the time of publication of this work, 80% of the course has already been translated (fig.1). And this work continues. After its completion, all stakeholders will have access to educational materials.

python-novice-gapminder

python-novice-gapminder

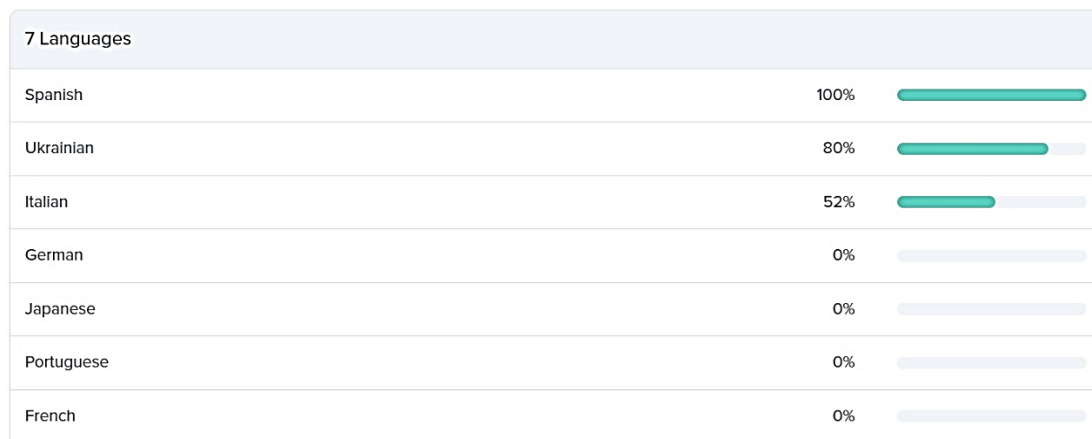


Fig. 1 General course translation statistics

Conclusions

As a generalized model of such an information environment for educational and research community interaction we can consider a system of a kind:

$$S = (< M, R>, P),$$

where M – is a set of elements of the information environment, R – is a set of connections between elements from M , P – is a set of system S efficiency criteria.

Each of the components of the system S has a hierarchical nature. Thus, a discrete set M of elements is a set of courses $\{M_1, \dots, M_N\}$, each of which, in turn, has a modular structure:

$$M_j = \{M_{j1}, \dots, M_{jK}\}, j=1, \dots, N.$$

The discrete set of properties P has the form:

- Training time, t ;
- Amount of acquired knowledge, V ;
- Possibility of further self-study, F ;
- Level of presentation, L :
 - L_1 – is an adequacy in terms of important concepts.
 - L_2 – is a possibility of direct application for the purposes of scientific research.
- Appearance of course, I :
 - I_1 – is an interactivity.
 - I_2 – is an ease of perception.
- Degree of openness (open source) of resources, D ;

- Possibility of refactoring, Re;
- Supporting open science paradigm [9], OS.

The formalized presentation of a set of properties P forms a set of partial efficiency criteria K: the scale of measurement of which in the general case must be established on the basis of expert knowledge processing. These partial criteria form a general criterion – the quality of education – in a world that is constantly changing.

REFERENCES

1. A Memorandum on Lifelong Learning // Commission of the European Communities. – URL: http://arhiv.acs.si/dokumenti/Memorandum_on_Lifelong_Learning.pdf. (date of request 19/11/2022).
2. Yuna Yao, Ping Wang, Yu JunJiang, Qiang Li, Yingji Li Innovative online learning strategies for the successful construction of student self-awareness during the COVID-19 pandemic: Merging TAM with TPB. Journal of Innovation & Knowledge. 2022. Vol.7. Issue 4. 100252.
3. Software Carpentry | Software Carpentry. URL: <https://software-carpentry.org/> (date of request 19/11/2022).
4. Coursera | Online Courses & Credentials from Top Educators. Join for Free. Coursera. URL: <https://www.coursera.org/> (date of request 19/11/2022).
5. edX | Free Online Courses by Harvard, MIT, & more. edX. URL: <https://www.edx.org/> (date of request 19/11/2022).
6. Prometheus | Prometheus – Найкращі онлайн-курси України та світу. URL: <https://prometheus.org.ua/> (дата звернення 19/11/2022).
7. Udemy | Online Courses – Learn Anything, On Your Schedule. URL: <https://www.udemy.com/> (date of request 19/11/2022).
8. Жмай О.В., Чепурна Л.В. Переваги мікронавчання в сучасних умовах online-освіти «Сучасна молодь в світі інформаційних технологій»: Матеріали II Всеукр. наук.-практ. інтернет-конф. м. Херсон, 14 травня 2021р. Херсон, 2021. С. 115 -116.
9. Open innovation, Open Science, open to the world. A vision for Europe. Brussels: European Commission, Directorate-General for Research and Innovation. 2016. 108 pp.

CHUB Olga Igorivna – PhD in Economics, Associate Professor of the Department of Theoretical and Applied System Engineering; V. N. Karazin Kharkiv National University, 6 Svobody Sq., Kharkiv, Ukraine, 61022;

PhD in Economics, Academic Visitor, School of Computer Sciences; University of St Andrews, College Gate, St Andrews, KY16 9AJ, Scotland

Main scientific interests are: logistic models, operations research, modelling of complex socio-economic and ecological systems

e-mail: o.i.chub@karazin.ua;

ORCID: 0000-0002-1216-856X.

NOVOZHYLOVA Maryna Volodymyrivna – Doctor of Sciences (Mathematics), Head of the Department of Computer Science and Information Technologies; O.M.Beketov National University of Urban Economy in Kharkiv, 17, Marshal Bazhanov Street, Kharkiv, Ukraine, 61002.

Main scientific interests are: mathematical modeling and solving multicriteria multidimensional resource planning constrained problems in time uncertainty conditions; supply chains; multicriteria nonlinear optimization problems of cutting, placement and packing, where items being placed have variable metric characteristics and spatial shape; optimization models and methods of resource management of development and redevelopment projects; nonlinear optimization, linearization technologies.

e-mail: marina.novozhilova@kname.edu.ua;

ORCID: 0000-0002-9977-7375.

ШЕВЧЕНКО Д.О., ЛИХАЧ О.Ю., УГРЮМОВ М.Л.

МЕТОДИ ВИЯВЛЕННЯ ВИКИДІВ В СИСТЕМАХ АВТОМАТИЗОВАНОГО ЗБОРУ ДАНИХ

Актуальність: у роботах, присвячених вирішенню завдань виявлення викидів, відсутнє визначення щодо найкращих серед існуючих математичних моделей або методів для виявлення аномалій та не враховується метрика для визначення точності цих моделей.

Таким чином, виникає потреба у виборі метрик для оцінювання правильності виявлення викидів, найкращих математичних моделей, методів та засобів реалізації інформаційної технології при виявленні аномалій у часових рядах та вибірках даних в системах автоматизованого збору даних.

Мета: ця робота присвячена вибору метрик, найкращих математичних моделей та методів для вирішення проблеми виявлення викидів у часових рядах та вибірках даних в системах автоматизованого збору даних.

