



Національний технічний університет
"Харківський політехнічний інститут"

National Technical University
"Kharkiv Polytechnic Institute"



Сучасні Інформаційні Системи

Advanced Information Systems

Sučasnì ìnformacìjnì sistemi

Том 7, № 2

Volume 7, No. 2

**Щоквартальний
науково-технічний журнал**

**Quarterly
scientific and technical journal**

Заснований у березні 2017 року

Founded in March 2017

У журналі публікуються результати досліджень з експлуатації та розробки сучасних інформаційних систем у різних проблемних галузях. Журнал призначений для наукових працівників, викладачів, докторантів, аспірантів, а також студентів старших курсів відповідних спеціальностей.

The journal publishes the research study from the usage and development of advanced information systems in various problem areas. The journal is intended for researchers, lecturers, doctoral students, postgraduate students, and for senior students of the corresponding specialties.

Засновник і видавець:

Національний технічний університет
"Харківський політехнічний інститут"

Founder and publisher:

National Technical University
"Kharkiv Polytechnic Institute"

Кафедра:

Комп'ютерна інженерія та програмування,
вул. Кирпичова, 2, 61002, м. Харків, Україна

Department:

Computer Engineering and Programming,
61002, Ukraine, Kharkiv, Kyrpychova str., 2

Телефон: +38 (057) 707-61-65

Phone: +38 (057) 707-61-65

E-mail редколегії:

Heorhii.Kuchuk@khpi.edu.ua

E-mail of the editorial board:

Heorhii.Kuchuk@khpi.edu.ua

Інформаційний сайт:

<http://ais.khpi.edu.ua>

Information site:

<http://ais.khpi.edu.ua>

*Затверджений до друку Вченою Радою Національного технічного університету
"Харківський політехнічний інститут" (протокол від 2 червня 2023 року № 5).*

Свідоцтво про державну реєстрацію KB № 22522-12422P від 13.01.2017 р.

Включений до "Переліку наукових фахових видань України, в яких можуть публікуватися результати дисертаційних робіт на здобуття наукових ступенів доктора наук, кандидата наук та ступеня доктора філософії" до категорії Б – наказом Міністерства освіти і науки України від 07.05.2019 № 612 (додаток 7, п. 33)

Харків • 2023

Редакційна колегія

Головний редактор:

СОКОЛ Євген Іванович
(*д-р техн. наук, проф., Харків, Україна*).

Заступник головного редактора:

ЗАКОВОРОТНИЙ Олександр Юрійович
(*д-р техн. наук, проф., Харків, Україна*).

Члени редакційної колегії:

БАЙРАМОВ Азад Агахар огли
(*д-р фіз.-мат. наук, проф., Баку, Азербайджан*);

БЛАУНШТЕЙН Натан Шаєвич
(*д-р фіз.-мат. наук, проф., Беер-Шева, Ізраїль*);

ГНАТЮК Сергій Олександрович
(*д-р техн. наук, проф., Київ, Україна*);

ЗАЙЦЕВА Єлена
(*д-р наук, проф., Жиліна, Словаччина*);

КАРПІНСЬКИЙ Микола Петрович
(*д-р техн. наук, проф., Бельсько-Бяла, Польща*);

КОВАЛЕНКО Андрій Анатолійович
(*д-р техн. наук, проф., Харків, Україна*);

КУЧУК Георгій Анатолійович
(*д-р техн. наук, проф., Харків, Україна*);

ЛИТВИН Василь Володимирович
(*д-р техн. наук, проф., Львів, Україна*);

ЛУКІН Володимир Васильович
(*д-р техн. наук, проф., Харків, Україна*);

МАМУЗИЧ Ілля
(*д-р техн. наук, проф., Загреб, Хорватія*);

МИГУЩЕНКО Руслан Павлович
(*д-р техн. наук, проф., Харків, Україна*);

МОХАММЕД Амін Саліх
(*д-р наук, доц., Ербіль, Ірак*);

ОСТАПОВ Сергій Едуардович
(*д-р фіз.-мат. наук, проф., Чернівці, Україна*);

ПОВОРОЗНЮК Анатолій Іванович
(*д-р техн. наук, проф., Харків, Україна*);

РАСКІН Лев Григорович
(*д-р техн. наук, проф., Харків, Україна*);

САРАВАНА Балайі Баласубраманіан
(*PhD, доц., Ербіль, Ірак*);

СЕМЕНОВ Сергій Геннадійович
(*д-р техн. наук, проф., Краків, Польща*).

СЕРКОВ Олександр Анатолійович
(*д-р техн. наук, проф., Харків, Україна*);

СМІРНОВ Олексій Анатолійович
(*д-р техн. наук, проф., Кропивницький, Україна*);

СТЕЦЬКО Юрій
(*канд. фіз.-мат. наук, Тампа, Флоріда, США*);

СУЧКОВ Григорій Михайлович
(*д-р техн. наук, проф., Харків, Україна*);

УШЕНКО Юрій Олександрович
(*д-р фіз.-мат. наук, проф., Чернівці, Україна*);

ФІЛАТОВА Ганна Євгенівна
(*д-р техн. наук, проф., Харків, Україна*);

ХАРЧЕНКО Вячеслав Сергійович
(*д-р техн. наук, проф., Харків, Україна*);

ШВАЧИЧ Геннадій Григорович
(*д-р техн. наук, проф., Дніпро, Україна*);

ШИШАЦЬКИЙ Андрій Володимирович
(*канд. техн. наук, Київ, Україна*);

Відповідальний секретар:

ПОДОРОЖНЯК Андрій Олексійович
(*канд. техн. наук, доц., Харків, Україна*).

Технічний секретар:

ГРЕБЕНЮК Дарина Сергіївна
(*магістр комп. інж., Харків, Україна*).

Editorial board

Editor-in-Chief:

Yevgen SOKOL
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*).

Associate editor:

Oleksandr Zakovorotniy
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*).

Editorial board members:

Azad Agalar oğlu BAYRAMOV
(*Dr. Sc. (Ph-Math.), Prof., Baku, Azerbaijan*);

Nathan BLAUNSTEIN
(*Dr. Sc. (Ph-Math.), Prof., Beer Sheva, Israel*);

Sergiy GNATYUK
(*Dr. Sc. (Tech.), Prof., Kyiv, Ukraine*);

Elena ZAITSEVA
(*Dr. (Comp. Eng.), Prof., Zilina, Slovakia*);

Mikolay KARPINSKI
(*Dr. Sc. (Tech.), Prof., Bielsko-Biala, Poland*);

Andriy KOVALENKO
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Heorhii KUCHUK
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Vasyl LYTVYN
(*Dr. Sc. (Tech.), Prof., Lviv, Ukraine*);

Volodymyr LUKIN
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Ilya MAMUZIĆ
(*Dr. Sc. (Tech.), Prof., Zagreb, Croatia*);

Ruslan MYGUSHCHENKO
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Amin Salih MOHAMMED
(*Dr. (Comp. Eng.), Ass. Prof., Erbil, Iraq*);

Serhii OSTAPOV
(*Dr. Sc. (Ph-Math.), Prof., Chernivtsi, Ukraine*);

Anatoliy POVOROZNYUK
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Lev RASKIN
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Balasubramanian Balaji SARAVANA
(*PhD (Comp. Eng.), Ass. Prof., Erbil, Iraq*);

Serhii SEMENOV
(*Dr. Sc. (Tech.), Prof., Krakow, Poland*).

Aleksandr SERKOV
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Alexey SMIRNOV
(*Dr. Sc. (Tech.), Prof., Kropyvnytskyi, Ukraine*);

Yuri STETSKO
(*PhD (Ph-Math.), Tampa, Florida, USA*);

Hryhorii SUCHKOV
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Yuri USHENKO
(*Dr. Sc. (Ph-Math.), Prof., Chernivtsi, Ukraine*);

Hanna FILATOVA
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Viacheslav KHARCHENKO
(*Dr. Sc. (Tech.), Prof., Kharkiv, Ukraine*);

Hennadii SHVACHICH
(*Dr. Sc. (Tech.), Prof., Dnipro, Ukraine*);

Andrii SHYSHATSKYI
(*PhD (Tech.), Kyiv, Ukraine*).

Responsible secretary:

Andrii PODOROZHNIAK
(*PhD (Tech.), Ass. Prof., Kharkiv, Ukraine*).

Technical secretary:

Daryna HREBENIUK
(*MSD of Comp. Eng., Kharkiv, Ukraine*).

З М І С Т

ПРОБЛЕМИ ІДЕНТИФІКАЦІЇ В ІНФОРМАЦІЙНИХ СИСТЕМАХ

Гоюшова У. М.

Алгоритми пошуку непересічних доріг
на зображеннях (eng.) 5

*Хорошун Г.М., Рязанцев О.І.,
Черницький М.В.*

Кластеризація та аномальність даних індексу
волатильності фондового ринку США 9

МЕТОДИ СИНТЕЗУ ІНФОРМАЦІЙНИХ СИСТЕМ

*Александров Є. Є., Александрова Т. Є.,
Костяник І. В., Моргун Я. Ю.*

Параметричний синтез інваріантної системи
курсвої стійкості автомобіля (eng.) 16

Воловик А. Ю.

Синтез квазіоптимальних швидкодіючих фільтрів
з позицій критерію найменших квадратів 21

*Гадецька С. В., Дубницький В. Ю.,
Кушнерук Ю. І., Ходирев О. І.*

EXCEL-орієнтований калькулятор для обчислення
результатів ентропійного аналізу даних,
що розподілені по категоріях 28

*Кошман С. О., Краснобаєв В. А.,
Нікольський С. Б., Ковальчук Д. М.*

Структура комп'ютерної системи
у залишкових класах (eng.) 41

ДОСЛІДЖЕННЯ ІНФОРМАЦІЙНИХ СИСТЕМ

*Аль-Мудхафар Акіл Абдулхуссейн М.,
Смірнова Т. В., Буравченко К. О.,
Смірнов О. А.*

Метод оцінки та підвищення користувальницького
досвіду абонентів в програмно-конфігурованих
мережах на основі використання
машинного навчання 49

Дашков Д. Є., Ляшенко О. С.

Захоплення руху за допомогою
датчиків MEMS (eng.) 57

*Раскін Л.Г., Сухомлін Л.В.,
Соколов Д.Д., Власенко В.В.*

Багатокритеріальна оцінка ефективності
багатофакторних стохастичних систем (eng.) 63

ІНТЕЛЕКТУАЛЬНІ ІНФОРМАЦІЙНІ СИСТЕМИ

Дунь Бао, Заковоротний О. Ю., Кучук Н. Г.

Генерація даних про обмінний курс валюти
на основі моделі Quant-GAN (eng.) 68

T A B L E O F C O N T E N T S

PROBLEMS OF IDENTIFICATION IN INFORMATION SYSTEMS

Ulyiyya Goyushova Malik

Algorithms for finding non-intersecting roads
on images 5

*Ganna Khoroshun, Oleksandr Ryazantsev,
Maksym Cherpitskyi*

Clustering and anomaly of the United States
stock market volatility index data (ukr) 9

METHODS OF INFORMATION SYSTEMS SYNTHESIS

*Yevgen Aleksandrov, Tetiana Aleksandrova,
Iryna Kostianyk, Yaroslav Morgun*

Parametric synthesis of the invariant system
of car course stability 16

Andrii Volovyk

Synthesis of quasi-optimal fast filters
by the least square criterion (ukr) 21

*Svitlana Gadetska, Valeriy Dubnitskiy,
Alexander Khodyrev, Yuri Kushneruk*

EXCEL-oriented calculator for calculating results
of entropy analysis of data distributed
by categories (ukr) 28

*Serhii Koshman, Victor Krasnobayev,
Serhii Nikolsky, Dmytro Kovalchuk*

The structure of the computer system
in the residual classes 41

INFORMATION SYSTEMS RESEARCH

*Al-Mudhafar Aqeel Abdulhussein M.,
Tetiana Smirnova, Kostiantyn Buravchenko,
Oleksii Smirnov*

The method of assessing and improving the user
experience of subscribers in software-configured
networks based on the use
of machine learning (ukr) 49

Dmytro Dashkov, Oleksii Liashenko

Motion capture
with MEMS sensors 57

*Lev Raskin, Larysa Sukhomlin,
Dmytro Sokolov, Vitalii Vlasenko*

Multi-criteria evaluation of the multifactor
stochastic systems effectiveness 63

INTELLIGENT INFORMATION SYSTEMS

Bao Dun, Oleksandr Zakovorotnyi, Nina Kuchuk

Generating currency exchange rate data
based on quant-GAN model 68

<i>Захарченко А. С., Степанець О. В.</i> Роль цифрового двійника в проектуванні інтелектуальних будівель (eng.) 75	<i>Anastasiya Zakharchenko, Oleksandr Stepanets</i> Digital Twin value in intelligent building development 75
МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЙНИХ СИСТЕМ	METHODS OF INFORMATION SYSTEMS PROTECTION
<i>Даценко С. С., Кучук Г. А.</i> Біометрична автентифікація, що використовує згорткові нейронні мережі (eng.) 87	<i>Serhii Datsenko, Heorhii Kuchuk</i> Biometric authentication utilizing convolutional neural networks 87
<i>Сальніков Д. В., Караман Д. Г., Крилова В. А.</i> Архітектура гнучко конфігурованих периферійних модулів на базі програмованих процесорних ядер (eng.) 92	<i>Dmytro Salnikov, Dmytro Karaman, Viktoriia Krylova</i> Highly reconfigurable soft-CPU based peripheral modules design 92
Алфавітний покажчик 98	Alphabetical index 98

✦

За достовірність викладених фактів, цитат та інших відомостей
відповідальність несе автор.

Географія статей цього номера



Азербайджан
Azerbaijan



Китай
China



Україна
Ukraine

Організації авторів

Інститут систем управління Міністерства науки і освіти
Азербайджану, Баку, Азербайджан
Банк Китаю, Пекін, Китай
Вінницький національний технічний університет,
Вінниця, Україна
Кременчуцький національний університет імені Михайла
Остроградського, Кременчук, Україна
Національний авіаційний університет, Київ, Україна
Національний технічний університет «Харківський
політехнічний інститут», Харків, Україна
Харківський національний університет ім. В. Н. Каразіна,
Харків, Україна
Національний технічний університет України «КПІ ім.
Ігоря Сікорського», Київ, Україна
Східноукраїнський національний університет імені
Володимира Даля, Київ, Україна
Харківський національний університет радіоелектроніки,
Харків, Україна
Харківський національний автомобільно-дорожній
університет, Харків, Україна
Харківський національний університет Повітряних Сил
імені Івана Кожедуба, Харків, Україна
Харківська гуманітарно-педагогічна академія, Харків,
Україна
Центральноукраїнський національний технічний
університет, Кропивницький, Україна

Authors affiliation

Institute of Control Systems of Azerbaijan National
Academy of Sciences, Baku, Azerbaijan
Bank of China, Beijing, China
Vinnytsia National Technical University, Vinnytsia,
Ukraine
Kremenchuk Mykhailo Ostrohradskyi National
University, Kremenchuk, Ukraine
National Aviation University, Kyiv, Ukraine
National Technical University "Kharkiv Polytechnic
Institute", Kharkiv, Ukraine
V.N. Karazin Kharkiv National University, Kharkiv,
Ukraine
National Technical University "Igor Sikorsky Kyiv
Polytechnic Institute", Kyiv, Ukraine
Volodymyr Dahl East Ukrainian National University,
Kyiv, Ukraine
Kharkiv National University of Radio Electronics,
Kharkiv, Ukraine
Kharkiv National Automobile and Highway University,
Kharkiv, Ukraine
Ivan Kozhedub Kharkiv National Air Force University,
Kharkiv, Ukraine
Kharkiv humanitarian-pedagogical academy, Kharkiv,
Ukraine
Central Ukrainian National Technical University,
Kropivnitskiy, Ukraine

АЛФАВІТНИЙ ПОКАЖЧИК

Александров Е. Е. (Yevgen Aleksandrov)	16	Краснобаев В. А. (Victor Krasnobayev)	41
Александрова Т. Е. (Tetiana Aleksandrova) .	16	Крилова В. А. (Viktoriiia Krylova)	92
Аль-Мудхафар Акіл Абдулхуссейн М. (Al-Mudhafar Aqeel Abdulhussein M.)	49	Кучук Г. А. (Heorhii Kuchuk)	87
Буравченко К. О. (Kostiantyn Buravchenko) .	49	Кучук Н. Г. (Nina Kuchuk)	68
Власенко В. В. (Vitalii Vlasenko)	63	Кушнерук Ю. І. (Yuri Kushneruk)	28
Воловик А. Ю. (Andrii Volovyk)	21	Ляшенко О. С. (Oleksii Liashenko)	57
Гадецька С. В. (Svitlana Gadetska)	28	Моргун Я. Ю. (Yaroslav Morgun)	16
Гоюшова У. М. (Ulviyya Goyushova Malik)	5	Нікольський С. Б. (Serhii Nikolsky)	41
Даценко С. С. (Serhii Datsenko)	87	Раскін Л. Г. (Lev Raskin)	63
Дашков Д. Є. (Dmytro Dashkov)	57	Рязанцев О.І. (Oleksandr Ryazantsev)	9
Дубницький В. Ю. (Valeriy Dubnitskiy)	28	Сальников Д. В. (Dmytro Salnikov)	92
Дунь Бао (Bao Dun)	68	Смірнов О. А. (Oleksii Smirnov)	49
Заковоротний О. Ю. (Oleksandr Zakovorotnyi) .	68	Смірнова Т. В. (Tetiana Smirnova)	49
Захарченко А. С. (Anastasiya Zakharchenko) .	75	Соколов Д. Д. (Dmytro Sokolov)	63
Караман Д. Г. (Dmytro Karaman)	92	Степанець О. В. (Oleksandr Stepanets)	75
Ковальчук Д. М. (Dmytro Kovalchuk)	41	Сухомлин Л. В. (Larysa Sukhomlyn)	63
Костяник І. В. (Iryna Kostianyk)	16	Ходирев О. І. (Alexander Khodyrev)	28
Кошман С. О. (Serhii Koshman)	41	Хорошун Г.М. (Ganna Khoroshun)	9
		Черпицький М.В. (Maksym Cherpitskyi)	9

Наукове видання

**Сучасні
інформаційні системи**

**Advanced
Information Systems**

Науковий журнал
Том 7, № 2

Відповідальний за випуск *О. Ю. Заковоротний*
Технічні редактори *Д. С. Гребенюк, А. П. Шиман*
Комп'ютерна верстка *Н. Г. Кучук*

Свідчення про державну реєстрацію КВ № 22522-12422Р від 13.01.2017 р.

Підписано до друку 06.06.2023 Формат 60×84/8. Ум.-друк. арк. 12,25. Тираж 100 прим. Зам. 606-23

Адреса редакції: Національний технічний університет "Харківський політехнічний інститут"
Кафедра КІП, вул. Кирпичова, 2, 61002, м. Харків, Україна, тел. + 38 (057) 707-61-65

Віддруковано з готових оригінал-макетів у друкарні ФОП Петров В.В.
Єдиний державний реєстр юридичних осіб та фізичних осіб-підприємців.
Запис № 2480000000106167 від 08.01.2009.

61144, м. Харків, вул. Гв. Широнінців, 79в, к. 137, тел. (057) 778-60-34
e-mail: bookfabrik@mail.ua

Problems of identification in information systems

UDC 004.942

doi: <https://doi.org/10.20998/2522-9052.2023.2.01>

Ulviyya Goyushova Malik

Institute of Control Systems of NAAS, Baku, Azerbaijan

ALGORITHMS FOR FINDING NON-INTERSECTING ROADS ON IMAGES

Abstract. Research relevance The use of GPS for UAVs to determine their current coordinates (navigation) leads to a number of problems. The development and improvement of the autonomous navigation method without the use of GPS is an actual issue, and this issue is considered in the article. **The subject** of study in the article is the issue of UAVs navigation based on video images. **The purpose** of work to develop an algorithm for autonomous navigation of UAVs that requires less resources. The following **tasks** are solved in the article: software modules are developed based on image processing algorithms; suitable model for invariant identification of curves was studied and applied. The following **results** were obtained: an algorithm was developed for the invariant identification of winding roads (curves) without intersections from different angles; developed software modules have been tested in the Republic of Azerbaijan on the basis of a map of the region called "Agsu Pass". **Conclusions:** the application of the method proposed in the article will enable rotation-invariant autonomous navigation of UAVs using less resources.

Keywords: Global Positioning System; unmanned aerial vehicle; navigation; autonomous; invariant; video; non-intersecting roads.

Introduction

The widespread use of unmanned aerial vehicles (UAVs) in modern times, both military and civilian, raises a number of issues. One such issue is the problem of locating, tracking and managing UAVs. These devices have different navigation methods. The following are some of the different navigation methods tested for UAVs.

One of the main approaches for navigating UAVs is the application of GPS technology. GPS - a global positioning system that operates continuously and transmits electromagnetic signals from position, speed and altitude meters to certain receivers. These signals allow you to determine the location of objects on the ground, the coordinates of the device with high accuracy. This technology, which is applied to the UAV, determines the location of the device and regulates its movement [1]. In [2], without human intervention, the UAV receives the current coordinates from the GPS receiver via a Bluetooth connection with the navigator computer. With this point, the optimal trajectory is drawn to the next destination. During the flight, the navigator computer provides information on which direction and how far to turn. This information is used in the control system of UAV engines.

To date, UAV navigation methods have been proposed in three main categories:

- 1) inertia navigation;
- 2) satellite navigation;
- 3) vision-based navigation. Visual-based navigation

from these methods provides richer information about the earth's surface and objects than others [3].

It is possible to navigate the UAV with GPS when flying in the open air, at an altitude without any obstacles. However, at lower altitudes, where obstacles may be present, this approach is not appropriate. [4] proposed optical stream navigation for autonomous movement of UAVs in such zones.

Problem statement

The main disadvantage of the widely used GPS method is that the operation of the UAV is completely dependent on GPS. Where there is no GPS, there are obstacles to the movement and control of these devices. For this reason, a GPS-independent control system is required for UAVs to be able to move autonomously.

Various methods have been proposed to regulate the autonomous movement of UAVs without the use of GPS. Some of the methods are shown in the introduction. One such method is video-based navigation. The main approach in this method is to determine the location of the objects based on the objects observed by the device and their coordinates. Observed facilities include important infrastructure such as power plants and roads. UAV recognizes these objects and determines its current coordinates and location based on their coordinates. The images and coordinates of the objects to be solved must be known in advance. This data is processed and placed in the database. The image obtained by UAV is also processed. The results obtained are compared with the data in the database to determine which of these objects corresponds to the database. The location of the device is determined by the coordinates of the image object.

Problem solving

Important infrastructure as reference objects is used to determine the location of the UAV. Roads were taken as such infrastructure in the study. Thus, the UAV must take a view of the earth's surface while driving without GPS and determine its approximate location relative to the road in the image. In this study, only non-separable, non-intersecting roads were considered.

In order to solve the problem, images of all the roads that the UAV can observe must first be obtained in advance. Certain features are determined by processing

these images. For winding, non-separable roads, such key features are considered to be the turning points of the roads. These features - the coordinates of the turning points are placed in the database as a factor indicating the appropriate path.