Методи дослідження: в ході виконання досліджень були використані методи обробки та підготовки даних, математичні моделі і методи виявлення викидів. Програмне забезпечення розроблено за допомогою мови Python. Також були використані наступні бібліотеки: scikit-learn, NumPy, Pandas, Tensorflow і інші.

Результати: проведено аналіз метрик, які використовуються при оцінці якості математичних моделей та методів виявлення викидів. Проведено аналіз результатів дослідження на 12 пробних вибірках щодо ефективності математичних моделей та методів.

Як результат, визначені найбільш ефективні математичні моделі, методи та метрика для вирішення проблеми виявлення викидів у часових рядах та вибірках даних в системах автоматизованого збору даних.

Висновки: визначено, що найкращою є метрика PR-AUC. Ця метрика була обрана найкращою, через те що вона фокусується на малих позитивних класах - в нашому випадку викидах, та дозволяє об'єктивно оцінити якість математичних моделей та методів.

На основі проведеного аналізу результатів дослідження, було визначено, що найбільш ефективним є метод ізоляційний ліс. Серед математичних моделей та методів глибокого навчання для виявлення викидів було визначено, що найбільш ефективним є метод, що базується на генеративних змагальних мережах (GAN).

Обираючи між ізоляційним лісом та GAN слід підкреслити, що обидва методи мають досить високі показники ефективності виявлення викидів на пробних вибірках. Вибір між ними залежить від обчислювальних можливостей та поставлених цілей, щодо якості виявлення викидів.

Ключові слова: виявлення викидів, машинне навчання, глибоке навчання, метрики оцінки якості моделей та методів.

Постановка проблеми та її актуальність

Будемо розглядати в якості об'єкту дослідження системи автоматизованого збору даних, наприклад, економічні системи, засновані на даних моніторингу контрольованих змінних стану. Результати моніторингу – це вибірки даних та часові ряди. Часові ряди, будучи дискретною моделлю контролю стану динамічних систем, зазвичай містять параметричну невизначеність, є нестационарними і зашумленими.

Задача виявлення викидів в результаті її декомпозиції повинна бути представлена, як послідовність вирішення взаємопов'язаних задач таких як: моніторинг стану системи (вибір та вимірювання значень контрольованих змінних стану системи через певні проміжки часу); попередня обробка даних, яка призводить до приведення даних моніторингу до виду, придатного для виявлення викидів; знаходження аномальних значень у часових рядах та вибірках даних моніторингу системи.

Розгляду завдань теорії та практики автоматизованого збору даних приділяється велика увага як науковцям в Україні, так і за її межами. На цей час опубліковано безліч робіт, присвячених опису математичних моделей та методів виявлення викидів у системах автоматизованого збору даних [1-15].

Розглядаючи класифікацію методів (моделей) виявлення викидів в системах можна визначити, що їх можна поділити на традиційні методи та методи глибокого навчання. До множини традиційних методів, моделей можна віднести: KNN [16], ODIN [17], LOF [18], SVM [19], Z-оцінку, DBSCAN [20] та Isolation Forest [21]. Серед множини методів глибокого навчання можна виділити: Автоенкодер [22], Варіаційний автоенкодер (VAE) [23]; Генеративні змагальні мережі (GAN) [24].

Аналіз існуючих літературних джерел показує, при розробці математичних моделей та методів вирішення завдань виявлення викидів у вибірках даних та часових рядах виникає низка

проблем: невизначеність вхідних даних (обмежений обсяг вибірок, наявність пропущених значень, корельованість змінних станів); велика розмірність множини змінних стану; невизначеність у виборі форми приведення вхідних даних до нормального вигляду, придатного для моделей виявлення викидів; невизначеність у виборі критеріїв якості математичних моделей та методів; невизначеність вибору правильних результатів рішень щодо аномальних значень.

Слід зазначити, що у роботах, присвячених вирішенню завдань виявлення викидів, відсутнє визначення щодо найкращих серед існуючих математичної моделі або методу для виявлення аномалій та не враховується метрика для визначення точності цих моделей.

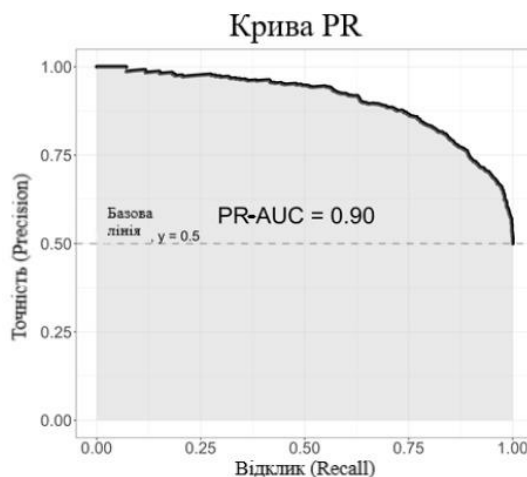
Таким чином, виникає потреба у виборі метрик для оцінювання правильності виявлення викидів, найкращих математичних моделей, методів та засобів реалізації інформаційної технології при виявленні аномалій у часових рядах та вибірках даних в системах автоматизованого збору даних.

Метрика для оцінювання якості виявлення викидів

Метрика в загальному сенсі машинного навчання – це еталон вимірювання якості математичного методу або моделі. Існує проблема у виборі єдиної метрики для визначення якості математичних методів, моделей виявлення викидів у пробних вибірках.

В даний час є велика кількість метрик [25-27], які підтвердили свою використовуваність за довгі роки у задачах виявлення викидів, такі як: точність (Accuracy); точність (Precision) і відклик (Recall); F; Кappa; ROC-AUC; PR-AUC; pAUC (Partial AUC); двосторонній pAUC (Partial AUC); LogLoss.

Проаналізувавши усі метрики, які використовуються при оцінці якості методів (моделей) виявлення викидів, можна визначити, що найкращою є метрика PR-AUC. Ця метрика розраховується так само, як ROC-AUC. Відмінною рисою є те, що крива будується на основі точності (Precision) і відклику (Recall). Крива PR зображена на рисунку, де відображена залежність точності від відклику. По осі ординат відображена міра точності, по осі абсцис — міра відклику. Площа під кривою PR визначає значення метрики PR-AUC. Для відмінної моделі значення цієї метрики буде дорівнювати одиниці.



Не всі значення від 0 до 1 є досяжними для PR-AUC. Залежно від даних, базовою лінією кривої PR є горизонтальна лінія, яка дорівнює значенню точності (Precision) — найменше значення точності. Коли поріг стає 0 (тобто крайня права точка на графіку), всі зразки класифікуються як позитивні. Тобто кількість зразків позитивного класу дорівнює кількості позитивних зразків. Таким чином, значення базової лінії зменшується, коли дані стають більш незбалансованими.

Цю метрику слід використовувати, коли модель повинна відмінно визначати або позитивні класи, або негативні. Тобто за допомогою PR-AUC можна сфокусуватися на правильному прогнозуванні одного із класів.

Ця метрика була обрана найкращою, через те що вона фокусується на малих позитивних класах, в нашому випадку викидах та дозволяє об'єктивно оцінити якість математичних моделей та методів.

Огляд результатів та ефективності методів

Для дослідження були обрані традиційні методи та методи глибокого навчання для виявлення викидів. Серед традиційних були обрані: Z-оцінка (стандартна оцінка спостереження); DBSCAN; ізоляційний ліс; Local Outlier Factor (LOF). Розглядаючи методи глибокого навчання були обрані: автоенкодер; варіаційний автоенкодер (VAE); генеративні змагальні мережі (GAN).

В якості наборів даних були обрані різні вибірки, які використовуються для оцінки якості методів виявлення викидів, наприклад ForestCover, Satellite та інші. Також було створено власний набір даних, який буде позначатися, як «Economical (наш)». Цей набір даних відображає соціально-економічний розвиток країн за 2012-2020 роки. Він представляє 115 альтернатив з 32 змінними стану. Перед використанням, ці набори даних були попередньо чисельно оброблені. Кожен набір даних не має пропущених значень та складається тільки з числових даних. Тобто, набори даних були приведені до виду, який підтримує кожна розроблена математична модель та метод.

Для тестування та оцінки якості було розроблено програмне забезпечення, яке дозволяє виконувати тренування математичних моделей та методів та розраховувати показник ефективності. Програмне забезпечення розроблено за допомогою мови Python. Також були використані наступні бібліотеки: scikit-learn, NumPy, Pandas, Tensorflow і інші.

Було проведено дослідження щодо ефективності традиційних математичних моделей та методів для виявлення викидів. В якості показника ефективності було значення обраної метрики PR-AUC.

Результати оцінки якості традиційних методів виявлення викидів на різних наборах даних приведені в таблиці 1.

Таблиця 1. Результати тестування традиційних методів для виявлення викидів

Набір даних	PR-AUC			
	Ізоляційний ліс	DBSCAN	LOF	Z-оцінка
Http (KDDCUP99)	0.90	0.60	0.45	0.50
ForestCover	0.85	0.83	0.72	0.60
Mulcross	0.89	0.33	0.37	0.41
Smtп (KDDCUP99)	0.83	0.80	0.65	0.60
Shuttle	0.81	0.60	0.55	0.67
Mammography	0.86	0.77	0.67	0.65
Satellite	0.82	0.68	0.72	0.66
Pima	0.71	0.65	0.52	0.60
Breastw	0.67	0.71	0.49	0.85
Arrhythmia	0.87	0.86	0.37	0.41
Ionosphere	0.80	0.78	0.73	0.63
Economical (наш)	0.84	0.73	0.65	0.54

Результати оцінки якості методів глибокого навчання для виявлення викидів на різних наборах даних приведені в таблиці 2.

Таблиця 2. Результати тестування методів глибокого навчання для виявлення викидів

Набір даних	PR-AUC		
	Автоенкодер	Варіаційний автоенкодер	GAN
Http (KDDCUP99)	0.90	0.97	0.95
ForestCover	0.88	0.93	0.97
Mulcross	0.92	0.95	0.76
Smtп (KDDCUP99)	0.89	0.91	0.90
Shuttle	0.93	0.95	0.94
Mammography	0.87	0.89	0.90
Satellite	0.85	0.87	0.88
Pima	0.81	0.84	0.86
Breastw	0.75	0.77	0.79
Arrhythmia	0.89	0.90	0.93
Ionosphere	0.85	0.86	0.90
Economical (наш)	0.88	0.90	0.86

Вибираючи найкращі методи серед традиційних та глибокого навчання, не важко бачити, що кращими є ізоляційний ліс та метод, що базується на генеративних змагальних мережах (GAN).

Результати та висновки

Проведено аналіз метрик, які використовуються при оцінці якості математичних моделей та методів виявлення викидів. Визначено, що найкращою є метрика PR-AUC. Ця метрика була обрана найкращою, через те що вона фокусується на малих позитивних класах - в нашому випадку викидах, та дозволяє об'єктивно оцінити якість математичних моделей та методів.

На основі проведеного аналізу результатів дослідження (було використано 12 пробних вибірок) щодо ефективності традиційних математичних моделей та методів, було визначено, що найбільш ефективним є метод ізоляційний ліс. Цей метод здобув найкращі результати показників ефективності. Перевагою методу є відсутність необхідності масштабувати дані в просторі та відсутність великої кількості параметрів, що робить його надійним та простим в підборі параметрів для покращення роботи методу.

Серед математичних моделей та методів глибокого навчання для виявлення викидів було визначено, що найбільш ефективним є метод, що базується на генеративних змагальних мережах (GAN). Він досяг найкращих показників ефективності серед усіх математичних моделей та методів виявлення викидів, які були розглянуті в роботі, використовуючи різні набори даних. Перевагами даного методу є те, що метод деталізує дані та може легко інтерпретуватися у різні версії. Цей метод потребує великі обчислювальні можливості, тому його слід використовувати при їх наявності.

Обираючи між ізоляційним лісом та GAN слід підкреслити, що обидва методи мають досить високі показники ефективності виявлення викидів на пробних вибірках. Вибір між ними залежить від обчислювальних можливостей та поставлених цілей, щодо якості виявлення викидів.

ЛІТЕРАТУРА

1. В.П. Шкодіреєв, К.І. Ягафаров, В.А. Баштовенко, Є.Е. Ільїна. Огляд методів виявлення аномалій в потоках даних. URL: http://ceur-ws.org/Vol-1864/paper_33.pdf (дата звернення: 10.11.2021).