The proposed method for solving the problem is to compare the turning points of the road imagined during the flight with the turning points of the known roads. The UAV captures video of the earth's surface during movement. The obtained images are processed, ways are found. The turning points of the road are determined, compared with the coordinates placed in the database, and the appropriate one is selected.

With the application of this method, UAV can provide autonomous traffic on roads without GPS. However, there are a number of problems here. During the research, it was found that UAV can capture video in different forms: at different distances, at different angles, in different directions and scales. In this case, it is known that the road image of the UAV will be processed differently from the road image in the database and different features will be obtained. As a result, although both images represent the same path, they will not match or will match with another image in the database. The problem of invariance was considered to solve this problem.

Each of the roads can be identified as a curve, as the study considers only non-separable, non-intersecting and mostly curved roads. There are different methods for identifying curves: interpolation polynomials, Bezier curves, etc. However, the B-spline model was chosen to identify the curves in the study. The main reason for choosing this method is that the B-spline is resistant to operations such as rotation and scaling. Application of the B-spline method provides invariant identification of curves [5].

Application of b-spline:

$$Q(u) = \sum_{i=0}^n P_i N_{i,k}(u) . \quad (1)$$

In this equation u is a parameter of the function representing the curve $u = n - k + 2$;

n - the number of nodes is determined in Fig. 1;

$P_i(x_i, y_i)$ - control points;

$N_{i,k}(u)$ - basis function;

k is the number of points on each segment where the curve is divided. This number satisfies the condition $2 \leq k \leq n + 1$.

The basis function $N_{i,k}(u)$ is the basic of the construction of the B-spline. The procedure for setting up basis functions is given in (2).

$$N_{i,k}(u) = \frac{u - t_i}{t_{i+k} - t_i} \cdot N_{i,k-1}(u) + \frac{t_{i+k+1} - u}{t_{i+k+1} - t_{i+1}} \cdot N_{i+1,k-1}(u), \quad (2)$$

$$t_i = 0, 1, 2, \dots, n + k$$

As can be seen from the construction of the base functions, these functions do not depend on the control

points. The B-spline model is applied to both base images and road images obtained by the UAV.

Image processing. As mentioned, the placement of pre-known images in the database means the identification and placement of certain features from these images. Comparison between images also means comparing their respective features. The image is processed to obtain these features from the images and to establish a suitable spline. The main objects - roads - should be identified from the processed images and the image should be cleaned of "dirty", non-informative data. Image processing involves a few steps. These steps are applied to the image of the so-called "Agsu Pass" in Azerbaijan, taken from the Google Earth platform Fig. 1. The sequence of these steps is shown below.



Fig. 1. Taken image of "Agsu Pass"

Converting a color image to a grayscale image.

The images are usually rendered in RGB (red, green, blue) format. Due to the complexity of working with color images, these images turn into grayscale images. The transition to a grayscale image is an effective and important step in terms of speed, ease of coding, etc. in the subsequent processing stages. The procedure for converting an RGB image to a grayscale image is given below.

$$0.3 \cdot R + 0.59 \cdot G + 0.11 \cdot B . \quad (3)$$

In this formula, RGB represents the red, green, and blue color distribution in the image pixel, respectively [6]. The grayscale image of the captured image is shown in Fig. 2.

Defining the edges of objects in the image. At this stage, the Sobel Edge Detection algorithm was applied. Using this algorithm, unnecessary pixels are thrown out of the image. Defining the edges of objects in the image brings the image to the black and white format, and at the end of the step only black (or white) pixels play an informative role [7]. The result of applying the algorithm is shown in Fig. 3.



Fig. 2. Grayscale image

Clearing noise pixels from images. The number of pixels required in the black-and-white image obtained with the application of Edge Detection algorithms is slightly reduced. Based on the Connected Component Label algorithm, other additional pixels in the image are deleted, and as a result, the image consists mainly of roads [8]. The result of applying the algorithm is shown in Fig. 4.



Fig. 3. Binary image



Fig. 4. Clean image

Thinning. At the end of the processing, the image should be in black and white, consisting only of a curve. Despite the processing of the image, the curve in the final image may not be of the desired thickness. In the final stage, the curve in the image - the path should be taken as a sequence of pixels, not as a set of pixels. The construction of this sequence is done by checking the adjacent pixels of each pixel. To build this sequence for thicker curves is more difficult and impossible. For this reason, it is necessary to reduce the thickness of the curve

to 1 pixel. The Zhang-suen Thinning algorithm is used as a thinning algorithm in the research [9]. The essence of this algorithm is to check the adjacent black pixels of each black pixel. As a result, the object - the curve in the image is thinned to 1 pixel, and the number of black pixels is sharply reduced. Any curve taken in Fig. 4, the result after the thinning is shown in Fig. 6.

With the connected component label approach, the set of black pixels that represent the curve is brought to the pixel sequence. Properties - turning points are determined in a certain method from the sequence of obtained pixels.

Each of these steps is applied to known images and the image obtained by the UAV. The characteristics of the known images are placed in the database, and the characteristics of the image obtained by the UAV are compared with those in the database. The corresponding one from the known images is determined. The current position of the UAV is determined based on the coordinates and position of this road.

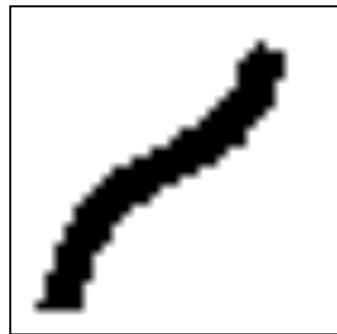


Fig. 5. Any curve taken

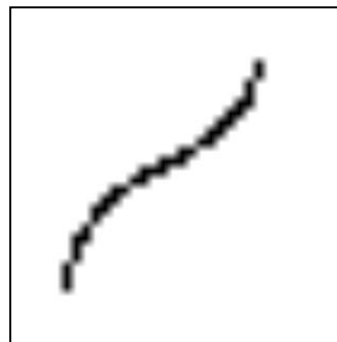


Fig. 6. The curve after the thinning

Aspects (opportunities) of practical application

The proposed method can be used as a navigation method for UAVs. The main essence of this method is that navigation is based on images. The main advantages of the method:

1. Having the ability to navigate the device without depending on GPS,
2. It requires quite a few resources during navigation.

The main limitations and disadvantages when applying the method are as follows:

1. The recognition process will not work in unfavorable weather conditions such as fog, rain, or any obstruction that restricts vision.

2. The proposed method can be applied only for winding roads without intersections.

Discussion of the results

To date, several navigation methods for UAVs have been reported in the literature. Most of these methods rely on the use of GPS. However, in environments where the use of GPS is restricted or there are external interferences to GPS, problems may arise when applying those methods. In such cases, it is more efficient to use the navigation method based on video images. There are also video-based navigation methods in the literature.

The advantage of the video-based navigation method proposed in the article over other video-based methods is that it requires less resources. Carrying out extensive processing on the images has significantly

reduced the number of points that play a role in the end. This has led to the requirement of less time and memory.

Conclusions

The article develops a method of navigating objects encountered in video images to ensure the autonomous movement of the UAV without the use of GPS. Non-separable, non-intersecting highways are taken as reference objects to determine the location of the UAV. This method was tested in the Republic of Azerbaijan on the basis of a picture taken of the region called "Agsu Pass". The proposed method can be used to ensure the autonomous movement and control of UAV in places where there are restrictions on the use of GPS. The application of the method is possible both in military zones and in other areas.

REFERENCES

1. (2008), *Global positioning system wide area augmentation system (WAAS) performance standard*, United States Department of Transportation; Federal Aviation Administration, 60 p., available at: <https://rosap.ntl.bts.gov/view/dot/16922>
2. Tae Hyun, Kim & Denny, Toazza (2009), *Navigation Control of an Unmanned Aerial Vehicle (UAV)*, Bachelor Thesis in Computer and Electrical Engineering, Sweden: School of Information Science, Computer and Electrical Engineering Halmstad University, available at: https://www.academia.edu/9237265/Navigation_Control_of_an_Unmanned_Aerial_Vehicle_UAV
3. Yuncheng, Lu, Zhucun, Xue, Gui-Song, Xia & Liangpei, Zhang (2018), "A survey on vision-based UAV navigation", *Geospatial Information Science*, Vol. 21(1), pp. 21-32, doi: <https://doi.org/10.1080/10095020.2017.1420509>
4. Aasish, C, Ranjitha, E., Razeen, Ridhwan U, Bharath, Raj S & Angelin, Jemi L. (2015), "Navigation of UAV without GPS", *Int. Conf. on Robotics, Automation, Control and Embedded Systems (RACE)*, Chennai, doi: <https://doi.org/10.1109/RACE.2015.7097260>
5. Sandeep, Pande & Manna, Chetty (2019), "Position Invariant Spline Curve Based Image Retrieval Using Control Points", *Int. Journal of Intelligent Engineering and Systems*, Vol. 12(4), pp. 177-191, doi: <https://doi.org/10.22266/ijes2019.0831.17>
6. Saravanan, C. (2010), "Color Image to Grayscale Image Conversion", *2010 Second International Conference on Computer Engineering and Applications*, pp. 196-199, Bali, Indonesia, ISBN 978-1-4244-6080-9.
7. Muthukrishnan, R & Radha M. (2011), "Edge Detection Techniques For Image Segmentation", *International Journal of Computer Science & Information Technology (IJCSIT)*, Vol. 3(6), pp. 259-267, doi: <https://doi.org/10.5121/ijcsit.2011.3620>
8. Reinhard, Klette (2014), *Concise Computer Vision*, Springer, doi: <https://doi.org/10.1007/978-1-4471-6320-6>
9. Paramjeet, Kaur & Harish, Rohil (2011), "A Comparative Study of Iterative Thinning Algorithms for BMP Images", *IJCSIT*, Vol. 2(5), pp. 2375 -2379, available at: <https://ijcsit.com/docs/Volume%202/vol2issue5/ijcsit20110205114.pdf>
10. Made, Sudarma & Ni Putu, Sutramiani (2014), "The Thinning Zhang-Suen Application Method in the Image of Balinese Scripts on the Papyrus", *International Journal of Computer Applications*, Vol. 91(1), pp. 9-13, doi: <https://doi.org/10.5120/15844-4726>

Надійшла (received) 04.02.2023

Прийнята до друку (accepted for publication) 26.04.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Ульвія Гоюшова Малік – молодший науковий співробітник Інституту систем управління Національної академії наук Азербайджану, Баку, Азербайджан;

Ulviyya Goyushova Malik – Junior scientific researcher at the Institute of Control Systems of NAAS, Baku, Azerbaijan; e-mail: kadaseva.ulviyyee@gmail.com; ORCID ID: <http://orcid.org/0000-0002-3646-4937>.

Алгоритми пошуку пересічних доріг на зображеннях

У. М. Гоюшова

Анотація. Дослідження присвячене розгляду питань забезпечення автономного руху безпілотних літальних апаратів. **Актуальність дослідження.** Використання GPS для БПЛА для визначення їх поточних координат (навігація) призводить до ряду проблем. Розробка та вдосконалення методу автономної навігації без використання GPS є актуальною проблемою, і це питання розглядається в статті. **Предметом дослідження** в статті є питання навігації БПЛА на основі відеозображень. **Метою роботи** є розробка алгоритму автономної навігації БПЛА, що вимагає менших ресурсів. У статті вирішуються **наступні завдання**: розроблено програмні модулі на основі алгоритмів обробки зображень; вивчена та застосована відповідна модель інваріантної ідентифікації кривих. Отримано **наступні результати**. Розроблено алгоритм інваріантної ідентифікації звивистих доріг (кривих) без перетинів з різних кутів. Алгоритм Zhang-Suen Thinning був застосований для зменшення товщини кривих на зображеннях до 1 пікселя. Розроблені програмні модулі пройшли випробування в Азербайджанській Республіці на основі карти регіону під назвою «Агсуський перевал». **Висновки:** застосування запропонованого в статті методу дозволить здійснювати ротаційно-інваріантну автономну навігацію БПЛА з меншими ресурсами.

Ключові слова: глобальна система позиціонування; безпілотний літальний апарат; навігація; автономний; незмінний; відео; дороги, що не перетинаються.

Г. М. Хорошун, О. І. Рязанцев, М. В. Черпіцький

Східноукраїнський національний університет імені Володимира Даля, Київ, Україна

КЛАСТЕРИЗАЦІЯ ТА АНОМАЛЬНІСТЬ ДАНИХ ІНДЕКСУ ВОЛАТИЛЬНОСТІ ФОНДОВОГО РИНКУ США

Анотація. **Актуальність.** Інвестування грошей це важливий спосіб покращення фінансового стану, як окремої людини, так і суспільства в цілому. Актуальною є проблема розуміння фінансових даних та прийняття рішень щодо інвестицій грошей в певний проект на даний момент часу. **Об'єктом** дослідження є процес встановлення залежності вартості активу від часу. **Предметом** дослідження є математичні моделі обробки даних, проведення кластеризації даних та пошуку аномалій. **Метою** даної роботи є розробка методу для ефективного інвестуванню грошей з використанням методів обробки та аналізу даних до значень індексу волатильності фондового ринку США, визначення кластерів по діям з активами, а також перевірка наявності аномальних даних. **Результати досліджень.** Обрані офіційні дані значень індексу волатильності та підготовлені для подальшого аналізу шляхом видалення неповних наборів та подальшої нормалізації. Проведена кластеризація часових рядів та розділено масив на п'ять однорідних груп. Кластери визначають діапазони індексу волатильності, що відображають різні настрої інвесторів на ринку та спонукають до відповідних дій з активами: продавати, очікувати, купляти, виводити гроші з проектів, що розвиваються і вкладати в стабільні, перечекаати. Програмно проведено сегментацію даних, застосування віконної функції, визначено центроїди для сегментів та проведено реконструкцію сигналу. Визначено точки аномалій даних. Проведено порівняльний аналіз за результатами побудованих початкових даних, реконструйованих та похибки реконструювання.

Ключові слова: кластеризація; аномальність; індекс волатильності.

Постановка та аналіз проблеми

В сучасному світі надлишок грошей в бюджеті підприємства або сім'ї прийнято інвестувати в бізнес, нерухомість та себе. Інвестиціями є всі види майнових та інтелектуальних цінностей, що вкладаються в об'єкти підприємницької та інших видів діяльності, в результаті якої створюється прибуток (доход) та/або досягається соціальний та екологічний ефект [1].

Існує багато показників в управлінні фінансовими ризиками, які використовуються для аналізу ринку. Одним з таких статистичних фінансових показників є індекс волатильності (VIX), що вказує на очікувані коливання вартості активу протягом наступних 30 днів [2]. Такий часовий період передбачення звужує світогляд до близького терміну. Індекс волатильності є показником настроїв інвесторів та їх прогнозів щодо волатильності основних котирувань активів з часом.

Отже, **метою** даної роботи є розробка методу для ефективного інвестуванню грошей з використанням методів обробки та аналізу даних для значень індексу волатильності.

Для цього пропонується розраховувати індекс волатильності, встановлюючи залежність вартості активу від часу. Для визначення трендів в однорідних групах даних корисно провести кластеризацію даних, а також визначити точки або області аномалій, що можуть бути помилковими даними або презентувати зародження нового кластеру.

Для роботи в галузі дослідження даних [3-7] нам необхідно обрати достовірні джерела інформації та обробити дані таким чином, щоб мати достатню кількість однотипних даних за обраний час з певною періодичністю та зберегти у зручному форматі для обробки. В цій роботі планується обробка економічних даних, тому ми звернулися до FRED — онлайн бази даних, що складається з сотень тисяч економічних

часових рядів даних з оцінок національних, міжнародних, державних та приватних джерел [8]. FRED створений і підтримується дослідним відділом при Федеральному резервному банку Св. Луїса. Корисною є можливість використання бази ALFRED (Archival Federal Reserve Economic Data), яка пропонує користувачам можливість доступу до старих даних для багатьох доступних серій FRED. По суті, ALFRED та FRED допомагають користувачам розповісти свої історії даних.

Теоретичні відомості

Індекс волатильності. Волатильність це статистичний показник змінювання цін, який є мірою ризику використання фінансового інструменту в даний період часу. Для розрахунку волатильності використовується статистична стандартна статистика відхилення, що дозволяє інвесторам визначити ризик придбання фінансового інструменту.

Для розрахунку індексу волатильності нам необхідно попередньо розрахувати стандартне відхилення σ_{sd} для вибірки значень ціни:

$$\sigma_{sd} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}. \quad (1)$$

В формулі (1) n – розмір вибірки; x_i – розмір індивідуального значення вибірки; \bar{x} – середнє арифметичне вибірки.

Найбільш часто розраховується середньорічна волатильність σ за формулою:

$$\sigma = \frac{\sigma_{sd}}{\sqrt{P}}, \quad (2)$$

де P – часовий проміжок у роках.

Волатильність σ_T на інтервал часу T (обчислюється в днях) розраховується на підставі річної середньої волатильності таким чином:

$$\sigma_T = \sigma\sqrt{T}. \quad (3)$$

Підготовка даних. Для аналізу даних необхідно перевірити їх на відповідність до загальних вимог, забезпечити їх повноту та логічність. Трансформація даних, наприклад нормалізація, може покращити точність та ефективність аналізу даних за допомогою алгоритмів машинного навчання. Такі методи дають кращі результати, якщо дані будуть нормалізовані, тобто масштабовані до певних діапазонів. Атрибут нормалізується шляхом масштабування його значень таким чином, щоб вони зменшувалися у невеликому заданому діапазоні, наприклад, від -1 до 1. Оберемо метод мін-макс нормалізації, яка виконує лінійне перетворення вхідних даних x у вихідні дані x_{new} та може бути обчислена за формулою:

$$V_{new} = \frac{(V - V_{min})}{(V_{max} - V_{min})}. \quad (4)$$

У виразі (4) V_{min} та V_{max} є відповідне мінімальне та максимальне значення змінної x .

Кластеризація даних часових рядів. За допомогою кластерного аналізу [9] можна зробити вибірку за ознакою, що вивчатиметься. Його основна задача полягає в розділенні багатовимірною масиву на однорідні групи. Парний коефіцієнт кореляції або Евклідова відстань між об'єктами за заданим параметром використовується як критерій групування. Найближчі значення групуються разом в єдиний кластер.

Кластерний аналіз включає наступні кроки. Вибір даних або їх перетворення для створення корисних і нових функцій. Розробка або вибір алгоритму кластеризації. Валідація кластера необхідна завдяки тому, що різні підходи зазвичай призводять до формування різних груп кластерів і навіть для одного алгоритму, параметри ідентифікації або послідовність введення шаблонів може вплинути на фінальні результати. Інтерпретація результатів дозволяє надати значущу інформацію з вихідних даних завдяки проведеної кластеризації.

Метод, який ми будемо використовувати, називається кластеризацією середніх або k -means. Алгоритм k -means кластеризує дані, намагаючись розділити вибірку на n груп однакової дисперсії, мінімізуючи критерій, відомий як інерція або сума квадратів всередині кластера. Цей алгоритм вимагає завдання кількості кластерів, які потрібно вказати. Він добре масштабується до великої кількості зразків і використовувався в багатьох областях.

Алгоритм k -means ділить набір з N зразків X на K непересічних кластерів C , кожен з яких описується середнім μ_i зразків у кластері, $\mu_i \in C$. Середнє зазвичай називають «центроїдами» кластера, і вони зазвичай не є точками з X , хоча вони знаходяться в одному просторі. Алгоритм k -means спрямований на вибір центроїдів, які мінімізуються критерій інерції або суми квадратів кластера:

$$\sum_{i=0}^n \min (\|x_j - \mu_i\|^2).$$

Інерція може сприйматися, як міра того, наскільки внутрішньо когерентними є кластери.

Методи пошуку аномалій. Загальне завдання виявлення аномалій у часових рядах часто поділяється на дві окремі задачі: виявлення викидів (Outlier Detection) та «нової поведінки» (Novelty Detection). Викиди є наслідком: помилок у даних (неточності вимірювання, округлення, неправильного запису тощо), наявності шумових об'єктів (неправильно класифікованих об'єктів), присутності об'єктів «інших» вибірок.

Для пошуку аномалій застосовуються статистичні методи для вивчення окремих ознак та відокремлення екстремальних та унікальних значень у вибірці, а саме: Z -оцінка та надлишковий ексцес. Також в роботі з визначення аномалій використовують модельний підхід, метричні, ітераційні методи та методи машинного навчання.

Виявлення аномалій включає: виявлення аномалій шляхом пошуку будь-яких значень за певним порогом; виявлення аномалій за структурою сигналу; більш тонкі похибки - зміна форми періодичної хвилі.

Результати роботи з кластеризації даних та їх аномальності. Використовуються дані у вигляді часових рядів для індексу волатильності, отриманих з бази FRED з 02.01.1990 по 29.09.2022 з періодичністю в 1 день. Вивчаються дані волатильності фондового індексу США S&P 500 [10], що містить інформацію про 500 найпотужніших компаній. Графічне представлення даних індексу волатильності VIX в періоди стабільності та панічних настроїв інвесторів наведений на рис. 1.

Подальший аналіз даних проведемо за допомогою статистичних методів. Гістограма індексу волатильності початкових даних (рис. 2а) вказує на значну наявність нульових або пропущених даних. Також, ми бачимо, що після індексу волатильності 45 кількість таких подій значно зменшується. Отже, зменшемо кількість випадків подій до 20 та видалимо пропущені дані (рис. 2б). Згідно до графіку, мінімальний індекс волатильності становить 9 % і буде прийнятий нами за нижню межу при проведенні кластеризації, а 83% - за вищу межу.

Кластеризація. Згідно з проведеного статистичного аналізу та існуючого тлумачення значень індексу волатильності VIX можна створити для інтерпретації результатів обчислень наступних 5 кластерів. Введемо такі рекомендації до кластерів, що визначають діапазони індексу волатильності та відображають різні настрої інвесторів на ринку та спонукають до відповідних дій з активами: перший – продавати, другий – очікувати, третій – купляти, четвертий – виводити гроші з проектів, що розвиваються і вкладати в стабільні, п'ятий – перерекетати на зменшення індексу.

1. 9-20% - низька волатильність. Це свідчить про оптимістичний настрій учасників ринку. Чим менше значення показника падає, тим більша ймовірність швидкого реверсивного руху тренду і його реверсивного. Це часто є сигналом для інвестора продавати цінні папери і близькі позиції. У цьому випадку важливо зафіксувати прибуток до початку неминучого перелому.

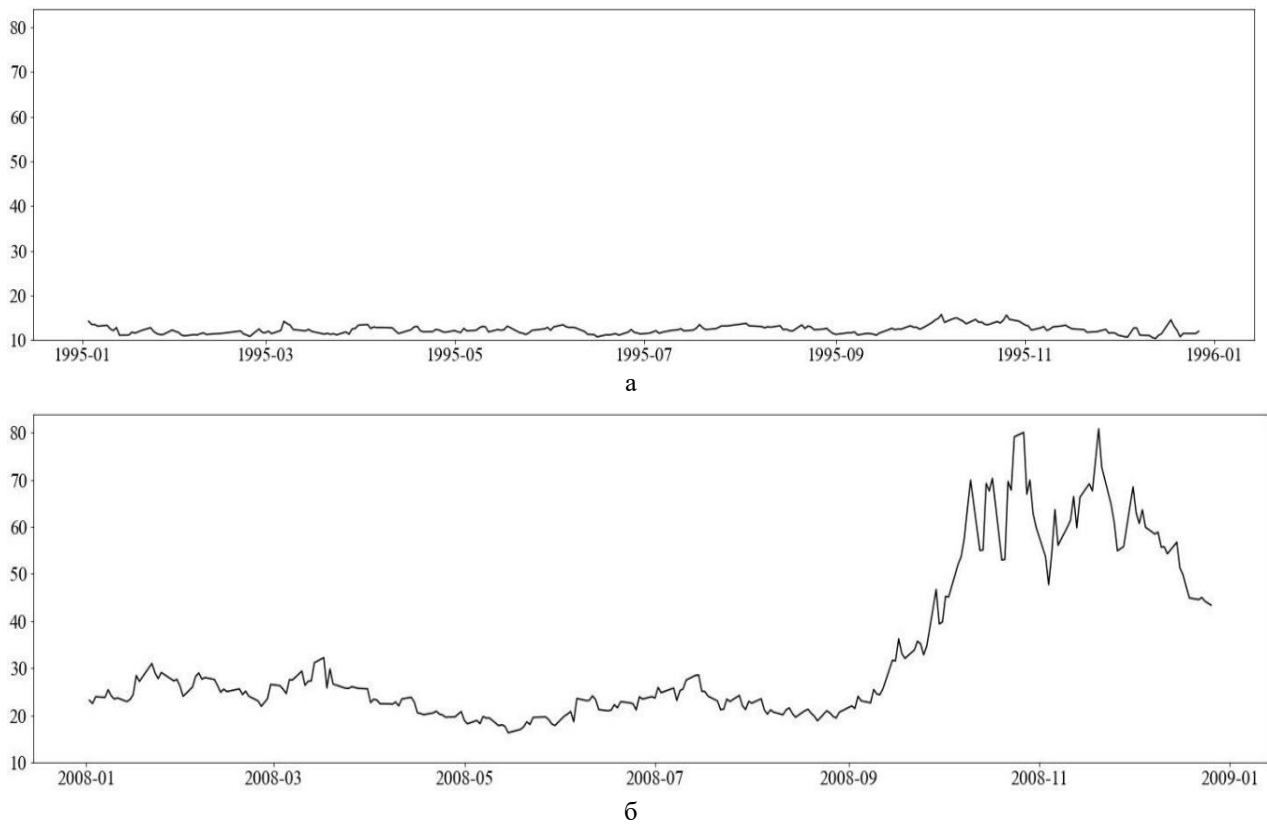


Рис. 1. Графічне представлення даних індексу волатильності VIX за стабільні (а) та передпанічні і панічні періоди часу (б)

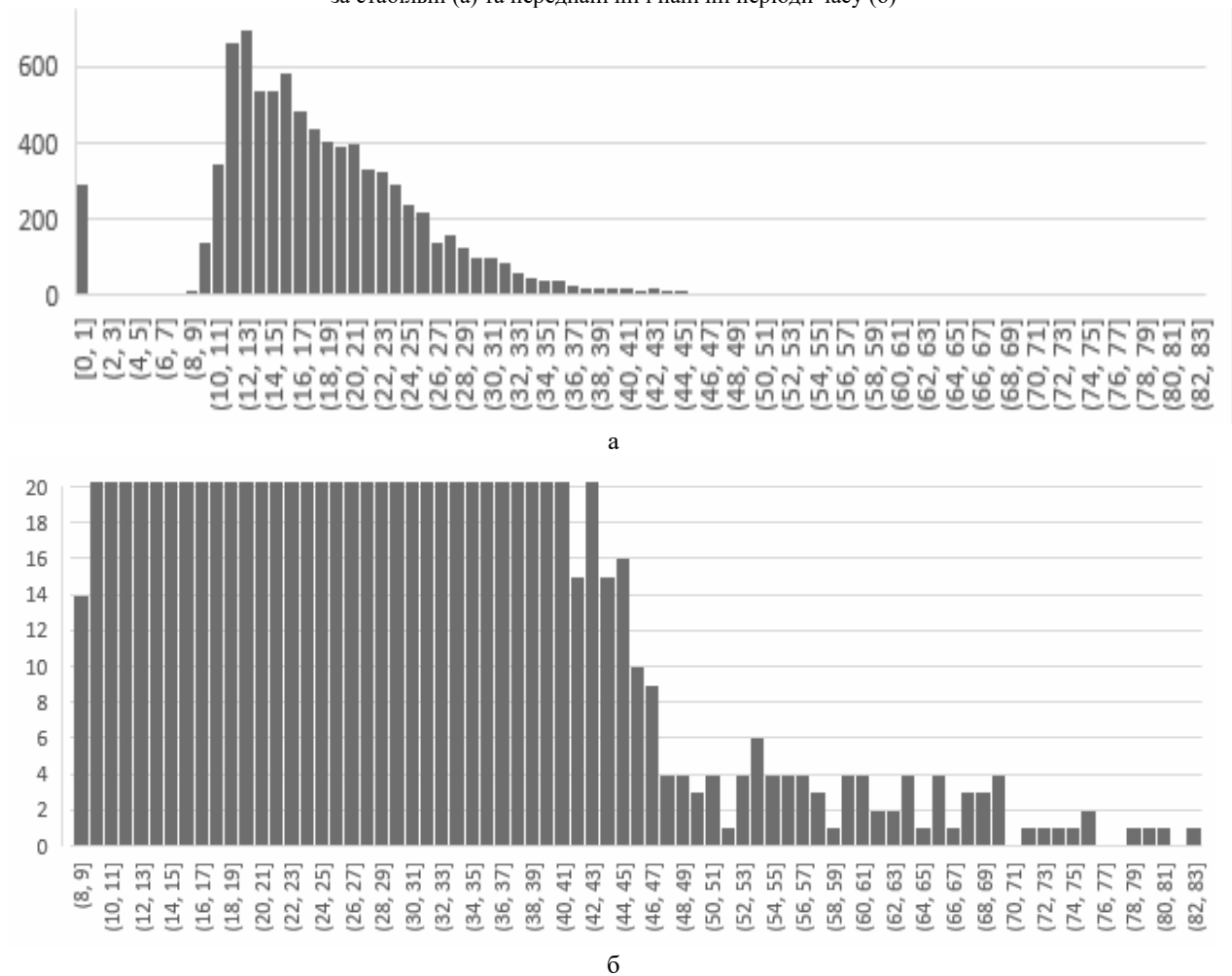


Рис. 2. Гістограми індексу волатильності початкових даних (а) та без пропущених даних з максимальним значенням (б)

2. 21-35% - середня волатильність. Коливання в межах цього діапазону не можуть дати інвестору будь-яких ознак дії.

3. 36-45% - ознака паніки на ринку. Такий стан справ зазвичай супроводжується швидким крахом цін на акції. Це сигнал для інвестора для пошуку точки входу на ринок. Після того, як температура знижується, ціна акцій знову підніметься. Тому це найкращий час для купівлі цінних паперів.

4. 46-59% - ознака великої паніки на ринку.

5. 60-83% – ознака кризи на ринку. Як правило, індекс стрімко збільшується вгору.

Аномальність даних. Для подальшого аналізу та пошуку аномалії в даних оберемо дані за роки відносно стабільності ринку. Кластерний аналіз використовується для визначення аномалії як певної моделі хвилі, яка раніше не спостерігалася. Створена бібліотека нормальних форм сигналу використовується для спроби реконструювати форму сигналу для тестування. Якщо реконструкція погана, то форма хвилі, ймовірно, містить щось ненормальне і, отже, є аномалією. Виявлення аномалій проводиться за наступним алгоритмом.

1. Розділити форму хвилі на сегменти по n зразків.

2. Застосувати віконну функцію до даних, яка приводить початок і кінець сигналу до нуля. Сформувати простір в n вимірі, де кожен сегмент представляє одну точку.

3. Провести кластеризацію точок сегментів і визначити центроїди кластерів. Центроїди кластерів забезпечують бібліотеку нормальних форм хвилі.

4. Реконструювати форму хвилі для тестування з використанням кластерних центроїдів, отриманих під час навчання.

5. Визначити області аномалій даних.

Проведемо сегментацію даних за 2012-2014 роки на 235 частин, з яких 7 сегментів наведені у якості прикладу на рис. 3.

Сегментовані фрагменти даних необхідно обробити за допомогою віконної функції. В якості віконної функції обрано частину синусоїди, яка наведена на рис. 4. Подальша обробка даних призводить до трансформації даних у вигляді на рис. 5.

Програмно визначено центроїди для сегментів та проведено їх реконструкцію. Результат реконструйованого сигналу наведений суцільним сірим кольором на рис. 6. Також на ньому легко провести порівняльний аналіз з початковими даними та похибкою реконструкції.

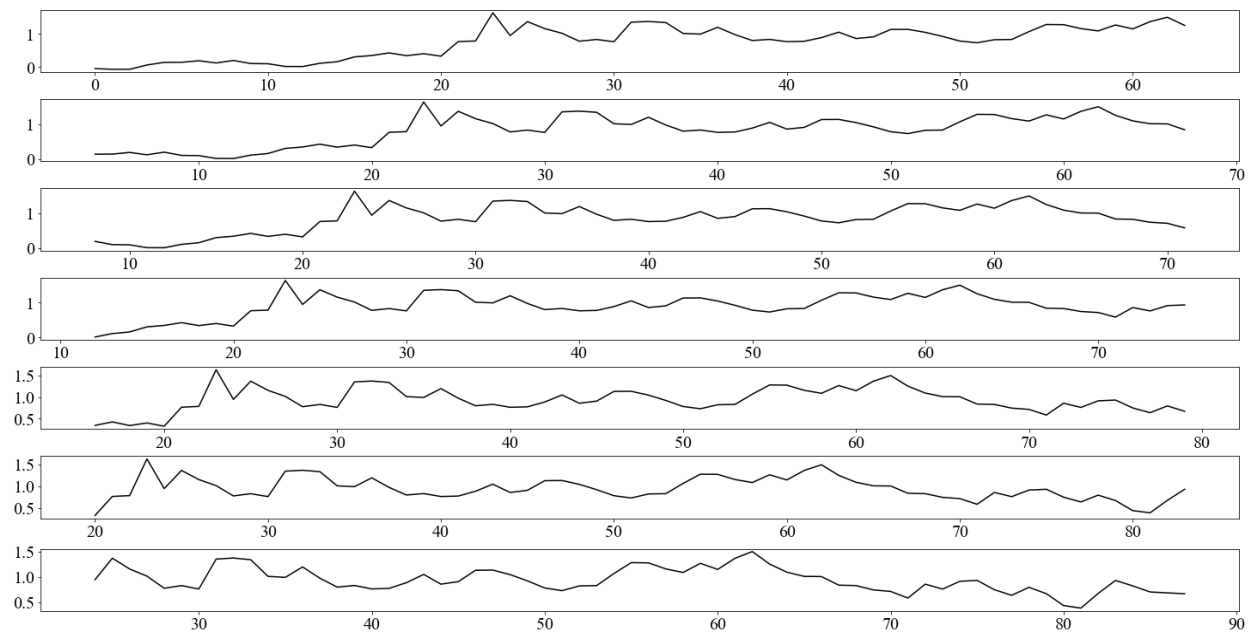


Рис. 3. Сегментовані фрагменти даних

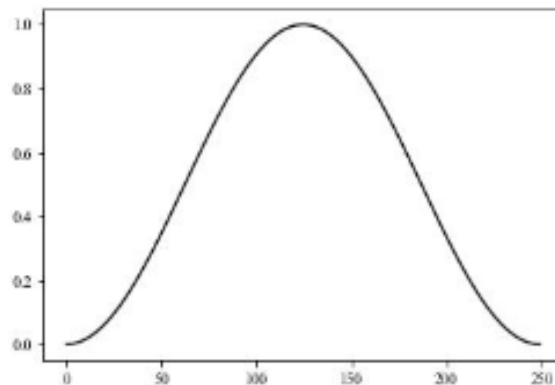


Рис. 4. Віконна функція для обробки даних

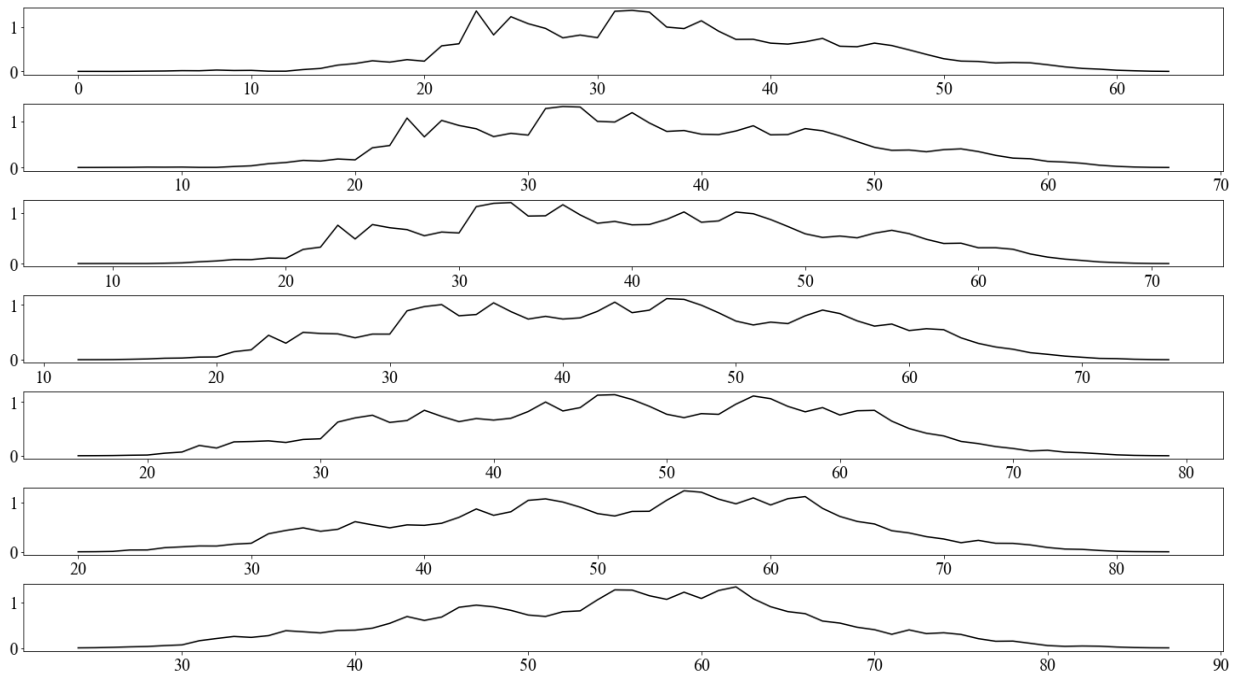


Рис. 5. Сегментовані дані після проходження віконної функції

Візуалізації початкових та оброблених даних наведена у графічному форматі з використанням різних типів та кольорів ліній. Як ми бачимо з графіків сигнали співпадають доволі добре на всьому протязі осі X.

Можна визначити області аномалій для даних, які в нашому випадку є точками з аномальними значеннями. Ці точки змінюють форму сигналу настільки що різниця між початковими даними та ре-

конструйованими більше ніж 0.1 по модулю. Різниця є похибкою реконструкції і позначена на рис. 6 та 7 суцільним сірим кольором.

Форма основного сигналу може значно збільшуватися або зменшуватися у порівнянні з реконструйованим сигналом.

Аномальні точки позначені колами на рис.6 можуть вказувати на непередбачуваність фінансових процесів, а також на появу нових трендів.

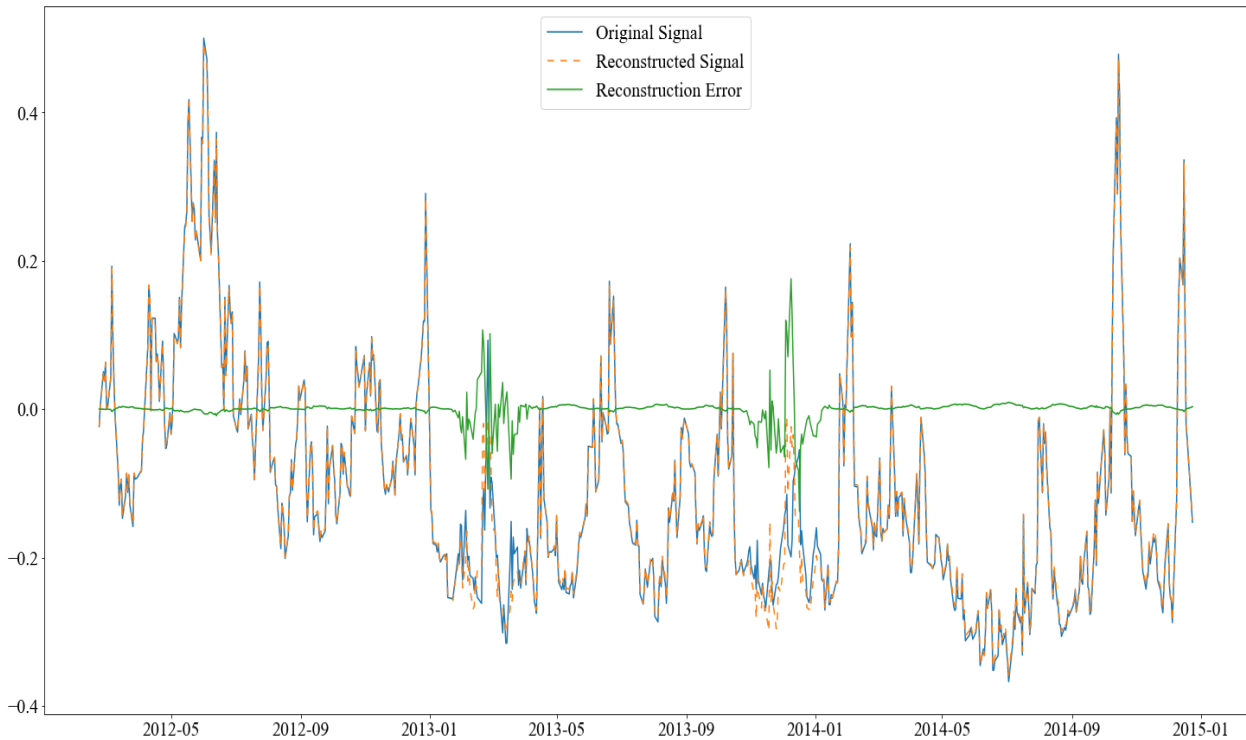


Рис. 6. Візуалізація фрагменту початкових даних (original signal) наведено суцільним чорним кольором, реконструйовані дані (reconstructed signal) пунктирною лінією, похибка реконструкції (reconstruction error) показана суцільним сірим кольором

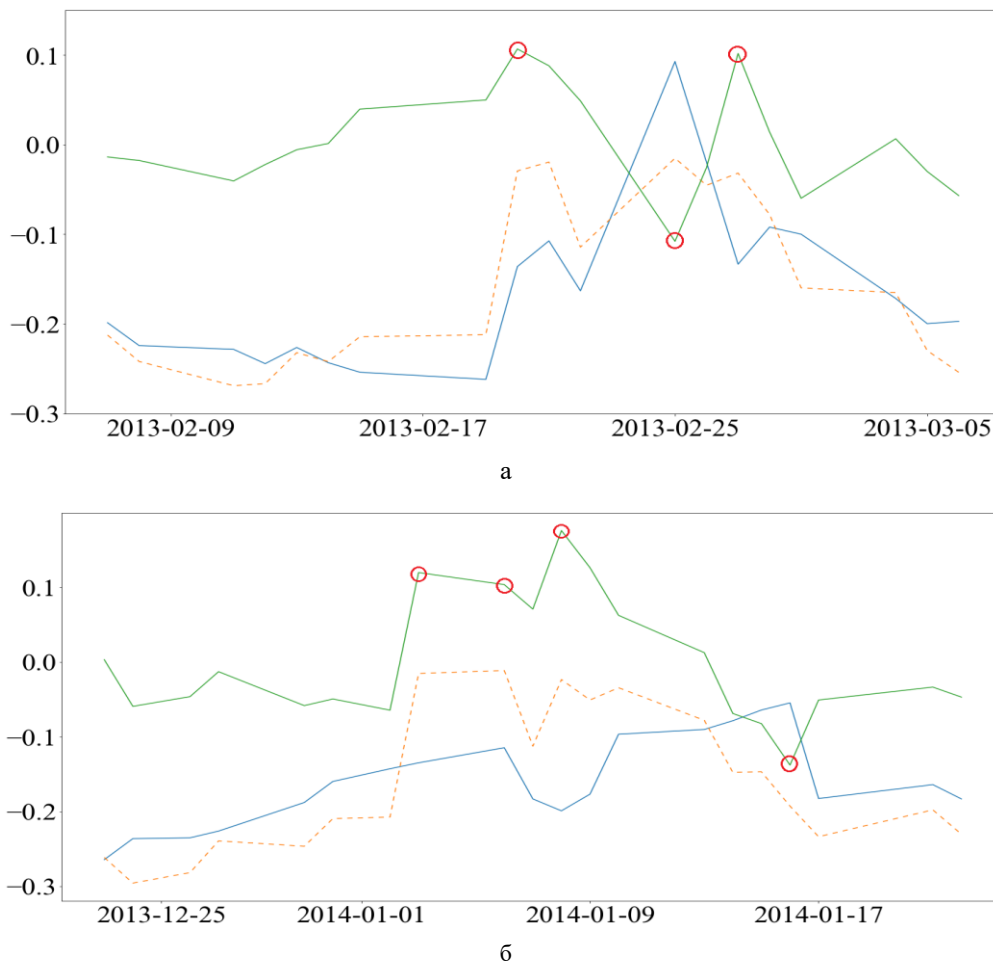


Рис. 7. Визначення аномальних даних, позначених колом, за перевищенням порогового значення похибки реконструкції

Висновки

Роботу присвячено аналізу даних по темі інвестування грошей з використанням індексу волатильності, який встановлює залежність вартості активу від часу. За допомогою статистичних методів проведено первісну обробку даних, проведено кластеризацію даних на п'ять кластерів, які відображають рі-

зні настрої інвесторів на ринку та спонукають до відповідних дій з активами. Також визначено точки аномалій даних завдяки кластерному аналізу та алгоритму реконструкції сигналу. Програмно визначено центроїди для сегментів та проведено їх реконструкцію. Проведено порівняльний аналіз за результатами побудованих початкових даних, реконструйованих та похибки реконструювання.