2. М.В. Ломоносова. Виявлення аномалій у роботі механізмів методами машинного навчання. URL: <http://ceur-ws.org/Vol-2022/paper59.pdf> (дата звернення: 10.11.2021).
3. Chalapathy R., Chawla S. Deep Learning for Anomaly Detection: A Survey. URL: <https://arxiv.org/abs/1901.03407> (дата звернення: 20.10.2022)
4. Thudumu, S., Branch, P., Jin, J. et al. A comprehensive survey of anomaly detection techniques for high dimensional big data. J Big Data 7, 42 (2020). DOI: <https://doi.org/10.1186/s40537-020-00320-x>. (дата звернення: 20. 10. 2022)
5. Deep Learning for Anomaly Detection: A Review: ACM Computing Surveys: Vol 54, No 2. URL: <https://dl.acm.org/doi/10.1145/3439950> (дата звернення: 20. 10. 2022).
6. Muruti G., Rahim F., bin Ibrahim Z. A Survey on Anomalies Detection Techniques and Measurement Methods // 2018 IEEE Conference on Application, Information and Network Security (AINS). 2018.
7. Agrawal, S. and Agrawal, J. (2015) Survey on anomaly detection using data mining techniques, Procedia Computer Science. Elsevier. URL: <https://www.sciencedirect.com/science/article/pii/S1877050915023479> (дата звернення: 20. 10. 2022)
8. Guansong Pang, Chunhua Shen, Longbing Cao, and Anton Van Den Hengel. 2021. Deep Learning for Anomaly Detection: A Review. ACM Comput. Surv. 54, 2, Article 38 (March 2022), 38 pages. DOI: <https://doi.org/10.1145/3439950>. (дата звернення: 20. 10. 2022)
9. Nassif, A.B. (2022) Machine Learning for Anomaly Detection: A Systematic Review, IEEE Access. Institute of Electrical and Electronics Engineers (IEEE). URL: https://www.academia.edu/66087914/Machine_Learning_for_Anomaly_Detection_A_Systematic_Review (дата звернення: 20. 10. 2022).
10. Golan, I. and El-Yaniv, R. (2018) Deep anomaly detection using geometric transformations, arXiv.org. URL: <https://arxiv.org/abs/1805.10917> (дата звернення: 20. 10. 2022).
11. Braei, M. and Wagner, S. (2020) Anomaly detection in Univariate Time-Series: A Survey on the state-of-the-art, arXiv.org. URL: <https://arxiv.org/abs/2004.00433> (дата звернення: 20. 10. 2022).
12. Rehman, Atiq & Brahim Belhaouari, Samir. (2021). Unsupervised outlier detection in multidimensional data. Journal of Big Data. 8. URL: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00469-z> (дата звернення: 20. 10. 2022).
13. Hodge, V.J, Austin, J. (2004) A survey of outlier detection methodologies. Artificial Intelligence Review. pp. 85-126. ISSN 1573-7462. URL: <https://core.ac.uk/download/pdf/58585.pdf> (дата звернення: 20. 10. 2022).
14. Singh, Karanjit & Upadhyaya, Shuchita. (2012). Outlier Detection: Applications And Techniques. International Journal of Computer Science Issues. 9. URL: https://www.researchgate.net/publication/267964435_Outlier_Detection_Applications_And_Techniques (дата звернення: 20. 10. 2022).
15. Wang, Siqi and Zeng, Yijie and Liu, Xinwang and Zhu, En and Yin, Jianping and Xu, Chuanfu and Kloft, Marius (2019). Effective End-to-end Unsupervised Outlier Detection via Inlier Priority of Discriminative Network. URL: <https://proceedings.neurips.cc/paper/2019/hash/6c4bb406b3e7cd5447f7a76fd7008806-Abstract.html> (дата звернення: 20. 10. 2022).
16. Cunningham, Pdraig & Delany, Sarah. (2007). k-Nearest neighbour classifiers. Mult Classif Syst. 54. URL: https://www.researchgate.net/publication/228686398_k-Nearest_neighbour_classifiers (дата звернення: 20. 10. 2022).
17. Hsu, Y.-C. et al. (2020) Generalized Odin: Detecting out-of-distribution image without learning from out-of-distribution data, arXiv.org. URL: <https://arxiv.org/abs/2002.11297> (дата звернення: 20. 10. 2022).
18. Breunig, Markus & Kriegel, Hans-Peter & Ng, Raymond & Sander, Joerg. (2000). LOF: Identifying Density-Based Local Outliers. ACM Sigmod Record. 29. 93-104. URL: https://www.researchgate.net/publication/221214719_LOF_Identifying_Density-Based_Local_Outliers (дата звернення: 20. 10. 2022).

19. Evgeniou, Theodoros & Pontil, Massimiliano. (2001). Support Vector Machines: Theory and Applications. 2049. 249-257. URL: https://www.researchgate.net/publication/221621494_Support_Vector_Machines_Theory_and_Applications (дата звернення: 20. 10. 2022).
20. T. M. Thang and J. Kim, "The Anomaly Detection by Using DBSCAN Clustering with Multiple Parameters," 2011 International Conference on Information Science and Applications, 2011, pp. 1-5, DOI: 10.1109/ICISA.2011.5772437.
21. Liu, Fei Tony & Ting, Kai & Zhou, Zhi-Hua. (2009). Isolation Forest. 413 - 422. URL: https://www.researchgate.net/publication/224384174_Isolation_Forest (дата звернення: 20. 10. 2022).
22. Z. Chen, C. K. Yeo, B. S. Lee and C. T. Lau, "Autoencoder-based network anomaly detection," 2018 Wireless Telecommunications Symposium (WTS), 2018, pp. 1-5, DOI: 10.1109/WTS.2018.8363930.
23. R. Yao, C. Liu, L. Zhang and P. Peng, "Unsupervised Anomaly Detection Using Variational Auto-Encoder based Feature Extraction," 2019 IEEE International Conference on Prognostics and Health Management (ICPHM), 2019, pp. 1-7, DOI: 10.1109/ICPHM.2019.8819434.
24. Han, X., Chen, X. and Liu, L.-P. (2020) Gan Ensemble for Anomaly Detection, arXiv.org. URL: <https://arxiv.org/abs/2012.07988> (дата звернення: 20. 10. 2022).
25. Hossin, Mohammad & M.N, Sulaiman. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. International Journal of Data Mining & Knowledge Management Process. 5. DOI: 10.5121/ijdkp.2015.5201.
26. Davis, J. and Goadrich, M. (2006) "The relationship between precision-recall and ROC curves," Proceedings of the 23rd international conference on Machine learning - ICML '06. DOI: 10.1145/1143844.1143874.
27. Sokolova, Marina & Japkowicz, Nathalie & Szpakowicz, Stan. (2006). Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. AI 2006: Advances in Artificial Intelligence, Lecture Notes in Computer Science. Vol. 4304. 1015-1021. DOI: 10.1007/11941439_114.

ШЕВЧЕНКО Дмитро Олександрович – аспірант; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: dimyich24@gmail.com; ORCID: 0000-0002-7897-250X.

Наукові інтереси:

- методи обчислюваного інтелекту;
- прогнозування стану систем на основі даних моніторингу.

ЛИХАЧ Олег Юрійович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: lykhach2018@gmail.com; ORCID: 0000-0002-4598-2912.

Наукові інтереси:

- методи обчислюваного інтелекту;
- прогнозування стану систем на основі даних моніторингу.

УГРЮМОВ Михайло Леонідович – д.т.н., професор кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: ugryumov.mykhaylo52@gmail.com; ORCID: 0000-0003-0902-2735.

Наукові інтереси:

- методи обчислюваного інтелекту;
- прогнозування стану систем на основі даних моніторингу.

ШМАТКОВ С. І., ТИТАРЕНКО Т. О.

МЕТОД АДАПТИВНОЇ КЛАСТЕРИЗАЦІЇ ДАНИХ НА ОСНОВІ МЕТЕТИЧЕСКОГО АЛГОРИТМУ

Актуальність: кластеризацією є розбиття множини даних на групи за схожими ознаками. Кластеризація використовується при вирішенні різноманітних задач обробки даних, в тому числі при розпізнаванні образів, машинному навчанні, автоматичній класифікації, виробленні стратегій керування і т. д. До цих пір не було знайдено якогось універсального алгоритму, який був би ефективним на даних різної природи. Таким чином, виникає потреба у виборі метода для кластеризації даних різної природи.

Мета: метою роботи було модифікувати метод адаптивної кластеризації даних на основі мететичного алгоритму, який був би позбавлений недоліків традиційних підходів до кластеризації даних

Методи дослідження: в ході виконання досліджень були використані методи обробки та підготовки даних. Програмне забезпечення розроблено за допомогою мови Java. Також були використані наступні бібліотеки: Weka, DeepLearning4, MALLET, Math і інші.

Результати: експерименти підтвердили ефективність розробленого підходу, який характеризується простотою чисельного впровадження та досить високим коефіцієнтом конвергенції.

Як результат, було модифіковано, внесено на основі процедури оптимізації котячих зграй з поліпшеними властивостями за допомогою стохастичної оцінки градієнта

Висновки: Розглянуто проблему адаптивної кластеризації на основі імовірнісного, можливого та достовірного підходів. Повторні модифікації відомих пакетних процедур, призначених для вирішення проблем видобутку потоку даних, дозволяють обробляти інформацію в режимі онлайн як послідовне доповнення до розглянутої системи. Оскільки цільові функції адаптивної кластеризації в загальному випадку є багатоекстремальними, було запропоновано вдосконалити рішення, використовуючи алгоритми еволюційної оптимізації зграй.

Ключові слова: мететичний алгоритм, оптимізація, котячі зграї.

Постановка проблеми та її актуальність

Проблема адаптивної кластеризації масивів даних розглядається в умовах, коли сформовані кластери доволіно перекриваються в просторі ознак.[1-2] Вихідною інформацією для вирішення задачі є масив багатовимірних векторів даних, утворених набором векторних спостережень та глобальний екстремум функції. [3]Також проблема, пов'язана з векторними спостереженнями, кластеризацією, часто виникає у багатьох додатках інтелектуального аналізу даних, і насамперед у адаптивній кластеризації даних при обробці векторного спостереження з різними рівнями ймовірності, достовірності, тощо може належати більше ніж до одного класу.

Завдання класифікації в режимі кластеризації багатовимірних даних є важливою частиною традиційного інтелектуального аналізу, такого як видобуток даних, динамічний видобуток даних, видобуток потоків даних, видобуток великих даних, веб-видобуток. [4-6]

Оптимізація котячої зграї на основі мететичного алгоритму

Для пошуку глобального екстремуму функції достовірності доцільно використовувати алгоритми оптимізації зграй біо-еволюційних частинок [7]. Серед алгоритмів зграй найшвидшими є алгоритми зграй котів [8–9], які виявилися ефективними у вирішенні широкого кола задач з обробки даних. Модель поведінки котячої зграї (CS) передбачає, що кожна кішка зграї, що складається з Q осіб ($p = 1, 2, \dots, Q$) може бути в одному з двох станів: Режим пошуку (SM) та Режим трасування (TM). У цьому випадку режим пошуку пов'язаний з повільними рухами з невеликою амплітудою поблизу початкової позиції (сканування простору в районі поточної позиції) та режимом трасування, який визначається швидкими стрибками з великою амплітудою і дозволяє коту піти з місцевого екстремуму, якщо вона туди потрапила.[10] Поєднання локального сканування та різких змін у поточному стані робить більш імовірним пошук глобального екстремуму порівняно з традиційними методами багатоекстремальної оптимізації.

У цьому алгоритмі кожна кішка може мати два паралельних стани: режим пошуку та режим відстеження. Цей підхід забезпечує пошук глобального екстремуму в тому випадку, коли кількість котів у зграї достатня.[11-12]

Дослідження та огляд результатів та ефективності методів

Адаптивна кластеризація на основі меметичного алгоритму проводилась на чотирьох різних вибірках даних: Cat_1, Cat_2 and Cat_3. Кожен із наборів даних має ряд параметрів, представлених у табл. 1.

Таблиця 1. Параметричні характеристики наборів даних

Вибірка	Кількість кластерів	Кількість атрибутів	Кількість спостережень
Cat_1	2	9	683
Cat_2	3	13	178
Cat_3	6	8	214

Для налагодження алгоритму оптимізації котячих зграй, використовувались параметри, представлені в табл. 2

Таблиця 2. Параметри алгоритму оптимізації зграї котів

Параметри	Значення
SRD	Випадково [0,1]
Seeking memory Pool (SMP)	5
Population size	Кількість кластерів
R	Випадково
C	Const
SPC	Випадково [0,1]
Кількість ітерацій	Вручну

Результати обробки часу алгоритмів кластеризації, таких як нечіткий алгоритм с-середніх (FCM), оптимізація роя частинок (PSO), алгоритм Гауса - Зейделя (GSA), CSO та методу адаптивної кластеризації даних на основі меметичного алгоритму (ACMA) продемонстровано в табл. 3. Порівняльний аналіз демонструє достатньо високу швидкість роботи, не поступаючи більш відомим на сьогодні алгоритмам адаптивної кластеризації даних. За допомогою оптимізаційних процедур, які містять основі оптимізації зграї котів (CSO), збільшує швидкість роботи запропонованого методу в кілька разів. [13-14]

Таблиця 3. Порівняльні результати обробки часу алгоритмів кластеризації

Вибірка	FCM	PSO	GSA	CSO	ACMA
Cat_1	0.009	0.138	0.204	0.026	0.01
Cat_2	0.010	0.431	0.432	0.021	0.02
Cat_3	0.009	0.282	0.098	0.076	0.02

Результати та висновки

Розглянуто метод адаптивної кластеризації даних, можливого та достовірного підходів. Повторні модифікації відомих пакетних процедур, призначених для вирішення проблем видобутку потоку даних, дозволяють обробляти інформацію в режимі онлайн як послідовне доповнення до розглянутої системи. Оскільки цільові функції кластеризації в загальному випадку є багатоекстремальними, було запропоновано вдосконалити рішення, використовуючи алгоритми меметичної оптимізації зграї.

Запропоновано модифікацію, внесену на основі процедури оптимізації котячих зграй з поліпшеними властивостями за допомогою стохастичної оцінки градієнта. Експерименти підтвердили ефективність розробленого підходу, який характеризується простотою чисельного впровадження та досить високим коефіцієнтом конвергенції.