REFERENCES

- (2022), *On investment activity*, Law of Ukraine dated September 18, 1991 No. 1560-XII: as of October 10, 2022, available at: <https://zakon.rada.gov.ua/laws/show/1560-12#Text>
- (2023), *SPX / S&P 500 Index Overview | MarketWatch*, URL: <https://www.marketwatch.com/investing/index/spx>
- Wolf, Andrew (2022), *Machine Learning Simplified: A gentle introduction to supervised learning*, 199 p., available at: <https://themlsbook.com>
- Maheshwari, A. (2014), "Business intelligence and data mining", *Business Expert Press*, available at: <https://www.amazon.com/Business-Intelligence-Data-Mining-Analytics/dp/1631571206>
- Yang, X. S. (2019), *Introduction to Algorithms for Data Mining and Machine Learning*, Academic Press, doi: <https://doi.org/10.1016/C2018-0-02034-4>
- Fernandes, M. (2008), *Statistics for business and economics*, Bookboon, available at: <https://bookboon.com/en/statistics-for-business-and-economics-ebook?mediaType=ebook>
- Vercellis, C. (2009), *Business intelligence: data mining and optimization for decision making*, Wiley, New York, 420 p., available at: <https://www.amazon.com/Business-Intelligence-Mining-Optimization-Decision/dp/0470511397>
- (2023), *What is FRED? | Getting To Know FRED*, available at: <https://fredhelp.stlouisfed.org/fred/about/about-fred/what-is-fred/>
- Smith, L. I. (2002), *A tutorial on Principal Components Analysis*, Computer Science Technical Report No. OUCS-2002-12, available at: <http://hdl.handle.net/10523/7534>
- (2023), Chicago Board Options Exchange, CBOE Volatility Index: VIX [VIXCLS], retrieved from FRED, Federal Reserve Bank of St. Louis; *Economic Research*, February 3, 2023, available at: <https://fred.stlouisfed.org/series/VIXCLS>

СПИСОК ЛІТЕРАТУРИ

1. Про інвестиційну діяльність: Закон України від 18.09.1991 № 1560-XII : станом на 10 жовт. 2022 р. URL: <https://zakon.rada.gov.ua/laws/show/1560-12#Text>
2. SPX | S&P 500 Index Overview | MarketWatch. URL: <https://www.marketwatch.com/investing/index/spx>
3. Wolf Andrew. Machine Learning Simplified: A gentle introduction to supervised learning. 2022. 199 p. URL: <https://themlsbook.com>
4. Maheshwari A. Business intelligence and data mining. *Business Expert Press*, 2014. URL: <https://www.amazon.com/Business-Intelligence-Data-Mining-Analytics/dp/1631571206>
5. Yang X. S. Introduction to Algorithms for Data Mining and Machine Learning. Academic Press, 2019. Doi: <https://doi.org/10.1016/C2018-0-02034-4>
6. Fernandes M. Statistics for business and economics. Bookboon, 2008. URL: <https://bookboon.com/en/statistics-for-business-and-economics-ebook?mediaType=ebook>
7. Vercellis C. Business intelligence: data mining and optimization for decision making. New York : Wiley, 2009. 420 p. URL: <https://www.amazon.com/Business-Intelligence-Mining-Optimization-Decision/dp/0470511397>
8. What is FRED? | Getting To Know FRED. URL: <https://fredhelp.stlouisfed.org/fred/about/about-fred/what-is-fred/>
9. Smith L. I. A tutorial on Principal Components Analysis (Computer Science Technical Report No. OUCS-2002-12). 2002. Retrieved from <http://hdl.handle.net/10523/7534>
10. Chicago Board Options Exchange, CBOE Volatility Index: VIX [VIXCLS], retrieved from FRED, Federal Reserve Bank of St. Louis. *Economic Research*. February 3, 2023. URL: <https://fred.stlouisfed.org/series/VIXCLS>

Надійшла (received) 10.03.2023

Прийнята до друку (accepted for publication) 17.05.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Хорошун Ганна Миколаївна – кандидат фізико-математичних наук, доцент, доцент кафедри комп'ютерних наук та інженерії, Східноукраїнський національний університет імені Володимира Даля, Київ, Україна;

Ganna Khoroshun – candidate of physical and mathematical sciences, Associated Professor, Associated Professor of Computer Science and Engineering Department, Volodymyr Dahl East Ukrainian National University, Kyiv, Ukraine; e-mail: an_khor@i.ua; ORCID ID: <http://orcid.org/0000-0002-1272-1222>.

Рязанцев Олександр Іванович – доктор технічних наук, професор, завідувач кафедри комп'ютерних наук та інженерії, Східноукраїнський національний університет імені Володимира Даля, Київ, Україна;

Oleksandr Ryazantsev – doctor of technical sciences, professor, head of Computer Science and Engineering Department, Volodymyr Dahl East Ukrainian National University, Kyiv, Ukraine; e-mail: ryazantsev@ukr.net; ORCID ID: <http://orcid.org/0000-0002-3740-3132>.

Черпильський Максим Вікторович – магістрант кафедри комп'ютерних наук та інженерії, Східноукраїнський національний університет імені Володимира Даля, Київ, Україна;

Maksym Cherpitskiy – master's student of Computer Science and Engineering Department, Volodymyr Dahl East Ukrainian National University, Kyiv, Ukraine; e-mail: m_cherpitskiy@ukr.net; ORCID ID: <http://orcid.org/0009-0001-8778-7576>.

**Clustering and anomalies
of USA stock market volatility index data**

Ganna Khoroshun, Oleksandr Ryazantsev, Maksym Cherpitskiy

Abstract. Actuality. Investing money is an important way to improve the financial condition of both an individual and society as a whole. The problem of understanding financial data and making decisions regarding the investment of money in a certain project at this moment in time is relevant. **The object** of the study is the process of establishing the dependence of the value of the asset on time. **The subject** of research is mathematical models of data processing, data clustering and anomaly detection. **The purpose** of this work is to develop a method for effective investment of money using data processing and analysis methods for CVOE volatility index values in the USA, determination of clusters based on actions with assets, as well as checking the presence of anomalous data. **Research results.** Official data of volatility index values were selected and prepared for further analysis by removing incomplete sets and further normalization. Clustering of time series was carried out and the array was divided into five homogeneous groups. Clusters determine the ranges of the volatility index, which reflect the different sentiments of investors in the market and encourage appropriate actions with assets: to sell, to wait of index increasing, to buy, to remove money from developing projects and invest in stable ones, to wait of index decreasing. Segmentation of data, application of window function, centroids for segments were determined and signal reconstruction was carried out. Data anomaly points were identified. A comparative analysis was carried out based on the results of constructed initial data, reconstructed data and reconstruction error.

Keywords: clustering; anomaly; volatility index.

Methods of information systems synthesis

UDC 629.3.017.5

doi: <https://doi.org/10.20998/2522-9052.2023.2.03>Yevgen Aleksandrov¹, Tetiana Aleksandrova², Iryna Kostianyk², Yaroslav Morgun²¹ Kharkiv National Automobile and Highway University, Kharkiv, Ukraine² National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

PARAMETRIC SYNTHESIS OF THE INVARIANT SYSTEM OF CAR COURSE STABILITY

Abstract. The problem of constructing the invariant stabilizer of the ESP (Electronic Stability Program) car course stability system is considered by implementing two control principles by the electronic brake force distribution unit EBD (Electronic Brakeforce Distribution) – the principle of control by deviation and the principle of control by external disturbance. The values of the variable parameters of the stabilization algorithm are selected from the conditions for minimizing both the static and dynamic errors of the system. Two brake fluid pressure sensors are introduced into the EBD structural-functional scheme in the brake lines of the right and left sides of the car. It is proven that the pressure difference of the brake fluid, which is measured by pressure sensors, is proportional to the external disturbance acting on the car body from the side of the road surface. Therefore, in order to give the ESP system the property of invariance to external disturbances, the control signal generated by the EBD electronic unit contains current information not only about the parameters of the disturbed movement of the car, namely, about the angle of deviation of the longitudinal axis of the car relative to the given direction of movement, about the angular velocity of rotation of the body relative to its vertical axis and about lateral displacement of the center of mass of the body, but also the current information about the external disturbance acting on the car body. Recommendations for choosing the values of variable parameters of the ESP system stabilizer are given, which ensure the minimization of both static and dynamic errors of the closed system in the emergency braking mode.

Keywords: car course stability system; invariant stabilizer; static error of system; dynamic error of system; variable parameters of stabilizer.

Introduction

Problem statement. The main task of the ESP (Electronic Stability Program) car course stability system is to help the driver in critical situations, namely, to maintain the course stability and the given trajectory of the car under the influence of random external disturbances [1-3].

For the first time, the idea of creating ESP arose in 1959 and was patented by the Daimler-Benz company. In its modified form, it first appeared in 1995 on the Mercedes 600 serial car. Currently, most modern cars of various manufacturers are equipped with the ESP system. Since November 2011, ESP, together with the anti-lock braking system ABS, has become one of the active safety systems that all new models of cars and trucks registered in the European Union must be equipped with. Since November 2014, this law applies to all new cars without exception.

The conventional ESP system contains sensors of the angular speed of rotation of the wheels of the car, which are integrated with the anti-lock braking system ABS and the ASR traction control, a sensor of the position of the steering wheel, sensors of pressure of brake fluid in the brake lines of the right and left sides of the car, a sensor of the angular speed of rotation of the body relative to the main central vertical axis inertia of the body, a sensor of linear accelerations of the body relative to the main central transverse axis and a sensor of the current speed of the car. Information from the listed sensors enters the digital electronic brake force distribution unit EBD (Electronic Brakeforce Distribution), which, in the event of a critical situation,

helps to restore the position of the car body by generating an electrical control signal that is fed to the input of the executive body of the ESP system.

The ESP executive body changes the brake fluid pressure in the brake lines of the right and left sides of the car [4, 5].

Most often, a critical situation related to the loss of stability and controllability of the car occurs in emergency acceleration and braking modes. If the driver sharply presses the fuel pedal during urgent acceleration, the driving wheels of the car may slip, which leads to a loss of stability during acceleration. In this case, the ASR system reacts to the processes of skidding by reducing the active moments transmitted from the engine to the driving wheels of the car.

If the driver presses the brake pedal sharply during emergency braking, some of the car's wheels may lock, which may also lead to a loss of stability. In this case, the ABS system comes into play, reducing the braking force on the locked wheels.

In cases where the anti-lock braking system (ABS) or the anti-skid system (ASR) is unable to ensure stable movement of the car, the ESP system comes in action. Most often, this situation occurs in the mode of emergency braking. It should be noted that the capabilities of ESP to correct skidding and stabilize the car body in a critical situation are not unlimited. And if the speed at which a skid occurs is excessively high, or the grip between the wheels and the road surface is excessively low, then even the ESP system can be helpless. The ESP system significantly reduces the risks of skidding and an emergency situation on the road, but does not eliminate them [6].

In the article [7], the authors solve the problem of parametric synthesis of the ESP system, which is optimal in terms of dynamic accuracy, of a car-refueling vehicle.

Experimental studies of the developed system proved the presence of a static error in the ESP system.

Therefore, the goal of this work is to reduce the static error while maintaining high dynamic accuracy of the system, that is, to develop an ESP system that is invariant to the action of external disturbances acting on the object of stabilization.

Main material

Mathematical model of disturbed car movement during emergency braking mode. In work [7], a mathematical model of the disturbed movement of a car during emergency braking mode was developed, which has the following form:

$$\dot{v}_x(t) = -\frac{2k_b}{M} p_0 - gf(t); \quad (1)$$

$$\ddot{\psi}(t) = -\frac{Bk_b}{I} \Delta p(t) - \frac{2H_m M}{I} v_x(t) \dot{\psi}(t) f(t); \quad (2)$$

$$\Delta \ddot{p}(t) = -\frac{f_r}{I_r} \Delta \dot{p}(t) - \frac{c_r}{I_r} \Delta p(t) + \bar{k}_u u(t); \quad (3)$$

$$\dot{y}(t) = -v_x(t) \psi(t), \quad (4)$$

where the following values are taken (Fig. 1):

$v_x(t)$ – the current speed of the car;

$\psi(t)$ – angle of deviation of the main longitudinal central axis of inertia of the car ox relative to the given direction of movement OX ;

$\Delta p(t)$ – brake fluid pressure difference in the brake lines of the right and left sides of the car

$$\Delta p(t) = p_r(t) - p_l(t);$$

$y(t)$ – lateral shift of the center of mass of the car body relative to the given direction of movement;

$u(t)$ – control signal generated by the EBD electronic unit;

M – mass of the car;

g – acceleration of gravity;

I – the moment of inertia of the car body relative to its own central vertical axis of inertia;

B – track width;

H_m – the distance from surface of movement of the car to its center of mass;

$f(t)$ – the car's movement resistance coefficient, which depends on the properties of the road surface and the tread of the car's wheels;

p_0 – brake fluid pressure in the main brake cylinder;

k_b – the proportionality factor between the braking force on one of the sides of the car and the pressure of the brake fluid in the line of the corresponding side;

I_r – the moment of inertia of the rocker arm of the EM electromagnet of the executive body of the ESP system (Fig. 2);

f_r – coefficient of liquid friction in the axis of rotation of the rocker arm;

c_r – stiffness coefficient of the fixing spring of the rocker arm;

\bar{k}_u is the reduced value of the gain coefficient of the executive body.

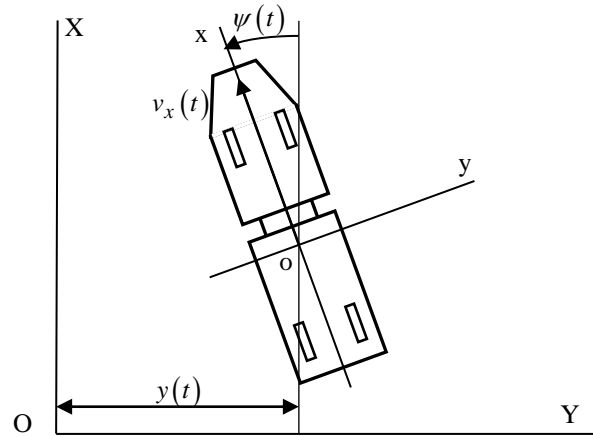


Fig. 1. For determination the current parameters of disturbed car movement

In works [7, 9], the control signal $u(t)$ was formed in the form

$$u(t) = k_\psi \psi(t) + k_{\dot{\psi}} \dot{\psi}(t) + k_y y(t), \quad (5)$$

where k_ψ , $k_{\dot{\psi}}$, k_y are varied EBD parameters, for the determination of which a combined algorithmic method of parametric synthesis is proposed in [10], which is a sequential combination of the Sobol grid method [10] and the Nelder-Mead method [11], implemented in the “Optimization Toolbox” extensions software package MATLAB and “Minimize” software package MATHCAD. Such a combination of two optimization methods allows, at the first stage of parametric synthesis, to bring the working point of the ESP system into the zone of the global minimum, and at the second stage to bring the working point to the point of the global minimum of the integral additive quadratic functional

$$I(k_\psi, k_{\dot{\psi}}, k_y) = \int_0^\tau [\beta_1^2 \psi^2(t) + \beta_2^2 \dot{\psi}^2(t) + \beta_3^2 y^2(t)] dt, \quad (6)$$

where β_1 , β_2 , β_3 are weighting coefficients of the functional (6), the values of which are calculated using a special algorithm implemented in the above-mentioned combined method of parametric synthesis.

In Fig. 2 shows the structural and functional scheme of the ESP executive body, where the following designations are adopted:

- EM – electromagnet;
- W1, W2 – electromagnet windings;
- R – rocker arm;
- Sp - fixing spring;
- N1, N2 – locking needles;
- MBC – main brake cylinder;
- WBC – working brake cylinder;
- HP – hydraulic pump;
- BP – brake pedal;

PSL, PSR – brake fluid pressure sensors in the brake lines of the right and left sides, respectively.

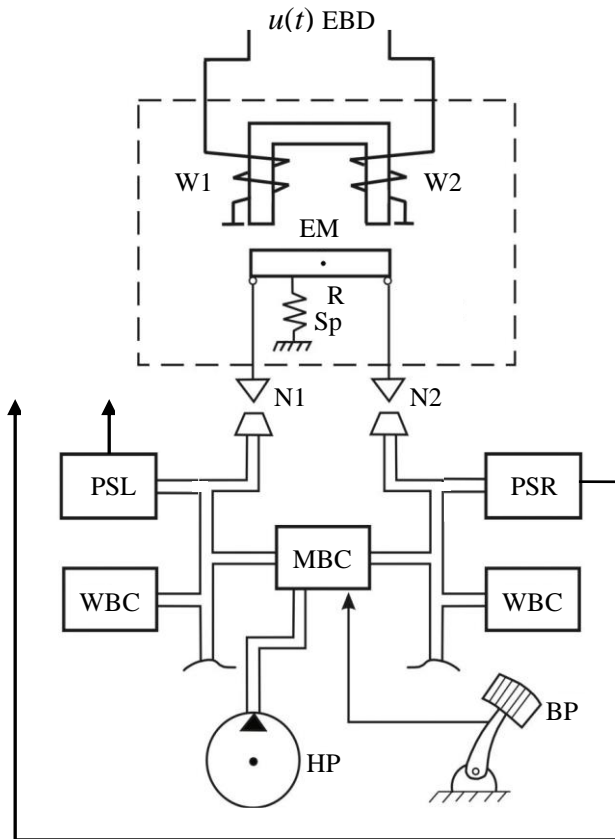


Fig. 2. Structural and functional scheme of the ESP executive body

From the EBD output, the control signal $u(t)$ is applied to one of the windings W1 or W2 of the EM.

If $u(t) > 0$, then the control signal module $|u(t)|$ is supplied to the winding W2, which leads to the rotation of the rocker arm R in the positive direction (counter-clockwise).

If $u(t) < 0$, then the control signal module $|u(t)|$ is supplied to the winding W1, which leads to the rotation of the rocker arm R in the negative direction (clockwise).

The rotation of R leads to the movement of the needles N1 and N2 in the corresponding direction.

When turning R counter-clockwise, the pressure of the brake fluid $p_l(t)$ in the left line increases, that is, the left side is braked at the same time, which causes the car to turn to the left.

When turning R clockwise, the right side of the car is braked, which causes it to turn to the right.

The structural and functional scheme of the ESP executive body shown in Fig. 2, is slightly different from the scheme given in [7]. In order to reduce the static error of the ESP system, that is, to give the system the property of invariance to external disturbances, it is necessary to implement two control principles in the ESP system, namely, the principle of deviation control and the control principle of external disturbance [14]. For this purpose, in the scheme of the ESP executive body, shown in Fig. 2, PSL and PSR sensors are used, the output signals of which are fed to the EBD inputs, and the control algorithm formed by the EBD electronic unit has the form

$$u(t) = k_\psi \psi(t) + k_{\dot{\psi}} \dot{\psi}(t) + k_y y(t) + k_p \Delta p(t), \quad (7)$$

moreover, the variable parameter should be chosen under the condition of minimizing the static error of the ESP.

Differential equations (1)–(4) together with relation (7) make up the mathematical model of the disturbed motion of the ESP closed system. This model is nonlinear because the right parts of the differential equations (2) and (4) have nonlinear terms.

Parametric synthesis of the invariant system of car course stability. In the emergency braking mode, the pressure of the brake fluid in the working cylinders reaches its maximum value p_0 , and the car begins to move with constant maximum acceleration a . The current speed of the car is

$$v_x(t) = v_0 - a \cdot t, \quad (8)$$

where v_0 is the initial speed at the beginning of the emergency braking mode. As a result, the disturbed motion of the car in the angular stabilization channel is described by differential equations

$$\ddot{\psi}(t) = -\frac{Bk_b}{I} \Delta p(t) - \frac{2H_m M}{I} (v_0 - a \cdot t) \dot{\psi}(t) f(t); \quad (9)$$

$$\Delta \ddot{p}(t) = -\frac{f_r}{I_r} \Delta \dot{p}(t) - \frac{c_r}{I_r} \Delta p(t) + \bar{k}_u u(t); \quad (10)$$

$$u(t) = k_\psi \psi(t) + k_{\dot{\psi}} \dot{\psi}(t) + k_p \Delta p(t). \quad (11)$$

Let's enter the notation

$$a_{\psi p} = \frac{Bk_b}{I};$$

$$a'_{\psi \dot{\psi}}(t) = \frac{2H_m M}{I} v_x(t) f(t);$$

$$a'_{pp} = \frac{f_r}{I_r};$$

$$a_{pp} = \frac{c_r}{I_r}.$$

Then the transfer function of the stabilization object is written as

$$W_0(s, t) = -\frac{a_{\psi p}}{s[s + a'_{\psi\psi}(t)]}. \quad (12)$$

Substitute relation (11) into the right part of the differential equation (10).

As a result, the differential equation (10) will be written as

$$\begin{aligned} \Delta\ddot{p}(t) + a'_{pp}\Delta\dot{p}(t) + (a_{pp} - \bar{k}_u k_p)\Delta p(t) = \\ = \bar{k}_u [k_{\psi}\psi(t) + k_{\dot{\psi}}\dot{\psi}(t)], \end{aligned} \quad (13)$$

and the transfer function of the stabilizer of the ESP system, which contains the EBD electronic unit and the executive body, is equal to

$$W_c(s) = \frac{\bar{k}_u(k_{\psi} + s \cdot k_{\dot{\psi}})}{s^2 + a'_{pp}s + (a_{pp} - \bar{k}_u k_p)}. \quad (14)$$

Let's write down the transfer function of the open circuit ESP in the angle stabilization channel

$$\begin{aligned} W_{oESP}(s, t) = W_0(s, t)W_c(s) = \\ = -\frac{\bar{k}_u a_{\psi p}(k_{\psi} + s \cdot k_{\dot{\psi}})}{s[s + a'_{\psi\psi}(t)][s^2 + a'_{pp}s + (a_{pp} - \bar{k}_u k_p)]}. \end{aligned} \quad (15)$$

The analysis of relation (15) allows us to conclude that the analyzed ESP in the angular stabilization channel is characterized by astatism of the first order, that is, the first coefficient of static error of the system is zero. If you put

$$a_{pp} - \bar{k}_u k_p = 0, \quad (16)$$

then the transfer function of the open ESP can be presented in the form

$$W_{oESP}(s, t) = -\frac{\bar{k}_u a_{\psi p}(k_{\psi} + s \cdot k_{\dot{\psi}})}{s^2[s + a'_{\psi\psi}(t)](s + a'_{pp})}. \quad (17)$$

When condition (16) is fulfilled, the ESP system is characterized by astatism of the second order, that is, the first two coefficients of the static error of such a system are equal to zero.

The degree of invariance of such a system to external disturbances is much higher than that of systems with a transfer function in an open state (15). Therefore, the value of the variable parameter k_p of the stabilization algorithm (7) should be chosen from the condition (16)

$$k_p = \frac{a_{pp}}{\bar{k}_u}. \quad (18)$$

It is known [14, 15] that a decrease in the static error leads to a decrease in the degree of stability of the closed system, that is, to an increase in the dynamic error.

Therefore, the choice of variable parameters of algorithm (7), namely k_{ψ} , $k_{\dot{\psi}}$ and k_y , should be aimed at minimizing the dynamic error of the closed ESP. It is advisable to select the values of these parameters using the combined algorithmic method of parametric synthesis, which is described in the article [7].