ЛІТЕРАТУРА

1. Wunsch Xu.R. Clustering / Xu.R. Wunsch, D.C. Hoboken, N.J. John. – New York: Wiley & Sons, 2009. – 234 p.
2. Aggarwal C.C. Data Mining / C.C. Aggarwal. – Berlin: Springer, 2015. – 43 p.
3. Kohonen T. Self-Organizing Maps / T. Kohonen. – Berlin: Springer-Verlag, 1995. – 121 p.
4. Bezdek J.C. Pattern Recognition with Fuzzy Objective Function Algorithms / J.C. Bezdek. – New York: Plenum Press, 1981. – 421 p.
5. Credibilistic clustering: the model and algorithms / J. Zhou, Q. Wang, C.C. Hung, X. Yi // Fuzziness and Knowledge-Based Systems. – 2015. – No. 4. – P. 545-564.
6. Zhou J. Credibilistic clustering algorithms via alternating cluster estimation / J. Zhou, Q. Wang, C.C. Hung // Journal of Intelligent Manufacturing. – 2017. – No. 5. – P. 727-738.
7. Young F.W. Theory and Applications of Multidimensional Scaling-Hillsdale / R.M. Hamer, F.W. Young. – New York: Erlbaum, 1994. – 26 p.
8. Кібернетика та системний аналіз
9. Zhou J. A generalized approach to possibilistic clustering algorithms / J. Zhou, C.C. Hung // Fuzziness and Knowledge-Based Systems. – 2007. – No. 15. – P. 117-138.
10. Fuzzy clustering of incomplete data by means of similarity measures / Zh. Hu, Ye. Bodyanskiy, O. Tyshchenko, A. Shafronenko // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering. – Lviv, 2-6 Jul 2019. – P. 149-152.
11. Grosan C. Swarm intelligence in Data Mining / C. Grosan, A. Abraham, M. Chis. – Berlin: Springer, 2006. – 123 p.
12. Chu S.C. Cat swarm optimization / S.C. Chu, P.W. Tsai, J.S. Pan. – Berlin: Springer-Verlag, 2006. – 210 p.
13. Chu S.C. Computational Intelligence based on the behavior of cats / S.C. Chu, P.W. Tsai // International Journal of Innovative Computing, Information, and Control. – 2007. – No 1. – P. 163-173.
14. Fuzzy Clusterization of Distorted by Missing Observations Data Sets Using Evolutionary Optimization / A. Shafronenko, Ye. Bodyanskiy, I. Pliss, K. Patlan // Proceedings "Advanced Computer Information Technologies. – Česke Budejovice, 5-7 June 2019. – P. 217-220. <https://doi.org/10.1109/ACITT.2019.8779888>.
15. Shafronenko A. The Fast Modification of Evolutionary Bioinspired Cat Swarm Optimization Method / A. Shafronenko, Ye. Bodyanskiy, I.P. Pliss // IEEE 8th International Conference on Advanced Optoelectronics and Lasers (CAOL). – Sozopol, 6-8 September 2019. – P. 548-552. <https://doi.org/10.1109/CAOL46282.2019.9019583>.

ШМАТКОВ Сергій Ігорович – завідувач кафедри теоретичної та прикладної системотехніки; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: s.shmatkov@karazin.ua; ORCID: 0000-0002-0298-7174.

Наукові інтереси:

- методи обчислюваного інтелекту;
- прогнозування стану систем на основі даних моніторингу.

ТИТАРЕНКО Тимур Олегович – студент; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, 6, Харків, Україна, 61022; e-mail: ха12283450@student.karazin.ua; ORCID: 0000-0001-8731-7198.

Наукові інтереси:

- методи обчислюваного інтелекту;
- прогнозування стану систем на основі даних моніторингу.

ДАТЧИКИ ВИМІРЮВАННЯ ВІДСТАНЕЙ ЯК ОСНОВА 3D-СКАНУВАННЯ

В теперішній час 3D-сканування має достатньо широкий спектр застосувань такі як: анімація, медицина, археологія, біометрія обличчя, сканування об'єктів на конвеєрах.

Тема роботи: дослідження використання датчиків вимірювання відстаней для 3D-сканування.

Мета роботи: дослідити різновиди датчиків вимірювання відстаней. Порівняти їх і знайти особливості застосування. Побудувати примітивний макет для 3D-сканування.

Результат: проведено дослідження різних типів сенсорів та побудований макет для сканування.

Висновок: з результату сканування макету та проведених досліджень щодо застосування даного типів датчиків, можна зробити висновок що, датчики вимірювання відстаней чудово виконують задачу 3D-сканування

Ключові слова: датчики вимірювання відстаней, ToF-сенсори, LIDAR, 3D-сканування

Різновиди датчиків вимірювання відстаней

Далекомір – це пристрій, що визначає відстань до об'єкту. Умовно поділяються на два типи: активні та пасивні(оптичні).

Принцип роботи оптичного далекоміру зводиться до задачі по рівнобедрений трикутник. Заздалегідь відомі відстань між окулярами (база), що являється основою трикутника, і кути при основі. З отриманих даних можна обрахувати відстань до об'єкта.

Активні далекоміри посилають сигнали будь-то світло чи звук, і вимірюють час за який сигнал повернеться на приймач. Даний тип приладів включають в себе передавач і приймач на певній відстані один від одного. Швидкість звуку чи світла вважають заздалегідь відомою. Датчики що посилають звуковий сигнал називаються – сонари, а ті що світловий сигнал ToF-сенсор (якщо випромінювач – світлодіод) та LIDAR (якщо випромінювач – лазер).

В даній роботі сонари не досліджувались тому, що звукова хвиля зі збільшенням відстані – розширюється і тому визначити відстань до певної точки простору зі збереженням деталізації є неможливим.

ToF-сенсор

Time-of-Flight (час польоту) – датчик активного типу. Випромінювання може створюватись як і лазером так і світлодіодом, в залежності від призначення використання. Сенсори в основі яких лазер, можна віднести до підкласу LiDAR, з відмінністю тим що сканування відбувається за принципом точка-точка, коли в лідара випускається пучок високо потужних лазерів.

Але існують ToF-сенсори в яких випромінювач є матрицею випромінювачів, і аналогічно із приймачами що об'єднанні в матрицю. Це дозволяє вимірювати за раз декілька відстаней до різних точок, і за звичай такі сенсори мають можливість налаштування кута зору.

Типи пристроїв:

- Радіочастотна модуляція джерела світла з фазовим детектором

Випромінювач відправляє певну модульовану хвилю, а приймач знаходить максимум кореляції між первісною хвилею. Це і визначає час, що потратив сигнал, щоб потрапити на приймач.

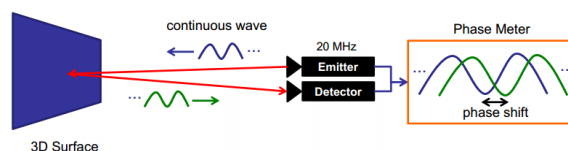


Рис.1 Схема модульованого ToF-сенсору[1]

Нехай випромінюваний сигнал $g(t) = \cos(\omega t)$ де ω - це модулюючий сигнал

Тоді отриманий сигнал $s(t) = b + a\cos(\omega t + \varphi)$ де b - деякий зсув, a - амплітуда

Кореляція $c(\tau) = s * g = \int_{-\infty}^{\infty} s(t) * g(t + \tau)dt$

$$c(\tau) = \frac{a}{2} \cos(\omega\tau + \varphi) + b$$

Але на повну кореляцію виконати складно, тому отриманий сигнал приймається на 4 сусідніх пікселя із зсувом 90° по фазі і тим самим корелює сам себе

Тоді зсув по фазі визначається наступним чином:

$$A_i = c\left(i * \frac{\pi}{2}\right), i = 0, \dots, 3$$

$$\varphi = \arctan2(A_3 - A_1, A_0 - A_2)$$

Після того як знайшли зсув по фазі, знаходимо дистанцію: $d = \frac{c}{4\pi\omega} \varphi$

- «Range gated imagers»

По своєму принципу це затворна камера. Припустимо що в нас затримка спрацювання затвору рівна 0. Час спрацювання затвору = t. Тоді об'єкти що знаходяться далі чим $t/2 * c$ просто не потраплять на камеру. І чим ближче буде розміщений об'єкт тим яскравішим він буде.

- «Direct Time-of-Flight imagers» або «Pulsed Modulation»

Принцип роботи простий і описаний вище, подається імпульс і в кожній точці матриці вимірюється точний час його повернення.

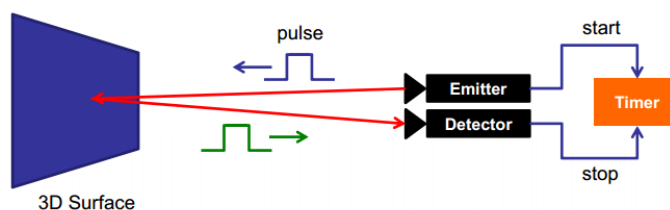


Рис. 2 Схема Direct ToF-сенсору[1]

Матриця складається із тригерів, що спрацьовують по фронту світлової хвилі, але цей метод потребує точного вимірювання часу спрацювання

LIDAR

Принцип роботи такий самий як у «Direct Time-of-Flight imagers», але лідари дозволяють отримати більш точну картину, на великих відстанях, у зв'язку з певними конструктивними особливостями. Перша конструктивна особливість це більш потужні лазери. І різні системи направлення пучка світла, хоча вони всі зводяться до того що є дзеркальце що обертається лише по одній осі. Але при використанні лише одного дзеркальця потрібно більше лазерів, як зображено на рисунку 3. Але можна використовувати лише один лазер і 2 дзеркальця щоб послідовно проходити площину сканування.

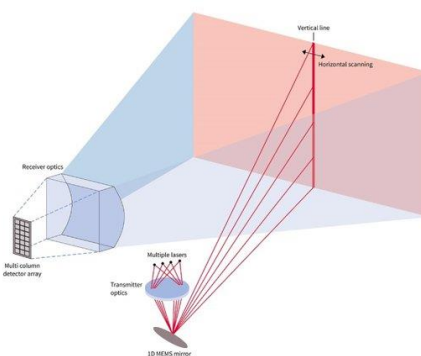


Рис. 3 Схема проходження скануючого пучка світла [2]

Також відмінність лідару від ToF-сенсору, що в нього є вбудований GPS, та IMU(інерційний вбудований пристрій), що дозволяє прив'язувати отриманні дані як і до GPS-

координат так і враховуючи положення самого приладу, такі як прискорення, гіроскоп, і т.д., що дозволяє прив'язувати отримані точки до внутрішніх координат x, y, z .

Сфери використання лідарів пов'язані із великими масштабами наприклад: геодезія, археологія, географія, архітектура, геологія, сейсмологія, атмосферна фізика тощо.

За класифікацією їх поділяють на 3 групи:

- Розташовані за напрямком. Наприклад: лідар-альтиметр дивиться лише вниз, а атмосферний – вгору
- Розташовані на скануючому механізмі. Лідар закріплений на платформі що робить оберт на 360 градусів, і таким чином формується повна картина
- Розташовані на платформі. Це загальна класифікація, адже лідар можна використовувати як і на суші так і в повітрі(напр. літак), також можливе використання на супутниках.

За типом обробки сигналу лідари умовно можна розділити на два типи: лідари, що опрацьовують дискретний сигнал, і на ті що – аналоговий сигнал. У випадку обробки дискретного сигналу приймаються близько чотирьох піків сигналу, на відміну від повного сигналу. Це робить лідари з даним принципом трішки дешевшими, але не такими точними.

Порівняння датчиків

Порівнюючи сенсори можна зробити висновок, що всі вони достатньо точно вимірюють відстань, але для деталізованого зображення більше підходять датчики в основі яких випромінювання світла, а саме LIDAR і ToF-сенсор. Але в залежності від особливостей середовища, де планується їхнє використання, потрібно віддавати перевагу одному чи іншому сенсору. Так LIDAR буде доцільніше використовувати на великих площах, або через наявність IMU його можна закріпити на дроні, і проводити сканування об'єктів в труднодоступних місцях із всіх ракурсів. А ToF-сенсор у зв'язку з обмеженням дальності вимірювання застосовують на не великих масштабах наприклад: сканування об'єктів на конвеєрі, біометрична автентифікація обличчя та відстежування рухів тіла. Сканування обличчя та відстежування рухів виконуються виключно ToF-сенсорами, адже високопотужні лазери LIDAR можуть пошкодити сітківку людини.

Побудова макету та обробка результату сканування

Для побудови примітивної моделі приладу було обрано:

- ToF-сенсор VL53L0X, через його дешевизну та доцільність використовувати в не великих масштабах.
- Мікроконтролер Arduino Uno, базуючись на тому що вона має достатній струм щоб в подальшому жити більш потужніші датчики, та досить пам'яті, для завантаження великих проектів
- 2 сервоприводи.

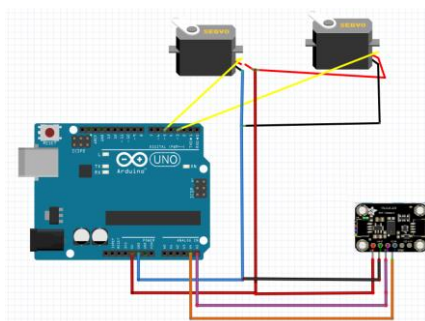


Рис 4. Макет для сканування

Але даний макет, можна вдосконалити за допомогою зміни датчика на більш дорогий та потужний, або змінити сам принцип сканування змінивши положення моторів.

Алгоритм програми:

- 1) Ініціалізація з'єднання з сенсором

- 2) Встановлення сервоприводу в позицію 0
- 3) Дочекатись сигналу сканування від Програми та отримати максимальні межі сканування
- 4) Зробити тричі замір відстані
- 5) Знайти середнє арифметичне
- 6) Збільшити положення сервоприводу на 1
- 7) Повторювати пункти 4-5 до того моменту коли положення сервоприводу буде рівне максимальному для даного сервоприводу.
- 8) Перемістити сервопривід в положення 0.
- 9) Збільшити положення другого сервоприводу на 1. Перейти в пункт 4.

Результат вимірювання є так званої хмарою точок (point cloud) яку можна візуалізувати за допомогою бібліотеки `rptk` для Python

У ході виконання даної роботи були розроблені дві програми. Перша програма керується мікроконтролером, а інша виконується на комп'ютері, що відповідає за рендер зображення і обробку даних

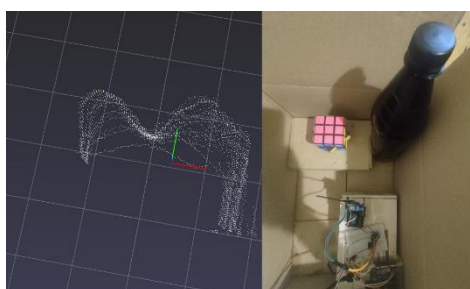


Рис. 5 Отриманий результат

Даний результат сканування не є ідеальним. Це обумовлено в першу чергу дешевизною датчика, тому в певних межах він вимірює з деякою похибкою, що спотворює результат. Більш промислові варіанти таких недоліків не мають.