Conclusions

To reduce the static error of the car course stability system, it is advisable to combine two control principles, namely, the principle of control by deviation and the principle of control by external disturbance;

The ESP system control algorithm, which is formed by the EBD electronic unit, contains information about the current values of the parameters of disturbed car movement $\psi(t)$, $\dot{\psi}(t)$ and $y(t)$, as well as indirect information about the current value of the external disturbance acting on the car body, which is obtained by measuring the brake fluid pressure in the brake lines of the right and left sides of the car $p_r(t)$ and $p_l(t)$ and the formation of the difference in these pressures $\Delta p(t)$;

The value of the variable parameter k_p is chosen from the condition of increasing the degree of astatism of the ESP system, which leads to a decrease in the static error of the system, and the value of the variable parameters k_{ψ} , $k_{\dot{\psi}}$ and k_y – from the condition of the minimum of its dynamic error, which is estimated by functional (6).

REFERENCES

1. Tavernini, D., Velenis, E. & Longo, S. (2015), "Model-based active brake force distribution for pitch angle minimization", *2015 54th IEEE Conf. on Decision and Control*, Osaka, Japan, pp. 197-202, doi: <https://doi.org/10.1109/CDC.2015.7402108>.
2. Nakamura, E., Soga, M., Sacai, A., Otomo, A., & Kobayashi, T. (2002), Development of Electronically Controlled Brake System for Hybrid Vehicle. *SAE 2002 World Congress*. Detroit, Michigan, March 4 – 7, 2002, Printed in USA. 6 p.
3. Park, G. & Choi, S. B. (2016), "Optimal Brake Distribution for Electronic Stability Control Using Weighted Least Square Allocation Method.", *16 International Conference on Control «Automation and Systems» (ICCAS 2016)*. Oct. 16 – 19, 2016 in HICO. Gyeongju, Korea, pp. 1420–1425, doi: <https://doi.org/10.1109/ICCAS.2016.7832492>.
4. Wu, Xinyu. (2015), "Control Strategy and Algorithm Study on Light Vehicle Electronic Mechanical Braking System", *5th International Conference on Education, Management, Information and Medicine (EMJM 2015)*, pp. 1453–1458, doi: <https://doi.org/10.2991/emim-15.2015.282>.
5. Her, H., Cho, W. & Yi, K. (2011), "Vehicle Stability Control Using Individual Break Force Based on Tire Force Information", *14th International IEEE Conference on Intelligent Transportation Systems*. October 5 – 7, 2011, Washington, USA, pp. 22–29, doi: <https://doi.org/10.1109/ITSC.2011.6082997>.
6. Aleksandrov Ye. Ye., Aleksandrov T. Ye., Ovcharenko Yu. Ye. (2019), *Enhancement of technical and ergonomic characteristics of movable objects of military purpose*, monograph, KHNADU Publ., Kharkiv, 176 p. (in Ukrainian).

7. Aleksandrov, Ye.Ye, Aleksandrova, T.Ye and Morgun, Ya.Yu. (2022), "Parametrychnyi syntez tsyfrovoho stabilizatora systemy kursovoi stiihosti avtomobilia-palyvozapravnyka [Parametric synthesis of the digital stabilizer of the course stability system of the fuel-filling vehicle]", *Problemy keruvannia ta informatyky*, No. 2, pp. 69–91 (in Ukrainian), doi: <http://doi.org/10.34229/2786-6505-2022-2-5>.
8. Aleksandrova, T. Ye. (2013), "K voprosu sinteza lineinykh invariantnykh sistem stabilizatsii [On the question of the synthesis of linear invariant stabilization systems]", *Visnyk Natsional'noho tekhnichnoho universytetu «KhPI». Seriya : Systemnyi analiz, upravlinnia ta informatsiini tekhnologii*, NTU «KhPI», Kharkiv, No. 62 (1035), pp. 152–158 (in Russian).
9. Aleksandrov, Ye., Aleksandrova, T. & Morgun, Ya. (2019), "Parametric Synthesis of the Electronic Control Unit of the Course Stability System of the Car", *Eastern-European Journal of Enterprise Technologies*, No. 6/9(102), pp. 39–45, doi: <https://doi.org/10.15587/1729-4061.2019.188185>.
10. Sobol I. M. (1973), *Numerical Monte Carlo Methods*] Nauka, Moscow, 311 p.
11. Khimmelblau D. (1980). *Prikladnoe nelineinoe programmirovaniie* [Applied Nonlinear Programming], Mir, Moscow, 280 p. (in Russian).
12. Voronin A.N. (1992), *Mnokokriterialnyi sintez dinamicheskikh sistem* [Multicriteria Synthesis of Dynamic Systems], Naukova dumka, Kyiv, 160 p. (in Russian).
13. Voronin A. N. (2001), *Kompromiss i konsensus v teorii vektornoii optimizatsii* [Compromise and consensus in vector optimization theory], *Problemy upravleniia i informatiki*, No. 5, pp. 14–16 (in Russian).
14. Besekerskiy, V.A. & Popov Ye.P. (2004), *Teoriia sistem avtomaticheskogo upravleniia* [Theory of automatic control systems], Professia, Saint-Petersburg, 752 p. (in Russian).
15. Aleksandrov, Ye.Ye., Kozlov, Ye.P. and Kuznyeczov, B.I. (2002), *Avtomatychne keruvannia ruhomymy obyektamy i tehnologichnymy procesamy. Tom 1. Teoriia avtomatichnoho keruvannia* [Automatic control of moving objects and technological processes. Volume 1. Theory of automatic control], NTU «KhPI», Kharkiv, 476 p. (in Ukrainian).

Надійшла (received) 20.03.2023

Прийнята до друку (accepted for publication) 18.05.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Александров Євген Євгенович – доктор технічних наук, професор кафедри автомобілів, Харківський національний автомобільно-дорожній університет, Харків, Україна;

Yevgen Aleksandrov – Doctor of Technical Sciences, Professor Department of automobiles, Kharkiv National Automobile and Highway University, Kharkiv, Ukraine;

e-mail: aleksandrov.ye.ye@gmail.com; ORCID ID: <https://orcid.org/0000-0001-7525-6383>.

Александрова Тетяна Євгенівна – доктор технічних наук, професор кафедри «Системний аналіз та інформаційно-аналітичні технології», Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Tetiana Aleksandrova – Doctor of Technical Sciences, Professor Department of Systems Analysis and Information-Analytical Technologies, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;

e-mail: aleksandrova.t.ye@gmail.com; ORCID ID: <https://orcid.org/0000-0001-9596-0669>.

Костяник Ірина Віталіївна – кандидат технічних наук, доцент кафедри підійомно-транспортних машин і обладнання, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Iryna Kostianyk – PhD, Associate Professor Department of Lifting and Transporting Mashines and Equipment, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;

e-mail: kostyanik-irina@ukr.net; ORCID ID: <http://orcid.org/0000-0003-0289-2869>.

Моргун Ярослав Юрійович – аспірант кафедри «Системний аналіз та інформаційно-аналітичні технології», Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Yaroslav Morgun – graduate student Department of Systems Analysis and Information-Analytical Technologies, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine;

e-mail: yarki95@gmail.com; ORCID ID: <https://orcid.org/0000-0002-7399-4937>.

Параметричний синтез інваріантної системи курсової стійкості автомобіля

Є. Є. Александров, Т. Є. Александрова, І. В. Костяник, Я. Ю. Моргун

Анотація. Розглядається задача побудови інваріантного стабілізатора системи курсової стійкості автомобіля ESP (Electronic Stability Program) шляхом реалізації електронним блоком розподілення гальмівних зусиль EBD (Electronic Brakeforce Distribution) двох принципів керування – принципу керування за відхиленням і принципу керування за зовнішнім збуренням. Значення варійованих параметрів алгоритму стабілізації обираються з умов мінімізації як статичної, так і динамічної похибок системи. В структурно-функціональну схему EBD вводяться два датчики тиску гальмівної рідини в гальмівних магістралях правого і лівого бортів автомобіля. Доводиться, що різниця тисків гальмівної рідини, яка вимірюється датчиками тиску, пропорційна зовнішньому збуренню, що діє на корпус автомобіля з боку дорожнього покриття. Тому, для надання системі ESP властивості інваріантності до дії зовнішніх збурень сигнал керування, що формується електронним блоком EBD, містить поточну інформацію не тільки про параметри збуреного руху автомобіля, а саме, про кут відхилення поздовжньої осі автомобіля відносно заданого напрямку руху, кутову швидкість повороту корпусу відносно його вертикальної осі та бічний зсув центру мас корпусу, але й поточну інформацію про зовнішнє збурення, що діє на корпус автомобіля. Додаються рекомендації про вибір значень варійованих параметрів стабілізатора системи ESP, які забезпечують мінімізацію як статичної, так і динамічної похибок замкнутої системи в режимі термінового гальмування.

Ключові слова: система курсової стійкості автомобіля; інваріантний стабілізатор; статична похибка системи; динамічна похибка системи; варійовані параметри стабілізатора.

А. Ю. Воловик

Вінницький національний технічний університет, Вінниця, Україна

СИНТЕЗ КВАЗІОПТИМАЛЬНИХ ШВИДКОДІЮЧИХ ФІЛЬТРІВ З ПОЗИЦІЙ КРИТЕРІЮ НАЙМЕНШИХ КВАДРАТІВ

Анотація. Предметом досліджень статті є спеціальні методи обробки сигналів засновані на теорії оптимальної дискретної фільтрації. **Мета** – підвищення ефективності модельно-орієнтованих методів обробки інформаційних сигналів шляхом скорочення обчислювальних витрат і підвищення швидкодії алгоритмів оптимальної дискретної фільтрації. **Застосовувані методи:** опис динамічних процесів у термінах простору станів з використанням елементів векторно-матричної алгебри, метод зважених найменших квадратів, елементи теорії оптимальної дискретної фільтрації за Калманом, базові поняття теорії функціональних відновників О'Рейлі–Луенбергера, елементи теорії ймовірностей, статистичне моделювання за методом Монте-Карло. **Отримані результати:** запропоновано новий метод скорочення обчислювальних витрат, що використовує апроксимацію часової залежності матриці передачі фільтра Калмана заданими кусково-лінійними функціями відповідно до критерію найменших квадратів. Оцінена ефективність запропонованих варіантів апроксимацій. На основі порівняльного аналізу визначено декілька, прийнятних для практичної реалізації алгоритмів. **Практична значущість роботи** полягає у подальшому розвитку методів синтезу квазіоптимальних фільтрів підвищеної швидкодії. Працездатність запропонованих квазіоптимальних швидкодіючих фільтрів підтверджено на прикладі лінійної динамічної системи другого порядку. Показано, що сумарна економія в числі операцій множення й додавання може досягати десятків раз за рахунок незначних втрат у точності процесу фільтрації.

Ключові слова: фільтр Калмана; матричний коефіцієнт передачі; апроксимація; критерій найменших квадратів.

Вступ. Аналіз джерел

Вже більш ніж пів століття методи оптимальної лінійної фільтрації, розроблені Калманом і Бьюси в рамках концепції змінних стану, знаходять саме широке застосування в різних галузях науки й техніки, пов'язаних із проблемою оцінювання станів динамічних систем у присутності випадкових збурень. Основний результат даних робіт є загально відомим і представлений у вигляді системи різнице-вих рівнянь для фільтра і його коефіцієнту підсилення [1, 2].

$$\mathbf{x}(k/k) = \mathbf{A}(k, k-1)\mathbf{x}(k/k-1); \quad (1)$$

$$\begin{aligned} \mathbf{x}(k/k) &= \mathbf{x}(k/k-1) + \\ &+ \mathbf{K}(k) [\mathbf{y}(k) - \mathbf{C}(k)\mathbf{x}(k/k-1)]; \end{aligned} \quad (2)$$

$$\begin{aligned} \mathbf{K}(k) &= \mathbf{P}(k/k-1)\mathbf{C}(k) \times \\ &\times [\mathbf{C}(k)\mathbf{P}(k/k-1)\mathbf{C}(k) + \mathbf{R}(k)]; \end{aligned} \quad (3)$$

$$\begin{aligned} \mathbf{P}(k/k-1) &= \\ &= \mathbf{A}(k, k-1)\mathbf{P}(k-1/k-1)\mathbf{A}(k, k-1) + \\ &+ \mathbf{B}(k, k-1)\mathbf{Q}(k-1/k-1)\mathbf{B}(k, k-1); \end{aligned} \quad (4)$$

$$\mathbf{P}(k/k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{C}(k)]\mathbf{P}(k/k-1). \quad (5)$$

З часом були досліджені зв'язки з іншими процедурами оцінювання, наприклад, з байєсівськими, максимуму правдоподібності, мінімуму середньоквадратичної помилки, методом Гауса й іншими [3–6]. Показово, що незалежно від типу процедури оцінювання структура рівнянь рекурентного фільтра залишалася інваріантною, за винятком чисельних значень матриці передачі, обумовлених доступними апріорними даними. До теперішнього часу прийнято вважа-

ти, що тема лінійної дискретної фільтрації докладно й всебічно вивчена. Поряд з цим, у ряді практичних застосувань, що стосуються високо динамічних систем реального часу, виникають питання, пов'язані обчислювальними витратами, необхідним обсягом оперативної пам'яті, числом операцій множення й додавання, у кінцевому рахунку визначаючих швидкодію алгоритму фільтрації. На початку цього шляху користувалися загально відомим фактом, що при обчисленнях матричного коефіцієнта передачі оптимального лінійного фільтра результати фактичних вимірів не використовуються і обчислення оцінок засновано винятково на апріорно заданих системних матрицях. Отже, частина алгоритму (2)–(4) може бути виконана до початку процедури фільтрації, а обчислена послідовність матриць передачі повинна зберігатися в оперативній пам'яті аж до її використання в певний момент часу. При цьому властивість оптимальності фільтра зберігається, а найбільш відчутний вираш в рахунок часу буде спостерігатися для систем з постійними параметрами.

Подальша економія обчислювальних витрат може бути отримана при представленні оптимального фільтра Калмана у вигляді стохастичного спостерігача Луенбергера [7]

$$\begin{aligned} \mathbf{x}(k/k) &= [\mathbf{A}(k, k-1) - \mathbf{K}(k)\mathbf{C}(k)] \times \\ &\times \mathbf{x}(k-1/k-1) + \mathbf{K}(k)\mathbf{y}(k). \end{aligned} \quad (6)$$

Тут компоненти матриці передачі $\mathbf{K}(k)$ визначаються шляхом завдання власних значень системної матриці $(\mathbf{A} - \mathbf{K}\mathbf{C})$ в межах кола одиничного радіуса, що гарантує її стійкість і водночас надає більшу вільність у виборі цих коефіцієнтів. Для збереження прийнятної якості оцінювання рекомендується вибирати вищезгадані власні значення поблизу полюсів досліджуваної системи. Отриману структуру

спостерігача повного порядку можливо ще трохи спростити використовуючи спостерігач заниженого порядку. У цілому такі фільтри є квазіоптимальними, а їх параметри залишаються незмінними в часі. Цей підхід отримав подальший розвиток у роботі [8], де запропоновано так звану каскадну схему спостерігача, що складається із сукупності послідовно з'єднаних спостерігачів заниженого порядку. Її застосування виявилось ефективним стосовно до нелінійних систем високого порядку, які були призначені для роботи в ковзних режимах. Відомі методи декомпозиції розширеного фільтра Калмана, засновані на застосуванні ортогональних $U-V$ перетвореннях [9–12], що приводять коваріаційні матриці помилок екстраполяції й фільтрації до діагонального виду. Це дозволяє апроксимувати розширений фільтр Калмана паралельною структурою у вигляді банку автономно працюючих фільтрів зниженої розмірності, що в остаточному підсумку сприяє скороченню обчислювальних витрат, збільшуючи тим самим швидкість обробки результатів спостережень. Оскільки вихід вищезгаданої паралельної структури не є еквівалентним до виходу розширеного фільтра Калмана, то результат синтезу також є квазіоптимальним. В роботі [13] було запропоновано локально-оптимальний варіант фільтра, у якому оптимальність, у змісті мінімуму середньоквадратичної помилки, зберігається в межах дії уведених обмежень, що накладаються на структуру системних матриць. Нескладно помітити, що вищенаведені методи засновані на процедурах спрощення вихідної структури оптимального фільтра й відмови від оптимальності перехідного процесу, що побічно сприяє збільшенню швидкодії за рахунок прийнятної втрати точності оцінювання для систем з відносно повільною динамікою. На відміну від вищевикладеного, у даній роботі пропонується систематична процедура безпосередньої апроксимації матричного коефіцієнта передачі оптимального фільтра Калмана, що заснована на критерії найменших квадратів. Подібні питання частково розглядалися в роботах [14–16]. Так, наприклад, в [14] використовувалася апроксимація імпульсної характеристики фільтра в далекомірній системі кусково-постійними функціями із кроком $2-N$, де N – розрядність регістру пам'яті. Однак, вибір цих функцій визначався винятково зручностями реалізації фільтра на елементах середнього ступеня інтеграції, а точність апроксимації не розглядалася. Пізніше, у роботах [15, 16] вибір прийнятних значень матричного коефіцієнта передачі рекурсивного фільтра здійснювався на основі розв'язку системи лінійних рівнянь методом ітерацій, при цьому для знаходження коваріаційних матриць помилок оцінювання й екстраполяції застосовувався метод розширення вектора стану і його фундаментальної матриці. Практика застосування рекурсивних фільтрів показує, що у багатьох випадках складність ітераційних обчислень не представляється виправданою для відносно простих траєкторій зміни вектора стану контрольованої системи.

$$P(k/k) = [I - K(k)C(k)]P(k/k-1)[I - K(k)C(k)] + K(k)R(k)K(k). \quad (7)$$

Постановка задачі та результати

Сутність запропонованого методу буде продемонстрована на прикладі системи другого порядку, спрощена модель якої описує еволюцію кута місця повітряного судна, що заходить на посадку в режимі горизонтального польоту із застосуванням навігаційного встаткування посадкової системи сантиметрового діапазону і описується наступними рівнями

$$\begin{aligned} x(k+1) &= A(k+1, k)x(k) + B(k+1, k)w(k); \\ y(k) &= C(k)x(k) + v(k); \end{aligned} \quad (7)$$

де $x(k) = [x_1(k), x_2(k)]$ – вектор стану системи, що складається з значення кута місця повітряного судна та швидкості його зміни відповідно; $y(k)$ – результат вимірювань кута місця. Априорні дані й обґрунтування моделі детально описані в монографії [17]:

$$\begin{aligned} A(k+1, k) &= \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}; \quad B(k+1, k) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}; \\ C(k) &= [1 \quad 0] - \end{aligned}$$

системні матриці, відповідних розмірностей, T – інтервал дискретизації. $w(k)$, $v(k)$ – некорельовані гаусові білі послідовності з нульовими середніми значеннями і заданими коваріаційними матрицями $Q(k)$, $R(k)$, відповідно. Початкові дані для проведення моделювання наведені в табл. 1.

Таблиця 1 – Априорні дані для дільниці горизонтального польоту

Параметр	Позначення	Числове значення	Розмірність
Початкове значення оцінки кута місця	$x_1(0/0)$	0.7000	град.
Початкове значення оцінки швидкості	$x_2(0/0)$	0.1500	град./с
Дисперсія значення кута місця	$P_{11}(0/0)$	0.4900	(град.) ²
Дисперсія швидкості зміни кута місця	$P_{22}(0/0)$	0.1600	(град./с) ²
Інтервал дискретизації	T	0.0247	с
Коваріаційна матриця збурень	$Q(k)$	0.0001	(град./с) ²
Коваріаційна матриця помилок вимірів	$R(k)$	0.0081	(град.) ²

Результати оцінювання у відповідності з (1)–(4) відображені на рис. 1.

Перш ніж переходити від оптимальної структури пристрою обробки результатів кутових вимірів до його практичного аналога доцільно здійснити пошук варіантів побудови дискретних фільтрів з підвищеною швидкістю, зрівняти їхні якісні показники точності з оптимальними, і на цій основі сформулювати обґрунтовані вимоги до обчислювальної потужності процесора. В рахунок втрати оптимальності, коваріаційна матриця помилок фільтрації повинна розраховуватися на підставі формули, слушної для довільної матриці передачі [1]:

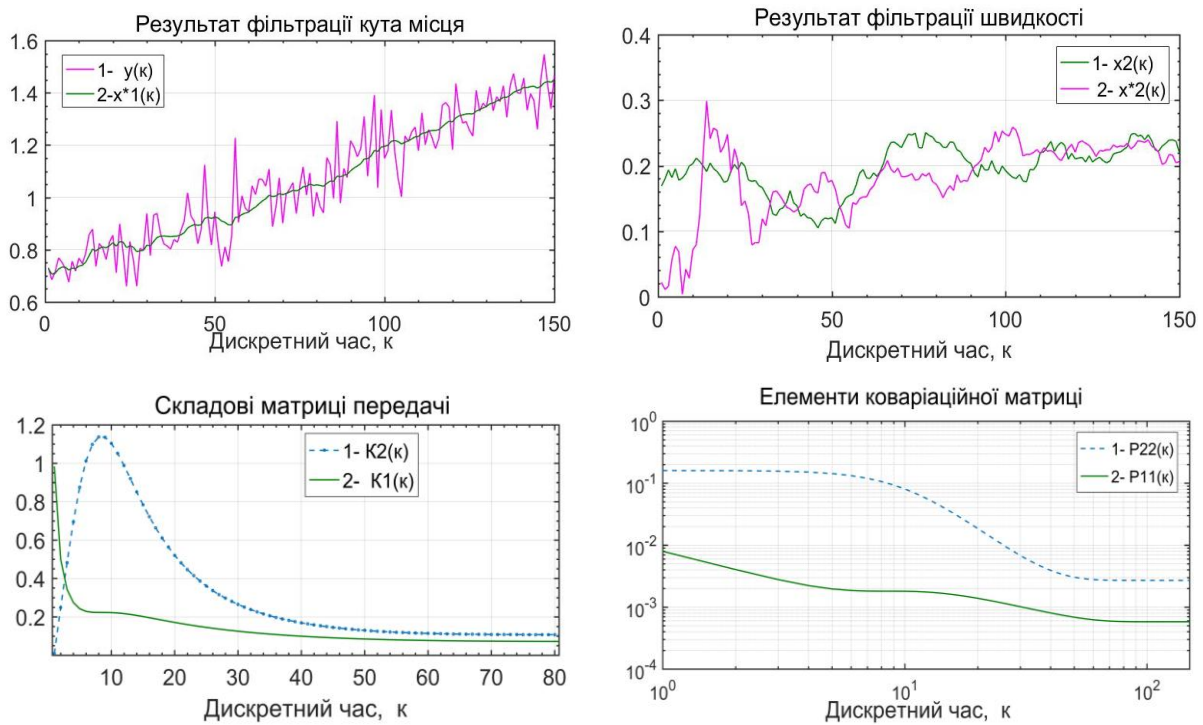


Рис. 1. Оцінки параметрів посадкової траєкторії фільтром Калмана
(Fig. 1. Estimates of landing trajectory parameters using the Kalman filter)

Ефективність варіантів апроксимацій матричного коефіцієнта передачі оцінювалася шляхом нормування показників точності до відповідних показників точності оптимального фільтра Калмана, отриманих за результатами імітаційного моделювання з використанням апріорних даних табл. 1. У процесі пошуку розглядалися наступні варіанти побудови фільтрів підвищеного швидкодії:

Варіант 1. Тут використовується фільтр із постійними в часі параметрами, який мінімізує середньоквадратичну помилку у сталому режимі. Skorиставшись методикою роботи [18], можливо показати, що при заданих умовах польоту й апріорних даних щодо системних шумів збурень $Q(k)=0,0001$ і шумів збурень вимірів $R(k)=0,0081$ елементи матриці передачі фільтра Вінера будуть дорівнювати $K_1(k) = 0,24$ і $K_2(k) = 0,02$, а тривалість перехідного процесу становить приблизно 0,5 с. Щодо дисперсії помилок оцінювання кута місця, то вони починаючи з 200-го такту приблизно однакові, а у сталому режимі збігаються (див. рис. 2а). Що стосується перехідного режиму для оцінок швидкості зміни кута місця повітряного судна, то тут фільтр Калмана має якісні показники більш ніж на порядок кращі. Повільну збіжність оцінок швидкості, частково, можна пояснити відсутністю відповідних вимірів, але головна причина все таки криється в тому, що фільтр Вінера не призначений для оптимізації помилок у перехідному режимі.