Література

1. Habr – Time of Flight [Електронний ресурс]. - Режим доступу: URL: <https://habr.com/ru/post/224605/>
2. N. Druml, I. Maksymova, T. Thurner, D. van Lierop, M. Hennecke, and A. Foroutan, “1D MEMS Micro-Scanning LiDAR,” in International Conference on Sensor Device Technologies and Applications (SENSOR-DEVICES), 2018

ЮЩЕНКО Владислав Сергійович – бакалавр з комп'ютерної інженерії; студент-магістр; Харківський національний університет імені В.Н. Каразіна, майдан Свободи, , м. Харків, 61022; vladyslav.yushchenko@student.karazin.ua ; ORCID: 0000-0002-2516-8163

Наукові інтереси:

-Технології, методи та моделі 3D-моделювання

V.N. KARAZIN KHARKIV NATIONAL UNIVERSITY
NATIONAL SCIENCE CENTER KHARKIV INSTITUTE OF PHYSICS AND TECHNOLOGY
MAX PLANCK INSTITUTE OF MICROSTRUCTURE PHYSICS
TARAS SHEVCHENKO NATIONAL UNIVERSITY OF KYIV
INSTITUTE OF NUCLEAR CHEMISTRY AND TECHNOLOGY (Warsaw, Poland)
RIVNE STATE HUMANITARIAN UNIVERSITY
NATIONAL AEROSPACE UNIVERSITY NAMED AFTER N.E. ZHUKOVSKY (Kharkiv)
CJSC "INSTITUTE OF INFORMATION TECHNOLOGY" (Kharkiv)
KHERSON NATIONAL TECHNICAL UNIVERSITY
TEAM INTERNATIONAL SERVICES, INC. (Lake Mary, USA)

PROGRAM COMMITTEE:

Azarenkov M.O., Acad. Of the NAS of Ukraine, Prof., Dr. Sc, Kharkiv, Chairman
 Bardachov Yu. M., prof., Prof., Dr. Sc, Kherson
 Bomba A.Ya., Prof., Dr. Sc, Rivne
 Buy D. B., Prof., Dr. Sc, Kyiv
 Vanin V.A., Prof., Dr. Sc, Kharkiv
 Gorbenko I.D., Prof., Dr. Sc, Kharkiv
 Dolya G.M., Prof., Dr. Sc, Kharkiv
 Zholtkevich G.M., Prof., Dr. Sc., Kharkiv
 Kuklin V.M., Prof., Dr. Sc, Kharkiv
 Lazurik V.T., Prof., Dr. Sc, Kharkiv
 Rossomakhin S.G., Prof., Dr. Sc, Kharkiv
 Savula Y. H., Prof., Dr. Sc, Lviv
 Sporov O. E., Assoc., Ph.D., Kharkiv
 Styervoyedov A. Dr., Halle, Germany
 Styervoyedov M.G., Assoc., Ph.D., Kharkiv
 Tolstoluzhska O.G., Prof., Dr. Sc, Kharkiv
 Tkachuk M.V., Prof., Dr. Sc, Kharkiv
 Kharchenko V.S., Prof., Dr. Sc, Kharkiv
 Khomchenko A.N. Prof., Dr. Sc, Mykolaiv
 Shmatkov S.I., Assoc., Dr. Sc, Kharkiv
 Shulga M.F., Acad. Of the NAS of Ukraine, Prof., Dr. Sc, Kharkiv
 Zimek Z., Ph.D., Warsaw, Poland
 Yanovsky V.V., Prof., Dr. Sc, Kharkiv

ORGANIZING COMMITTEE:

<i>Lazurik V.T.</i> , Dr. Sc, Prof., Dean of the CSD of V.N. Karazin Kharkiv National University, Chairman, <i>Sporov O. E.</i> , Assoc., Ph.D., V.N. Karazin Kharkiv National University, Deputy Chairman, <i>Tolstoluzhska O.G.</i> , Dr. Sc, Prof. of V.N. Karazin Kharkiv National University, Deputy Chairman, <i>Tkachuk M.V.</i> , Dr. Sc, Prof., Head of MST Dept of V.N. Karazin Kharkiv National University, <i>Kuklin V.M.</i> , Dr. Sc, Prof. Head of the AIS Dept of V.N. Karazin Kharkiv National University, <i>Dyuldyia S.V.</i> , Ph.D., Kharkiv Institute Of Physics And Technology, <i>Esin V.I.</i> , Dr. Sc, Prof. of V.N. Karazin Kharkiv National University, <i>Artyukh A.A.</i> , Head of Laboratory of V.N. Karazin Kharkiv National University,	<i>Zholtkevich G.M.</i> , Dr. Sc, Prof. Dean of the MCS of V.N. Karazin Kharkiv National University, <i>Vanin V.A.</i> , Dr. Sc, Prof. NTU "KhPI", <i>Zinoviev D.V.</i> , Senior Lecturer of V.N. Karazin Kharkiv National University, <i>Rassomakhin S.G.</i> , Dr. Sc, Prof., Head of ISTS Dept of V.N. Karazin Kharkiv National University, <i>Styervoyedov A. Dr.</i> , Max Planck Institute of Microstructure Physics (Germany), <i>Petersen S.</i> , CEO TEAM International (Kharkiv), <i>Styervoyedov M.G.</i> , Ph.D., Head of ECS Dept of V.N. Karazin Kharkiv National University, <i>Shmatkov S.I.</i> , Dr. Sc, Prof. of V.N. Karazin Kharkiv National University, <i>Kruhol M.M.</i> , assistant of Power Plant Department, NTU KhPI.
---	--

[http:// www.univer.kharkov.ua](http://www.univer.kharkov.ua)
[http:// www-csd.univer.kharkov.ua](http://www-csd.univer.kharkov.ua)

Наукове видання

**Комп'ютерне моделювання в наукоємних технологіях
(КМНТ-2022)**

*Збірник наукових праць
міжнародної науково-технічної конференції
(23-25 листопада 2022 року, м. Харків, Україна)*

Українською та англійською мовами

Комп'ютерне верстання: О. О. Афанасьєва

Підписано до друку 26.12.2022 р.
Формат 60×84/8. Папір офсетний. Друк ризограф.
Ум. друк. арк.- 30,2
Обл.- вид.арк.- 38,4
Наклад 20 пр. Зам. № 144/2023
Ціна договірна

Видавець і виготовлювач
Харківський національний університет імені В. Н. Каразіна,
61022, м. Харків, майдан Свободи, 4.
Свідоцтво суб'єкта видавничої справи ДК № 3367 від 13.01.09

Видавництво ХНУ імені В. Н. Каразіна