Крім того, передбачається, що результати фільтрації повинні надходити в систему керування польотом. Це означатиме, що на них поширюються обмеження пов'язані зі швидкодією й стійкістю контуру керування в цілому. Оскільки виконання посадкового маневру у вертикальній площині вимагає обмеження тривалості перехідного процесу фі-

льтра до величини 0,5 с. при ручному керуванні, і 0,1–0,25 с. при застосуванні автопілота [19], то високий рівень помилок у перехідному режимі виключає використання фільтра Вінера у швидкоплинному процесі посадки повітряного судна.

Варіант 2. За аналогічною методикою можливо розраховувати елементи матриці передачі квазіоптимального фільтра $K_1(k)$ і $K_2(k)$, виходячи із заданої тривалості перехідного процесу, наприклад, $t_n=0,2$ с. Вони становлять 0,267 і 0,183, відповідно. Посилаючись на рис. 2б, можна помітити, що істотних змін у перехідному режимі не відбулося. Вище наведені приклади показують, що вимоги заданої тривалості перехідного процесу й мінімізації середньоквадратичних помилок є взаємно суперечливими, а можливого компромісу, можна досягти лише за допомогою фільтрів зі змінними параметрами.

Далі пошук варіантів побудови квазіоптимальних фільтрів зі змінними параметрами здійснювався шляхом апроксимації часових залежностей $K_1(k)$ і $K_2(k)$ кусочно-постійними функціями за критерієм найменших квадратів. Результати пошуку відображені в табл. 2.

Варіанти 3–4. У варіанті 3 вдало поєднуються простота процедури керування коефіцієнтами передачі $K_1(k)$ і $K_2(k)$ з високою точністю оцінювання кута місця повітряного судна. Відповідно до табл. 2 для збереження апроксимованих часових залежностей $K_1(k)$ і $K_2(k)$ необхідно не більш п'яти комірок оперативної пам'яті: дві для $K_1(k)$ і три для $K_2(k)$. Щодо точності оцінювання кута місця повітряного судна, то процедура фільтрації за варіантом 3 уступає фільтру Калмана в перехідному режимі не більше ніж на 40%. У встановленому режимі ці показники практично збігаються.

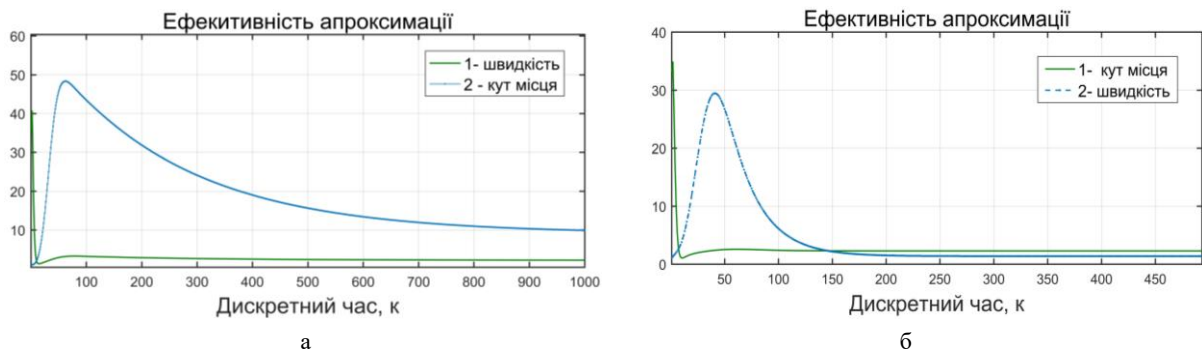


Рис. 2. Ефективність апроксимації матричного коефіцієнта передачі постійними величинами (Fig. 2. Effectiveness of the approximation of the matrix transmission coefficient by constant values): а – для фільтра Вінера; б – фільтра із заданою тривалістю перехідного процесу

Таблиця 2 – Види апроксимації коефіцієнтів $K_1(\kappa)$ і $K_2(\kappa)$ кусково-постійними функціями

	Вид апроксимації $K_1(\kappa)$	Значення $K_1(\kappa)$	Вид апроксимації $K_2(\kappa)$	Значення $K_2(\kappa)$
3		$K_1(1)=1.0;$ $K_1(2) - K_1(15)=0.32;$ $K_1(16) - K_1(300)=0.1.$		$K_2(1) - K_2(2)=0.0;$ $K_2(3) - K_2(14)=1.0;$ $K_2(15) - K_2(34)=0.34;$ $K_2(35) - K_2(300)=0.12.$
4		$K_1(1) - K_1(2)=0.95;$ $K_1(3) - K_1(5)=0.25;$ $K_1(6) - K_1(20)=0.2;$ $K_1(21) - K_1(40)=0.15;$ $K_1(41) - K_1(300)=0.07.$		$K_2(1) - K_2(2)=0.0;$ $K_2(3) - K_2(5)=0.80;$ $K_2(6) - K_2(15)=1.08;$ $K_2(16) - K_2(31)=0.40;$ $K_2(32) - K_2(50)=0.17;$ $K_2(51) - K_2(300)=0.11.$
5		$K_1(1) - K_1(2)=1.0;$ $K_1(3) - K_1(300)=0.1.$		$K_2(1) - K_2(3)=0.0;$ $K_2(4) - K_2(23)=1.0;$ $K_2(24) - K_2(300)=0.10.$

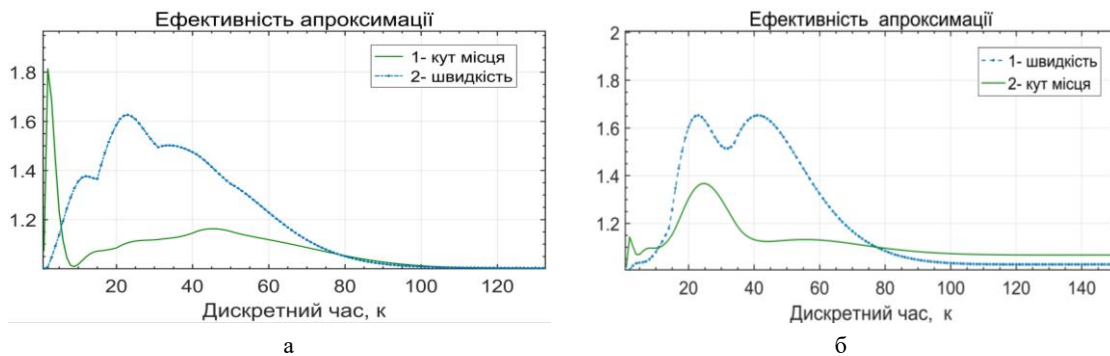


Рис. 3. Ефективність апроксимацій за варіантами 3–4 (Fig. 3. Efficiency of approximations according to options 3–4)

Більш детальна апроксимація залежностей $K_1(\kappa)$ і $K_2(\kappa)$ властива варіанту 4 (табл. 2). Деталізація ускладнює процес керування матричним коефіцієнтом передачі й на фоні відсутності істотного поліпшення точності оцінювання кута місця. Більше того, 10-15% підвищення точності фільтрації в порівнянні з варіантом 3 (рис. 3б) обмінюється на трикратне збільшення обсягу пам'яті. Тому з погляду простоти технічного розв'язку поставленого завдання перевага слід віддавати варіанту за номером 3.

Варіант 5 ілюструє помилковий розв'язок розглянутого завдання, а саме бажання звести перехідний процес у системі до декількох ітерацій за рахунок імпульсного розширення смуги пропускання

згладжуючого фільтра. На цих тактах коефіцієнти передачі $K_1(\kappa)$ і $K_2(\kappa)$ одночасно прирівнюються до одиниці (див. табл. 2), що відповідає абсолютній вірогідності результатів вимірювання. Внаслідок цього функція згладжування втрачається. Незважаючи на такі вжиті заходи, точність оцінювання кутів координат у порівнянні з фільтром Калмана стає гірше в 3-4 рази (рис. 4), а перехідний процес виходить надмірно затяжним. Причина цього полягає в тому, що раптове розширення смуги пропускання фільтра дійсно швидко зменшує динамічну складову помилки оцінювання кута місця повітряного судна, але за рахунок різкого зростання шумової складової.

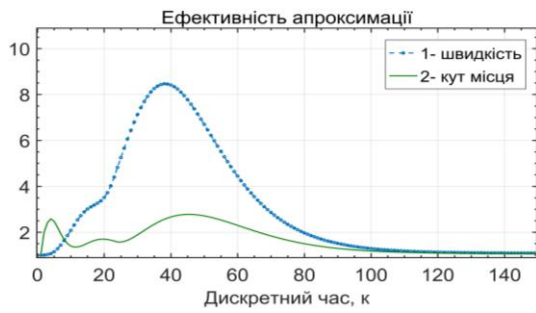


Рис. 4. Ефективність апроксимації за варіантом 5
(Fig. 4. Efficiency of approximation according to option 5)

Для варіантів 3-4, що представляють найбільший практичний інтерес, проведемо більш ретельну оцінку обчислювальних витрат, скориставшись при цьому результатами роботи [20, 21] з наступним порівнянням отриманих показників аналогічними показниками фільтра Калмана. Кількість комірок оперативної пам'яті, необхідних для реалізації рівнянь (1)-(5), визначалося розмірністю відповідних векторів і матриць. Ці відомості наведені в табл. 3. Результати підрахунку числа операцій додавання й множення, необхідних для обчислення оптимальних оцінок станів на окремому ітераційному кроці, наведено в табл. 4.

Таблиця 3 – Потреба у комірках пам'яті ОЗП для фільтра Калмана

Змінна	$x(k/k-1)$	$x(k/k)$	$A(k, k-1)$	$K(k)$	$y(k)$	$C(k)$	$R(k)$	$Q(k)$	$B(k, k-1)$	$P(k-1)$	$P(k)$
Розмірність	$n \times 1$	$n \times 1$	$n \times n$	$n \times m$	$m \times 1$	$m \times n$	$m \times m$	$q \times q$	$n \times q$	$n \times n$	$n \times n$
Число комірок пам'яті ОЗП	n	n	n^2	nm	m	mn	m^2	q^2	nq	n^2	n^2

Таблиця 4 – Потреба в числі операцій додавання й множення на окремому ітераційному кроці

Обчислювальне значення	Рівняння фільтрації	Етапи обчислення	Число множень	Сумарне число множень	Число додавань	Сумарне число додавань
$x(k-1/k-1)$	(1)	$A(k, k-1)x(k-1/k-1)$;	n^2	n^2	$(n-1)n$	$(n-1)n$
$x(k/k)$	(2)	$C(k)x(k-1/k-1)$ $[y(k) - C(k)x(k-1/k-1)]$ $K(k)[y(k) - C(k)x(k-1/k-1)]$ $x(k/k)$	mn - nm -	- - 2 -	$(n-1)m$ m $n(m-1)$ n	$2nm$
$K(k)$	(3)	$P(k-1)C^T(k)$ $C(k)[P(k-1)C^T(k)]$ $C(k)P(k-1)C^T(k) + R(k)$ $[C(k)P(k-1)C^T(k) + R(k)]^{-1}$ $K(k)$	n^2m nm^2 - nm^3 n^3	- - $m(m^2 + 2nm + n^2)$ -	$nm(n-1)$ $(n-1)m^2$ m^3 $nm(m-1)$	$nm(n+2m-2) + m^3$ $nm(n+2m-2) + m^3$
$P(k-1)$	(4)	$P(k-1)A^T(k, k-1)$ $A(k, k-1)[P(k-1)A^T(k, k-1)]$ $Q(k-1)B^T(k, k-1)$ $B(k, k-1)[Q(k-1)B^T(k, k-1)]$ $A(k, k-1)[P(k-1)A^T(k, k-1)] +$ $+B(k, k-1)[Q(k-1)B^T(k, k-1)]$	n^3 nq^2 n^2q -	- - $2n^3 + nq^2 + n^2q$ -	$n^2(n-1)$ $n^2(n-1)$ $nq(q-1)$ $n^2(q-1)$ n^2	$2n^2(n-1) + nq(q+n-1)$
$P(k/k)$	(5)	$K(k)C(k)$ $I - [K(k)C(k)]$ $[I - K(k)C(k)]P(k-1)$	- n^3	- $n^2(n+m)$	$m(n-1)$ n^2 $n(n-1)$	$n^2(n+m-1)$
Загальне число операцій множення:			$M(n, m, q) = 3n^3 + 2n^2m + 2nm^2 + n^2q + nq^2 + 2nm + m^3 + n^2$			
Загальне число операцій додавання:			$A(n, m, q) = 3n^3 + 2n^2m + 2nm^2 + n^2q + nq^2 + m^3 - 2n^2 - nq$			

У тому випадку, коли через лінійність фільтра елементи матричного коефіцієнта передачі не залежать від результатів спостережень і можуть бути обчисленими на підставі апріорних даних, у реаль-

ному масштабі часу слід обчислювати тільки поточні оцінки за формулами (1)-(2). При цьому загальне число операцій множення й додавання скорочується до величин:

$$M(n, m, q) = n(n + 2m);$$

$$(n, m, q) = n(n + 2m - 1).$$

Стосовно до розглянутого прикладу, у якому $n = 2$, $m = q = 1$, ці величини рівні

$$M_I(2, 1, 1) = 8 \quad \text{й} \quad A_I(2, 1, 1) = 6$$

відповідно.

Для порівняння, повноформатному варіанту фільтра Калмана для отримання оптимальних оцінок вектора стану буде потрібно число множень $M(2, 1, 1) = 51$, а число додавань – $A(2, 1, 1) = 33$. Неважко помітити, що ріст обчислювальних витрат нелінійно залежить від порядку розглянутої системи й у першому наближенні оцінюється третім степенем від розмірності вектора стану.

Крім того, підсумковий обсяг обчислювальних витрат можна ще суттєво зменшувати, якщо врахувати, що запропонований метод апроксимації відповідно до результатів п. 3-5 табл. 2 скорочує число

необхідних ітерацій на етапі перехідного процесу як мінімум на порядок.

Висновки

1. Розглянуті методи скорочення обсягу обчислювальних витрат, що приводять у підсумку до спрощення структури оптимального фільтра Калмана й підвищенню його швидкодії.

2. Запропонований метод скорочення підсумкових обчислювальних витрат, заснований на апроксимації часової залежності матриці передачі фільтра Калмана кусково-постійними функціями за критерієм найменших квадратів.

3. Ефективність методу продемонстрована на прикладі динамічної системи другого порядку. Шляхом порівняльного аналізу запропоновано декілька варіантів шуканої апроксимації. Показано, що досяжна економія в числі операцій множення й додавання може досягати десятка разів при незначній втраті точності процесу фільтрації/

REFERENCES

1. Bozic, S.M. (2018), *Digital & Kalman Filtering. An Introduction to Discrete-Time Filtering & Optimum Linear Estimation*, Dover publications, INC: Mineola, New York, 240 p.
2. Chui, C. (2018), *Kalman Filtering: with Real-Time Applications*, Springer; Softcover reprint of the original 5th ed. 265 p.
3. Calin, O. (2021), *Informal Introduction To Stochastic Calculus With Applications*, WSPC; 2nd edition, 510 p.
4. Davis, B. and Blair, W. (2018), "Maximum likelihood mixture modeling for three-dimensional non-Gaussian measurements", *IEEE Radar Conference*, Oklahoma City, OK, USA, pp. 1027-1032, doi: <https://doi.org/10.1109/RADAR.2018.8378702>.
5. Gakis, G. and Smith, M. A (2022), "Deterministic Least Squares Approach for Simultaneous Input and State Estimation", *IEEE Transactions on Automatic Control*, pp. 1 – 16, doi: <https://doi.org/10.1109/TAC.2022.3209415>.
6. Yang, J.N., Pan, S. and Huang, H. (2007), "An adaptive extended Kalman filter for structural damage identifications. II: Unknown inputs", *Structural Control and Health Monitoring* 14, pp. 497–521, doi: <https://doi.org/10.1002/stc.171>.
7. O'Reilly, J. (2007), *Observers for Linear Systems*, Academic Press, 258 p.
8. Krasnova, S.A. and Utkyn, V.A. (2006), *Kaskadniy sintez nabliudatelei sostoiannya dynamicheskoykh system*, [Cascade Synthesis of Dynamical Systems State Observers] Yn-t problem upravleniya ym. V.A. Trapeznikova. RAN, Nauka, 272 p.
9. Alouani, A.T., Rice, T.R and Blair, W.D. (2007), "A two-stage filter for state estimation in the presence of dynamical stochastic bias", *Proceedings of the American Control Conference*, Chicago, IL, pp.1784–88. doi: <https://doi.org/10.23919/ACC.1992.4792418>.
10. Volovyk, A.Yu., Kychak, V.M. and Havrilov, D.V. (2021), "Discrete Kalman Filter Invariant to Perturbations". *Acta Polytechnica Hungarica*, Vol. 18, No. 10, pp. 21-41, doi: <https://doi.org/10.12700/APH.18.10.2021.10.2>.
11. Volovyk, A., Kychak, V., Osadchuk, A., Zhurakovskiy, B. (2023). "Fault Identification in Linear Dynamic Systems by the Method of Locally Optimal Separate Estimation", *Emerging Networking in the Digital Transformation Age. TCSET 2022. Lecture Notes in Electrical Engineering*, vol 965. Springer, Cham., pp.634-651, doi: https://doi.org/10.1007/978-3-031-24963-1_37.
12. Marelli, D., Sui, T. and Fu, M. (2021), "Distributed Kalman estimation with decoupled local filters", *Automatica*, Vol.132, 109793, doi: <https://doi.org/10.1016/j.automatica.2021.109724>.
13. Hsieh, C.S. (2007), "Extension of the robust two-stage Kalman filtering for systems with unknown inputs", *TENCON 2007, IEEE Region 10 Conference*, Taipei, Taiwan, pp. 1-4, doi: <https://doi.org/10.1109/TENCON.2007.4429133>.
14. Syromiatnykov, R.F. (1975), "Yssledovanye pomekhoustoichyvosty alhorytmov mnogokanalnogo yzmeritelia dalnosti v sisteme avtomatizirovannogo upravleniya poletamy", [Investigation of the noise immunity of algorithms for a multi-channel range meter in an automated flight control system], Kand. Dissertation, LETY, 251p.
15. Avtin, I. V., Baburov, V. I., and Ponomarenko, B. V. (2021), *Principles of Integrated Airborne Avionics*, Springer Aerospace Technology, Springer; 1st ed., 719 p.
16. Henningsson, T. (2007), "Recursive state estimation for linear systems with mixed stochastic and set-bounded disturbances", *47th IEEE Conference on Decision and Control*, Cancun, Mexico, 2008, pp. 678-683, doi: <https://doi.org/10.1109/CDC.2008.4739139>.
17. Kychak, V.M., Volovyk, A.Iu and Volovyk, Yu. (2011), *Metody ta prystroi obrobky radiosyhnaliv bortovykh aviatsiynkhn system posadky*, [Methods and devices for radio signals processing of airborne aircraft landing systems], Monohrafiia, Vinnytsia, VNTU, 208 p.
18. Pervachev, S.V., Valuev, A.A. and Chylykyn, V.M. (1973), *Statysticheskaia dynamika radyotekhnicheskikh slediashchikh system*, [Statistical dynamics of radio tracking systems], M.: Sov. radyo, 488 p.
19. Stengel, R. (2022), *Flight Dynamics: Second Edition 2nd Edition*, Princeton University Press; 2nd edition, 912 p.
20. Hsieh, C.S. and Chen, F.C. (1999), "Optimal solution of the two-stage Kalman estimator", *IEEE Transactions on Automatic Control*, vol. 44, no. 1, pp. 194-199, doi: <https://doi.org/10.1109/9.739135>.
21. Lo, K., Lu, Q., and Kwon, W. (2022), "Comments on "Optimal solution of the two-stage Kalman estimator"", *IEEE Transactions on Automatic Control*, vol. 47, no. 1, pp. 198-199, doi: <https://doi.org/10.1109/9.981745>.

СПИСОК ЛІТЕРАТУРИ

1. Bozic S. M. Digital & Kalman Filtering. An Introduction to Discrete-Time Filtering & Optimum Linear Estimation. Dover publications, INC: Mineola, New York. 2018, 240 p.
2. Chui C. Kalman Filtering: with Real-Time Applications. Springer; Softcover reprint of the original 5th ed, 2018, 265 p.
3. Calin O. Informal Introduction to Stochastic Calculus with Applications. WSPC; 2nd edition, 2021. 510 p.
4. Davis B., Blair W. Maximum likelihood mixture modeling for three-dimensional non-Gaussian measurements, *IEEE Radar Conference*. Oklahoma City, OK, USA. 2018. P. 1027-1032. DOI: <https://doi.org/10.1109/RADAR.2018.8378702>.
5. Gakis G., Smith M., A Deterministic Least Squares Approach for Simultaneous Input and State Estimation, *IEEE Transactions on Automatic Control*. 2022, P. 1 - 16. DOI: <https://doi.org/10.1109/TAC.2022.3209415>.
6. Yang J.N., Pan S., Huang, H. An adaptive extended Kalman filter for structural damage identifications. II: Unknown inputs. *Structural Control and Health Monitoring* 14, 2007. P. 497–521. DOI: <https://doi.org/10.1002/stc.171>
7. O'Reilly J. Observers for Linear Systems. Academic Press, 2012. 258 p.
8. Краснова С.А., Уткін В.А. Каскадний синтез спостерігачів стану динамічних систем. М.: Наука, 2006. 272 с.
9. Alouani A.T., Rice T.R., Blair W.D. A two-stage filter for state estimation in the presence of dynamical stochastic bias. *Proc. of the American Control Conference*. 1992. Chicago, IL. P. 1784–1788. DOI: <https://doi.org/10.23919/ACC.1992.4792418>.
10. Volovyk A.Yu., Kychak V.M., Havrilov D.V. Discrete Kalman Filter Invariant to Perturbations. *Acta Polytechnica Hungarica*. 2021. Vol. 18, No. 10. P. 21-41. DOI: <https://doi.org/10.12700/APH.18.10.2021.10.2>.
11. Volovyk A., Kychak V., Osadchuk A., Zhurakovskiy B. Fault Identification in Linear Dynamic Systems by the Method of Locally Optimal Separate Estimation. Emerging Networking in the Digital Transformation Age. TCSET 2022. *Lecture Notes in Electrical Engineering*. 2023. Vol 965. Springer, Cham. P. 634-651. DOI: https://doi.org/10.1007/978-3-031-24963-1_37.
12. Marelli D., Sui T., Fu M. Distributed Kalman estimation with decoupled local filters. *Automatica*. Vol. 132, October 2021, 109793. DOI: <https://doi.org/10.1016/j.automatica.2021.109724>.
13. Hsieh C.S. Extension of the robust two-stage Kalman filtering for systems with unknown inputs. *TENCON 2007, IEEE Region 10 Conference*. Taipei, Taiwan, 2007. P. 1-4. DOI: <https://doi.org/10.1109/TENCON.2007.4429133>.
14. Сиром'ятників Р.Ф. Дослідження завадостійкості алгоритмів багатоканального вимірювача дальності у системі автоматизованого управління польотами. Л.: ЛІТЕТ, 1975.
15. Avtin I. V., Baburov V. I., Ponomarenko B. V. Principles of Integrated Airborne Avionics. Springer Aerospace Technology. Springer; 1st ed. 2021. 719 p.
16. Henningsson T. Recursive state estimation for linear systems with mixed stochastic and set-bounded disturbances. *47th IEEE Conference on Decision and Control*. Cancun, Mexico, 2008. P. 678-683. DOI: <https://doi.org/10.1109/CDC.2008.4739139>.
17. Кичак В.М., Воловик А.Ю., Воловик Ю.М. Методи та пристрої обробки радіосигналів бортових авіаційних систем посадки: Монографія. Вінниця, ВНТУ, 2011. 208 с.
18. Первачов С. В., Валуев А. А., Чилікін В. М. Статистична динаміка радіотехнічних систем, що стежать. М: Рад. радіо, 1973. 488 с.
19. Stengel R. Flight Dynamics: Second Edition 2nd Edition. Princeton University Press; 2nd edition, 2022, 912 p.
20. Hsieh C.S., Chen F.C. Optimal solution of the two-stage Kalman estimator, *IEEE Transactions on Automatic Control*. Vol. 44, no. 1. P. 194-199, Jan. 1999. DOI: <https://doi.org/10.1109/9.739135>.
21. Lo K., Lu Q., Kwon W. Comments on "Optimal solution of the two-stage Kalman estimator". *IEEE Transactions on Automatic Control*. Vol. 47, no. 1. P. 198-199, Jan. 2002/ DOI: <https://doi.org/10.1109/9.981745>.

Надійшла (received) 27.02.2023

Прийнята до друку (accepted for publication) 19.04.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Воловик Андрій Юрійович – кандидат технічних наук, доцент, доцент кафедри інформаційних радіоелектронних технологій і систем, Вінницький національний технічний університет, Вінниця, Україна;

Andrii Volovyk – PhD in Technical Sciences, associate professor, associate professor of Department of Information Radioelectronic Technologies and Systems, Vinnytsia National Technical University, Vinnytsia, Ukraine;

e-mail: voland@vntu.edu.ua; ORCID ID: <http://orcid.org/0000-0002-8772-9843>.

Synthesis of quasi-optimal fast filters by the least square criterion

Andrii Volovyk

Abstract. The subject of the article research is special signal processing methods based on the optimal discrete filtering theory. The goal is to increase the efficiency of model-based methods for processing information signals by reducing computational costs and increasing the speed of optimal discrete filtering algorithms. Applied methods: description of dynamic processes in terms of state space using elements of vector-matrix algebra, weighted least squares method, elements of Kalman's theory of optimal discrete filtering, basic concepts of the O'Reilly–Luenberger theory of functional observers, elements of probability theory, statistical modeling by the Monte Carlo method. Results: a new method for reducing computational costs is proposed, which uses the approximation of the Kalman filter transfer matrix time dependence by given piecewise linear functions according to the least squares criterion. The effectiveness of the method was evaluated on the example of a second-order dynamical system. On the basis of a comparative analysis, several acceptable variants of the considered approximation are proposed. The practical significance of the work lies in the further development of methods for the synthesis of quasi-optimal high-speed filters. The operability of the proposed modifications is confirmed by the example of a second-order linear dynamic system. The efficiency of the algorithms was evaluated by the statistical modeling method according to the criterion "accuracy-computational costs". It is shown that the total savings in the number of multiplication and addition operations can reach tens of times due to insignificant losses in the accuracy of the filtering process.

Keywords: Kalman filter; matrix gain; approximation; least squares criterion.

С. В. Гадецька¹, В. Ю. Дубницький², Ю. І. Кушнерук³, О. І. Ходирев²

¹ Харківський національний автомобільно-дорожній університет, Харків, Україна

² ННІ “Каразінський банківський інститут” ХНУ імені В. Н. Каразіна, Харків, Україна

³ Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків, Україна

EXCEL-ОРІЄНТОВАНИЙ КАЛЬКУЛЯТОР ДЛЯ ОБЧИСЛЕННЯ РЕЗУЛЬТАТІВ ЕНТРОПІЙНОГО АНАЛІЗУ ДАНИХ, ЩО РОЗПОДІЛЕНІ ПО КАТЕГОРІЯХ

Анотація. Мета роботи. Розробка EXCEL-орієнтованого калькулятора для обчислення результатів ентропійного аналізу даних, які розподілені по категоріях. **Предмет дослідження** – гістограми довільних законів розподілу та таблиці спряженості 2×2 . **Методи дослідження:** Ентропійний та інформаційний аналіз гістограм довільних законів розподілу та таблиць спряженості 2×2 . **Отримані результати.** Запропоновано використовувати методи ентропійного аналізу для аналізу даних, що розподілені по категоріях та наведено відомості про структуру excel-орієнтованого калькулятора, призначеного для виконання таких розрахунків. Калькулятор дає можливість обчислювати ентропійні характеристики гістограм, виконувати попарне порівняння ентропій гістограм, визначати відстань між гістограмами, обчислювати інформаційний коефіцієнт кореляції, порівнювати розбіжності між гістограмами. Для таблиць спряженості 2×2 калькулятор дає можливість оцінювати значущість взаємодії фактору рядків та фактору стовпців. Калькулятор визначає значення умовних ентропій для таблиць спряженості 2×2 . Запропонований калькулятор у деякій мірі заповнює прогалини в існуючих програмних продуктах та може бути використаний для обробки методами ентропійного аналізу даних, що розподілені по категоріях.

Ключові слова: ентропія; ентропійний аналіз; інформаційний коефіцієнт кореляції; відстань Хелінгера; відстань Кулльбака-Леблера.

Вступ

Припустимо, що кожен елемент множини $X = \{x_i\}; i=1, 2, \dots, N$ має властивість S_s за умови, що $S_s \in S$ та $|S| \leq |X|$. Ці властивості можуть бути визначені в одній з наступних шкал вимірювань:

- 1) номінальній або шкалі найменувань;
- 2) порядковій, або шкалі рангів;
- 3) інтервальній, або шкалі інтервалів.

Властивості, які визначені в шкалі рівних відносин, в даній роботі не розглядаються. Алгебраїчні властивості цих шкал і переважні області застосування розглянуто в [1]. Також розрізняють шкали на метричні та неметричні. Поділ шкал вимірювань на метричні та неметричні залежить від типу досліджуваного емпіричного об'єкта, зокрема, від властивостей вимірюваної величини. Метричні шкали – це шкали, у яких ознаки мають одиниці вимірювання. Неметричні шкали – це шкали, у яких немає одиниць вимірювань, наприклад, якісні ознаки. У загальному випадку результати вимірювань, які виконані з використанням цих шкал, прийнято називати даними, що розподілені по категоріях. В [2] наведено для таких даних наступне визначення: категоризовані дані – це дані, які представлено у вигляді частот спостережень, що потрапили в деякі категорії або класи. Для даних, які визначено в інтервальній шкалі, це визначення співпадає з визначенням гістограми [3, 4] і визначенням групованих або частково групованих вибірок [5]. У тому випадку, коли на категорії розподілено можливі значення двовимірної випадкової величини, то отриману гістограму називають таблицею спряженості $r \times c$.

Прийнято, що r визначає кількість рядків, c – стовпчиків. Відповідно до [3, 4] інтервальним статистичним розподілом можливих значень одновимірної випадкової величини (гістограмою) називатимемо дані, які представлено у вигляді табл. 1.

У цій таблиці прийнято, що нижні індекси L і R визначають ліву та праву межі інтервалів, n_1, n_2, \dots, n_s – кількість спостережень у відповідному інтервалі за умови, що загальна кількість спостережень дорівнює N .

У тому випадку, коли індивідуальні значення x_i невідомі, але відома тільки кількість елементів у кожному інтервалі, які вказано в табл. 1, такий спосіб формування гістограми в [5] названий частковим групуванням. В [6, 7] показано, що при великій (понад п'ятисот) кількості спостережень необхідно застосовувати методи обробки даних, що використовують поняття теорії інформації. Аналогічні висновки зроблені в [8]. В [9] відмічено, що теоретико-інформаційний аналіз експериментальних даних зарекомендував себе як ефективний інструмент дослідження широкого класу процесів і систем. Мірою кількості інформації в цьому підході служить зміна ентропії даних залежно від досліджуваних чинників. Сукупність таких методів отримала назву ентропійного аналізу. Оскільки в [10] було показано, що результати цього аналізу залежать від кількості інтервалів в гістограмі, то в рамках даної роботи порівнюватись будуть лише гістограми, що містять однакову кількість інтервалів.

В практиці медичних та біологічних досліджень набули широкого поширення таблиці спряженості $r \times c$. Їх загальний вигляд наведено в табл. 2.

Таблиця 1 – Структура гістограми можливих значень одновимірної випадкової величини

h	$x_{1L} \leq x < x_{1R}$	$(x_{2L} = x_{1R}) < x < x_{2R}$...	$(x_{sL} = x_{(s-1)R}) < x \leq x_{sR}$
N	n_1	n_2	...	n_s

Таблиця 2 – Структура таблиці спряженості $r \times c$

X	Y						
	y_1	y_2	...	y_j	...	y_c	
x_1	n_{11}	n_{12}	...	n_{1j}	...	n_{1c}	$n_{1\bullet} = \sum_{j=1}^c n_{1j}$
x_2	n_{21}	n_{22}	...	n_{2j}	...	n_{2c}	$n_{2\bullet} = \sum_{j=1}^c n_{2j}$
...
x_i	n_{i1}	n_{i2}	...	n_{ij}	...	n_{ic}	$n_{i\bullet} = \sum_{j=1}^c n_{ij}$
...
x_r	n_{r1}	n_{r2}	...	n_{rj}	...	n_{rc}	$n_{r\bullet} = \sum_{j=1}^c n_{rj}$
	$n_{\bullet 1} = \sum_{i=1}^r n_{i1}$	$n_{\bullet 2} = \sum_{i=1}^r n_{i2}$...	$n_{\bullet j} = \sum_{i=1}^r n_{ij}$...	$n_{\bullet c} = \sum_{i=1}^r n_{ic}$	N
$N = \sum_{i=1}^r \sum_{j=1}^c n_{ij} = \sum_{i=1}^r n_{i\bullet} = \sum_{j=1}^c n_{\bullet j}$							

У рамках даної роботи розглянемо таблиці спряженості 2×2 . Раніше їх використання для оцінювання ефективності деяких видів фармакологічних досліджень розглянуто в [26].

Таблиця 3 – Структура таблиці спряженості 2×2

a	b	$\alpha = a + b = n_{1\bullet}$
c	d	$\beta = c + d = n_{2\bullet}$
$\gamma = a + c = n_{\bullet 1}$	$\delta = b + d = n_{\bullet 2}$	N

Слід зауважити, що символи в комірках табл. 2 та табл. 3 позначають кількість даних, які в них потрапили без вказівок їх чисельних значень.

Різні способи обробки категоризованих даних включені до складу багатьох програмних продуктів, наприклад, AtteStat, Statgraphics, SociometryPro 2.3, TextusPro 1.0, TextAnalyst 2.01, WordStat 1.1, ContentAnalyzer 0.52., OCA, Vortex 10.7, IBM SPSS Statistics 26, Stadia 8.0, Statistica 13.3, StatPlus 5.0, DA-System 5.0, X7.2009. Слід зауважити, що лише в системі AtteStat використовуються деякі методи ентропійного аналізу гістограм, причому тільки для одновимірного закону розподілу. Тому розробку програмного забезпечення для ентропійного аналізу таблиць спряженості вигляду $r \times c$ і виконаного у вигляді EXCEL-орієнтованого калькулятора можна, на нашу думку, вважати актуальним завданням.

Аналіз літератури. Необхідні теоретичні відомості про аналіз таблиць спряженості наведено в роботах [2, 11]. Детально відомості про чисельні методи цього аналізу викладено в [11, 12]. В роботі [13] розглянуто способи об'єднання ентропійного аналізу і аналізу таблиць спряженості. На думку автора цієї роботи найбільш ефективним є поєднання методів ентропійного і логлінійного аналізу таблиць спряженості. Основні принципи логлінійного аналізу описано в роботі [14]. Логлінійний аналіз –

це, в основному, метод дослідження багатовимірних таблиць спряженості, тобто таких таблиць, що мають вигляд $r \times c \times k$. Цей метод дозволяє перевірити статистичну значущість різних чинників, присутніх в таблиці спряженості і їх взаємодій. Докладно зміст алгоритмів і чисельні приклади їх застосування описано в [15], їх програмна реалізація доступна в системах SPSS та STATISTICA.

Основна статистична гіпотеза, яку перевіряють при аналізі таблиць спряженості – це гіпотеза про відсутність взаємодії між фактором рядків і фактором стовпців. Якщо ця взаємодія відсутня, то початкові дані повинні бути розподілені рівномірно. Для перевірки цієї гіпотези найчастіше обчислюють величину χ^2 -квadrat із використанням співвідношення:

$$\chi_{\phi}^2 = \frac{(n_{ij} - E_{ij})^2}{E_{ij}}, \tag{1}$$

де E_{ij} – очікувана кількість (частота) об'єктів, що мають ознаки i та j одночасно в припущенні справедливості рівномірного розподілу:

$$E_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n}. \tag{2}$$

Якщо це припущення, що приймають як гіпотезу H_0 справедливо, то величина χ^2 -квadrat має $\nu = (r - 1)(c - 1)$ ступеней свободи. В цьому випадку правила для прийняття гіпотез мають вигляд:

$$T_1 = \begin{cases} H_0, & \text{якщо } \chi_{\phi}^2 < \chi_{\alpha, \nu}^2; \\ H_1, & \text{якщо } \chi_{\phi}^2 > \chi_{\alpha, \nu}^2 \end{cases} \tag{3}$$

або
$$T_2 = \begin{cases} H_0, & \text{якщо } P_{\nu_{\phi}} > P_{\nu_{\alpha}}; \\ H_1, & \text{якщо } P_{\nu_{\phi}} < P_{\nu_{\alpha}}, \end{cases} \tag{4}$$

де α – прийнятий рівень довірчої ймовірності, Pv_ϕ – фактично отримана довірча ймовірність.

Починати аналіз таблиць спряженості рекомендовано з використання діагностики Сімонова-Цая. Для цього обчислюють величину:

$$SC = \frac{(\chi^2(v, \alpha))^{1/2}}{3(E^2)^{3/2}} \sum_{i=1}^r \sum_{j=1}^c \frac{|(n_{ij} - E_{ij})|^3}{E_{ij}^2}. \quad (5)$$

В умові (5) прийнято, що $\chi^2(v, \alpha)$ – значення оберненої функції розподілу χ^2 з v ступенями свободи і рівнем довірчої ймовірності α . Якщо величина $SC > 0,25$, то використання критеріїв, заснованих на застосуванні величини χ^2 , не рекомендовано. У подальшому аналізі використовують декілька критеріїв.

Критерій відношення правдоподібності є таким:

$$G^2 = 2 \sum_{i=1}^r \sum_{j=1}^c n_{ij} \ln \frac{n_{ij}}{E_{ij}}. \quad (6)$$

При застосуванні цього критерію прийнято, що у випадку коли $n_{ij} = 0$, тобто спостереження з даним поєднанням чинників відсутні, то величину критерію для цієї комірки таблиці спряженості приймають рівною нулю. Величина G^2 розподілена відповідно до розподілу хі-квадрат з $\nu = (r-1)(c-1)$ ступенями свободи.

Критерій Зелтермана має вигляд:

$$D_z^2 = \chi^2 - \sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}}{E_{ij}} + rc, \quad (7)$$

де значення χ^2 обчислено, за співвідношенням (1). Статистика цього критерію має χ^2 -розподіл з $\nu = (r-1)(c-1)$ ступенями свободи при рівні довірчої ймовірності α . В [2] показано, що обчислювати величину χ^2 слід за співвідношенням:

$$\chi^2 = \frac{N(ad - bc)^2}{\alpha\beta\gamma\delta}. \quad (8)$$

Для таблиці спряженості 2×2 кількість ступенів свободи дорівнює $\nu = 1$. Більш точний результат обчислення величини χ^2 можна отримати, якщо використовувати його незсунену оцінку, визначену за співвідношенням:

$$\chi^2(\text{незс}) = \frac{N \left[\left| ad - bc \right| - \frac{1}{2} N \right]^2}{\alpha\beta\gamma\delta}. \quad (9)$$

Ентропію дискретної випадкової величини, закон розподілу якої визначено в табл. 1, визначають за добре відомим співвідношенням:

$$H = - \sum_{i=1}^s \frac{n_i}{n} \ln \left(\frac{n_i}{N} \right) = - \sum_{i=1}^s p_i \ln p_i. \quad (10)$$

Для визначення незсуненої оцінки \hat{H} ентропії H випадкової величини, гістограма якої отримана за наслідками спостережень, в [16] отримано співвідношення:

$$\begin{aligned} \hat{H} = & - \sum_{i=1}^s p_i \ln p_i - \frac{s-1}{N} + \\ & 1 - \sum_{i=1}^s p_i^{-1} + \sum_{i=1}^s (p_i^{-1} - p_i^{-2}) \\ & + \frac{1}{12N^2} + \frac{1}{12N^3}. \end{aligned} \quad (11)$$

Оцінка (11) відрізняється від співвідношення (10) тим, що містить поправку на зсування. Дисперсію ентропії визначають за співвідношенням:

$$D_H^2 = \frac{1}{N} \left(\sum_{i=1}^s p_i \ln^2 p_i - H^2 \right); \quad (12)$$

її незсунену оцінку \hat{D}_H^2 визначають як:

$$\hat{D}_H^2 = \frac{1}{N} \left(\sum_{i=1}^s p_i \ln^2 p_i - H^2 \right) + \frac{s-1}{2N^2}. \quad (13)$$

Середньоквадратичне відхилення величини ентропії d_H і його оцінку \hat{d}_H визначають за співвідношеннями:

$$d_H = \sqrt{D_H^2}, \quad \hat{d}_H = \sqrt{\hat{D}_H^2}. \quad (14)$$

Нижню H_d і верхню H_u межі довірчого інтервалу оцінки ентропії \hat{H} в роботі [16] запропоновано визначати за співвідношеннями:

$$H_d = \hat{H} - \Omega \left(\frac{1+\beta}{2} \right) \cdot \frac{\hat{d}_H}{\sqrt{s}} = \hat{H} - \gamma \cdot \frac{\hat{h}_H}{\sqrt{s}}; \quad (15)$$

$$H_u = \hat{H} + \Omega \left(\frac{1+\beta}{2} \right) \cdot \frac{\hat{d}_H}{\sqrt{s}} = \hat{H} + \gamma \cdot \frac{\hat{h}_H}{\sqrt{s}}. \quad (16)$$

У співвідношеннях (15) та (16) прийнято, що $\Omega(\bullet)$ – обернена функція стандартного нормального розподілу, в подальших обчисленнях за умовчанням прийнято, що $\beta=0,975, \gamma=1,96$. В [17] запропоновано обчислювати відносні оцінки ентропії: величину відносної ентропії

$$h = \frac{H}{H \max} = \frac{H}{\ln s} \quad (17)$$

та інформаційний індекс різноманітності

$$I_{div} = \frac{1}{N} \cdot \ln \left(N! / \prod_{i=1}^s n_i! \right). \quad (18)$$

У [18] для порівняння двох ентропій запропоновано варіант t -критерію Стьюдента у вигляді:

$$t_{факт} = \frac{|H_1 - H_2|}{\sqrt{(D_{H1}^2)^2 / N_1 + (D_{H2}^2)^2 / N_2}}, \quad (19)$$

Кількість ступенів свободи визначають як:

$$df = \left[\frac{(D_{H1}^2 + D_{H2}^2)^2}{(D_{H1}^2)^2 / N_1 + (D_{H2}^2)^2 / N_2} \right], \quad (20)$$

де $[A]$ – ціла частина числа A .

Методика ентропійного аналізу даних, які представлені в табл. 2, викладена в [8, 13, 19]. Відповідно до рекомендацій цих робіт безумовну ентропію даних визначали за співвідношенням:

$$H(X, Y) = - \sum_{i=1}^r \sum_{j=1}^c \frac{n_{ij}}{N} \ln \left(\frac{n_{ij}}{N} \right). \quad (21)$$

Ентропійну міру дисперсії визначали за співвідношенням:

$$\varepsilon = \left(N \ln N - \sum_{i=1}^r \sum_{j=1}^c n_{ij} \ln n_{ij} \right) / (N \ln(rc)). \quad (22)$$

Умовні ентропії визначали за співвідношеннями:

$$H_y = - \sum_{j=1}^c \frac{n_{\bullet j}}{N} \ln \frac{n_{\bullet j}}{N}; \quad (23)$$

$$H_x = - \sum_{i=1}^r \frac{n_{i\bullet}}{N} \ln \left(\frac{n_{i\bullet}}{N} \right); \quad (24)$$

$$H(y/x_i) = - \sum_{j=1}^c \frac{n_{ij}}{n_{i\bullet}} \ln \left(\frac{n_{ij}}{n_{i\bullet}} \right), \quad i = \overline{1, r}; \quad (25)$$

$$H(x/y_j) = - \sum_{i=1}^r \frac{n_{ij}}{n_{\bullet j}} \ln \left(\frac{n_{ij}}{n_{\bullet j}} \right), \quad j = \overline{1, c}. \quad (26)$$

Методи ентропійного аналізу дозволяють визначати міру статистичного зв'язку між парою змінних (X, Y) незалежно від шкал, в яких вони визначені, та не залежать від форми зв'язку між ними і закону їх розподілу. В [20] ця характеристика названа інформаційним коефіцієнтом кореляції. Чисельне значення інформаційного коефіцієнта кореляції $R(X, Y)$ визначають із співвідношень:

$$I = \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}}{N} \ln \frac{N \cdot n_{ij}}{n_{i\bullet} \times n_{\bullet j}}; \quad (27)$$

$$R(X, Y) = \sqrt{1 - \exp(-2I(X, Y))}. \quad (28)$$

В [21] запропоновано при обчисленні співвідношення (27) робити поправку на зсування. В цьому випадку слід використовувати співвідношення:

$$\hat{I} = \sum_{j=1}^c \sum_{i=1}^r \frac{n_{ij}}{N} \ln \frac{N \cdot n_{ij}}{n_{i\bullet} \cdot n_{\bullet j}} - \frac{(r-1)(c-1)}{2N}; \quad (29)$$

$$\rho(X, Y) = \sqrt{1 - \exp(-2\hat{I}(X, Y))}. \quad (30)$$

Порівняння інформаційного коефіцієнта кореляції з коефіцієнтами кореляції Пірсона і Спірмена – одна із задач цієї роботи.

В [13] для порівняння гістограм розподілу об'єктів, властивості яких визначено в однакових шкалах та які мають однакову кількість інтервалів, описано використання відстаней Хелінгера і Кулльбака-Леблера. Властивості і особливості визначення відстані Хелінгера для гістограм, які приведені в табл. 4, описано в [22, 31].

Таблиця 4 – Загальний вид порівнюваних гістограм

Гістограма закону розподілу випадкової величини P					
n_1^P	n_2^P	·	n_i^P	...	n_s^P
p_1	p_2	·	p_i	...	p_s
Гістограма закону розподілу випадкової величини Q					
n_1^Q	n_2^Q	...	n_i^Q	...	n_s^Q
q_1	q_2	...	q_i	...	q_s

Загальна кількість спостережень в кожній з гістограм визначена співвідношеннями:

$$\sum_{i=1}^s n_i^P = N_P; \quad \sum_{i=1}^s n_i^Q = N_Q. \quad (31)$$

Частка спостережень у кожному з інтервалів гістограм визначають за співвідношеннями:

$$p_i = n_i^P / N_P; \quad q_i = n_i^Q / N_Q. \quad (32)$$

Відстань Хелінгера між гістограмами P і Q визначають за співвідношенням, яке наведено в [31]:

$$D_H(P, Q) = \frac{1}{\sqrt{2}} \|\sqrt{P} - \sqrt{Q}\| = \frac{1}{\sqrt{2}} \sqrt{\sum_{i=1}^s (\sqrt{p_i} - \sqrt{q_i})^2}. \quad (33)$$

Для гістограм, які наведені в табл. 4, відстань Кулльбака-Леблера визначають за співвідношеннями:

$$D_{KL}(P, Q) = \sum_{i=1}^s p_i \ln \left(\frac{p_i}{q_i} \right); \quad (34)$$

$$D_{KL}(Q, P) = \sum_{i=1}^s q_i \ln \left(\frac{q_i}{p_i} \right).$$

З (34) витікає, що відстань Кулльбака-Леблера має важливу особливість – воно не симетричне, тобто $D_{KL}(P, Q) \neq D_{KL}(Q, P)$. Приклад застосування цієї відстані для кількісної міри оцінки відмінності між гістограмами різних розподілів приведено в [24]. В [25] запропонована методика перевірки статистичних гіпотез, яка заснована на використанні відстані Кулльбака-Леблера. Для викладу її основних положень в співвідношеннях (36)...(41) і далі будуть використані терміни і позначення, що прийняті в цій роботі. Задачею однієї вибірки для поліноміальної популяції в [25] названа задача в наступній постановці. Припустимо, що результати спостережень представлено у вигляді гістограми, наведеної в табл. 5. У цій таблиці прийнято, що в кожному i -ому інтервалі гістограми, $i = \overline{1, s}$, міститься x_i спостережень. Прийнята для аналізу отриманої гістограми статистична модель дозволяє визначити величину p_i – розрахункову частку кількості спостережень в кожному інтервалі за умови, що:

$$\sum_{i=1}^s x_i = N; \quad \sum_{i=1}^s p_i = 1. \quad (35)$$

Таблиця 5 – Умови задачі однієї вибірки

N	x_1	x_2	...	x_i	...	x_s
P	p_1	p_2	...	p_i	...	p_s

Необхідно перевірити гіпотезу H_0 , яка полягає в тому, що розподіл спостережень випадкової величини N не протирічить теоретичному розподілу випадкової величини P . В цьому випадку гіпотеза H_1 – альтернативна гіпотеза. В такій постановці задача співпадає з перевіркою гіпотези щодо закону розподілу випадкової величини по критерію χ^2 . Для розв’язання поставленої задачі в [25] пропонується обчислити дві величини:

$$\hat{I}(*; 2; O_N) = \sum_{i=1}^s x_i \ln \frac{x_i}{Np_i}; \quad (36)$$

$$\hat{J}(*; 2; O_N) = \left[\sum_{i=1}^s \left(\frac{x_i}{N} - p_i \right) \ln \frac{x_i}{Np_i} \right], \quad (37)$$

які мають розподіл χ^2 з $(s - 1)$ ступенем свободи. Задачею двох вибірок для поліноміальної популяції названа задача в наступній постановці. Для гістограм, які наведено в табл. 4, необхідно перевірити гіпотезу H_0 про те, що ці розподіли належать одній популяції. Для розв’язання поставленої задачі в [25] запропоновано обчислити величини:

$$2\hat{I}(H_1 : H_0) \approx \frac{1}{N_P N_Q} \sum_{i=1}^s \frac{(N_Q n_i^P - N_P n_i^Q)^2}{n_i^P + n_i^Q}; \quad (38)$$

$$\begin{aligned} \hat{J}(H_1 : H_0) &= \frac{1}{2(N_P + N_Q)^2} + \\ &+ \frac{1}{2N_P N_Q} \sum_{i=1}^s \frac{(N_Q n_i^P - N_P n_i^Q)^2}{n_i^P + n_i^Q} + \\ &+ \sum_{i=1}^s \frac{(N_Q n_i^P - N_P n_i^Q)^2 (n_i^P + n_i^Q)}{n_i^P n_i^Q}, \end{aligned} \quad (39)$$

які мають розподіл χ^2 з $(s - 1)$ ступенем свободи. Для таблиць спряженості типів 2×2 (табл. 2) і $r \times c$ (табл. 3) при перевірці гіпотези H_0 , яка полягає в перевірці відсутності зв'язку між ефектами рядків і стовпців (альтернативна гіпотеза H_1 – присутній зв'язок між ефектами рядків і стовпців), слід обчислити співвідношення:

$$2\hat{I}(H_1 : H_0) = \sum_{i=1}^r \sum_{j=1}^c \left(n_{ij} - \frac{n_{i\bullet} n_{\bullet j}}{N} \right)^2 / \frac{n_{i\bullet} n_{\bullet j}}{N}; \quad (40)$$

$$\hat{J}(H_1 : H_0) = \frac{1}{2} \left[\sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - n_{i\bullet} n_{\bullet j} / N)^2}{n_{i\bullet} n_{\bullet j} / N} + \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - n_{i\bullet} n_{\bullet j} / N)^2}{n_{ij}} \right]. \quad (41)$$

В [25] показано, що чисельні значення співвідношень (40) і (41) мають розподіл χ^2 з $(r - 1)(c - 1)$ ступенями свободи. Гіпотези, які сформульовані за співвідношеннями (36, 37), (38, 39), (40, 41), приймаються тільки при сумісному виконанні кожної з гіпотез, що входить у відповідну пару.

Приклади застосування відстані Кульбака-Лейблера для аналізу фінансової інформації і рішення задач проектування критичних систем описані в [27, 28, 29, 31]. При підготовці даного повідомлення автори не змогли знайти відомості про програмні продукти, що реалізують ентропійні методи аналізу таблиць спряженості.

Мета роботи. Розробка EXCEL-орієнтованого калькулятора для обчислення результатів ентропійного аналізу даних, які розподілені по категоріях.

Предмет дослідження – гістограми довільних законів розподілу та таблиці спряженості 2×2 .

Методи дослідження: Ентропійний та інформаційний аналіз гістограм довільних законів розподілу та таблиць спряженості 2×2 .

Отримані результати

На думку авторів даного повідомлення, систему EXCEL можна розглядати як одну з найбільш зручних платформ для обробки даних про функцію щільності ймовірностей, які представлено в табличній формі, тобто у вигляді гістограми. В даній роботі процес обробки, даних побудовано на використанні "розумної таблиці" – інструменту, який дозволяє спростити формули з посиланнями на її елементи. Головною перевагою "розумної таблиці" є те, що при зміні кількості рядків у ній формули автоматично налаштовуються на ці зміни. Крім початкових даних таблиця містить стовпці з формулами для обчислення проміжних результатів.

На головному аркуші робочої книги Excel розміщується кнопка виклику меню програми, яке містить перелік задач (рис. 1). Як вікно меню, так і вікна з результатами задач створено за допомогою форм – стандартного інструменту MS Excel, що значно спрощує режим діалогу. Для виклику на екран результатів будь-якої задачі треба її виділити і натиснути кнопку "Пуск", розміщену у вікні меню праворуч. При цьому вікно меню з екрана вилучається, а на його місці виводиться вікно з результатами відповідної задачі. Після цього знову можна викликати вікно меню і отримати результати іншої задачі. Вікна з результатами різних задач можна розмістити на екрані поруч одне з одним, або закривати непотрібні.

Перший пункт меню відрізняється від інших задач, бо здійснює перехід на аркуш з таблицею, у яку треба ввести початкові дані. Введення у таблицю нових даних складається у заповненні стовпця "Кількість елементів n_i ". Якщо нових даних більше, ніж у існуючій таблиці, розрахункові формули автоматично додаються у нові рядки. Якщо в таблиці виявляються зайві рядки, то їх треба вилучити командою "вилучити комірки". Варіант застосування інструмента "Розумна таблиця" показано на рис. 2.

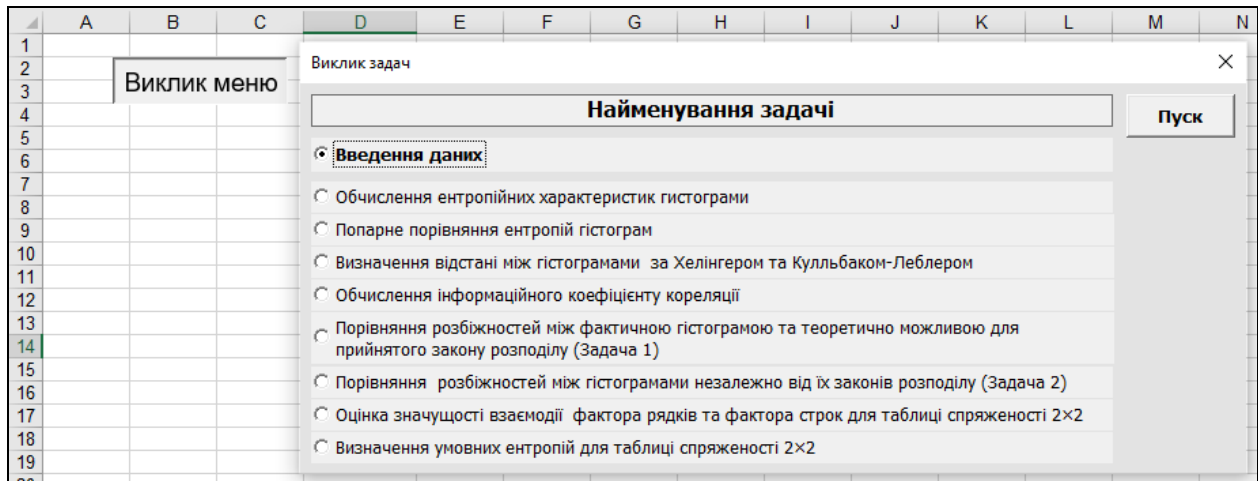


Рис. 1. Скрін-копія аркуша таблиці з меню програми
(Fig.1. Screen – a copy of the screen window for the table sheet from the program menu)

№ Класу	Кількість елементів n_i	$P_i = n_i / \sum n_i$	$\ln(P_i)$	$P_i \ln(P_i)$	$\ln(P_i)^2$	$P_i \ln(P_i)^2$	корінь(P_i)
1	10	0,058823529	-2,833213344	-0,166659608	8,027097853	0,472182227	0,242535625
2	20	0,117647059	-2,140066163	-0,25177249	4,579883184	0,538809786	0,34299717
3	50	0,294117647	-1,223775432	-0,35993395	1,497626307	0,440478326	0,542326145
4	35	0,205882353	-1,580450376	-0,325386842	2,49782339	0,514257757	0,453742606
5	28	0,164705882	-1,803593927	-0,297062529	3,252951053	0,535780173	0,405839725
6	15	0,088235294	-2,427748236	-0,21421308	5,893961497	0,520055426	0,297044263
7	12	0,070588235	-2,650891787	-0,187121773	7,027227268	0,496039572	0,265684466

Рис. 2. Скрін-копія вікна екрану варіанта застосування функції "Розумна таблиця"
(Fig. 2. Screen – a copy of the screen window of the option to use the "Dynamic Table" function)

Для виконання ентропійного аналізу одновимірних гістограм їх представляють у вигляді, який показано в табл. 6.

Таблиця 6 – Приклад подання гістограм

Гістограми	Комірки гістограм							Усього спостережень
	1	2	3	4	5	6	7	
Г1	10	20	50	35	28	15	12	170
Г2	24	24	24	25	25	24	24	170

Калькулятор обробляє одну, декілька або всі введені гістограми згідно з вказівкою користувача. Для прикладу застосування калькулятора обрано гістограму Г1, яку наведено в [4], та гістограму Г2, що відповідає рівномірному закону розподілу. Результати обчислення ентропійних характеристик гістограми Г1 показано на рис. 3. Для їх визначення було використано співвідношення (10)...(18).

Для попарного порівняння ентропій гістограм Г1 ($H1$) і Г2 ($H2$) використано критерій Хатчинсона, який визначений співвідношеннями (19), (20). Як нульова прийнята гіпотеза $H0: H1=H2$, як альтернативна прийнята гіпотеза $H1: H1 \neq H2$. У цьому випадку правила для прийняття гіпотез мають вигляд:

$$T_3 = \begin{cases} H_0, & \text{якщо } Pv(t_\phi) > Pv_\alpha; \\ H_1, & \text{якщо } Pv(t_\phi) < Pv_\alpha. \end{cases} \quad (42)$$

У співвідношенні (42) величину $Pv_\phi(t)$ визначають, використовуючи умову:

$$Pv_\phi(t) := \text{СТЬЮДРАСП}(t_\phi; df; 2). \quad (43)$$

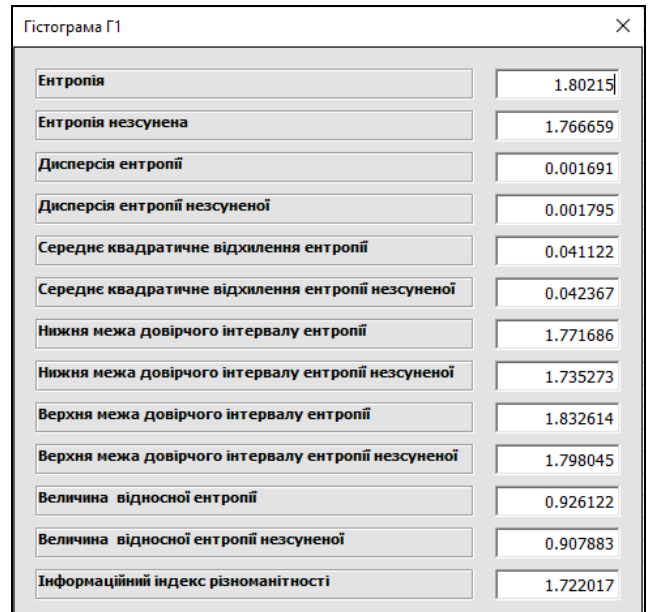


Рис. 3. Скрін-копія вікна екрану обчислення ентропійних характеристик гістограми Г1
(Fig. 3. Screen – a copy of the screen window for calculating the entropy characteristics of the histogram Г1)

Величину t_ϕ визначали, використовуючи співвідношення (19), величину df – (20). За замовченням величина $\alpha=0,05$. Результати порівняння гістограм Г1 і Г2 показані на рис. 4. Ентропія $H1$ відповідає гістограмі Г1, ентропія $H2$ відповідає гістограмі Г2. Всі обчислення виконують для зсуненої та незсуненої оцінок гістограм.

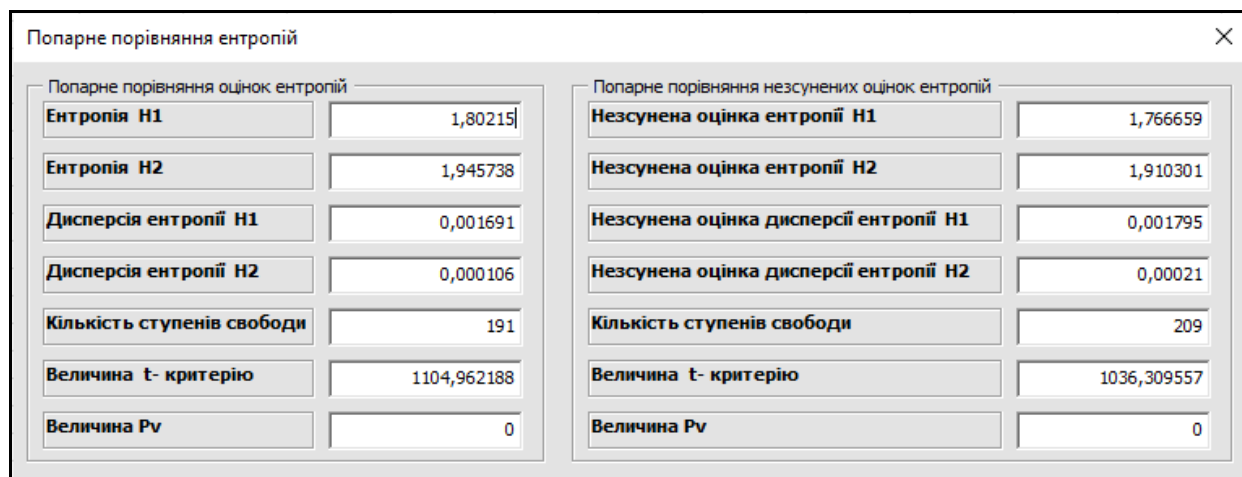


Рис. 4. Скрін-копія вікна екрану попарного порівняння ентропій гістограм Г1 (H1) і Г2 (H2)
 (Fig. 4. Screen – a copy of the window of the screen for pairwise comparison of the entropies of the histograms Г1 (H1) and Г2 (H2))

Як і слід було чекати, за наслідками порівняння слід прийняти альтернативну гіпотезу H1. Це не протирічить фізичному змісту отриманого результату тому, що гістограми Г1 і Г2 відповідають абсолютно різним законам розподілу. В рамках даної роботи прийняття статистичних гіпотез, пов'язане з використанням величини χ^2 , виконують згідно із співвідношенням (4), обґрунтування використання величини P_v та методи її визначення наведено в [37]. Для аналізу таблиць спряженості 2x2 цю величину визначають за співвідношенням:

$$P_v(\chi^2) := \text{ХИРАСП}(\chi^2; 1). \quad (44)$$

Для обчислення відстані між парами гістограм використовували відстань Хелінгера, яку визначали згідно із співвідношеннями (31)...(33) і відстань Кульбака-Левлера, яку визначали згідно із співвідношеннями (31), (32), (34). Результати обчислення показано на рис. 5.

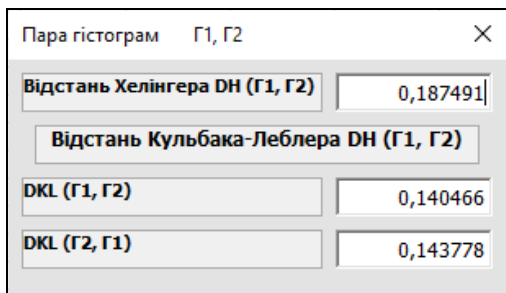


Рис. 5. Скрін-копія вікна екрану визначення відстані між гістограмами Г1 і Г2
 (Fig. 5. Screen – a copy of the screen window for determining the distance between histograms Г1 and Г2)

До методів ентропійного аналізу відносять також і обчислення інформаційного коефіцієнта кореляції. Способи його обчислення описано в [20, 21]. Для порівняння інформаційного коефіцієнта кореляції, обчисленого за співвідношеннями (27) та (28) з коефіцієнтом кореляції Пірсона було проведено чисельний експеримент. В роботах [32...34] наведено кореляційні таблиці і значення коефіцієнтів кореляції Пірсона або інформаційного коефіцієнта

кореляції. Для кожної з цих таблиць (табл. 7 ... 9) було обчислено недостатні коефіцієнти. У комірках цих таблиць поставлено відповідну кількість спостережень.

Для якісного оцінювання величини коефіцієнту кореляції використаємо так звану шкалу Чеддока (1879-1940), табл. 10. Незважаючи на те, що цю шкалу застосовують у багатьох роботах, наприклад [30], посилання на першоджерело як правило відсутні. Автори цього повідомлення із задоволенням заповнюють цю прогалину і відсилають читачів на роботу [35].

Результати обчислень показано у табл. 11.

Таблиця 7 – Кореляційна таблиця, побудована за даними роботи [32]

Умовні індекси змінної X	Умовні індекси змінної Y							
	A	B	C	D	E	F	G	H
1							1	
2							1	
3						1		
4						1	1	1
5					2	1		
6				1	1	4		
7			1	4	2			
8	1	3	9		3			
9		2	7	1	1			
10		1	5		1			
11			2		1	1		

Таблиця 8 – Кореляційна таблиця, побудована за даними роботи [33]

Умовні індекси змінної X	Умовні індекси змінної Y							
	A	B	C	D	E	F	G	H
1		1						
2	1	6						
3		5	9					
4		4	9	5	3			
5			7	9	3			
6			2	9	12	4	2	
7				7	8	6	4	
8				1	2	5	3	3
9						1	2	2

Таблиця 9 – Кореляційна таблиця, побудована за даними роботи [34]

Умовні індекси змінної X	Умовні індекси змінної Y							
	A	B	C	D	E	F	G	H
1	1	5	7	4				
2			1	5	2			
3			1	1	6	3	2	
4					2	4	2	
5					1	1		

Таблиця 10 – Якісна оцінка величини коефіцієнта кореляції (шкала Чеддока)

Значення коефіцієнта кореляції	[0,1...0,3)	[0,3...0,5)	[0,5...0,7)
Зв'язок	незначний	помірний	істотний
Рівень	1	2	3

Значення коефіцієнта кореляції	[0,7...0,9)	[0,9...0,99]	1,0
Зв'язок	високий	дуже високий	Функціональний
Рівень	4	5	6

Таблиця 11 – Чисельні значення коефіцієнту кореляції Пірсона та інформаційного коефіцієнту кореляції

Показники вимірів Статистичного зв'язку ^{*)}	Літературні посилання		
	[32]	[33]	[34]
\hat{I}	0,884	0,862	0,876
ρ	0,547(3)	0,782(4)	0,730(4)
r	0,669 (4)	0,832(4)	0,859(4)

^{*)}Жирним шрифтом позначена якісна оцінка величини коефіцієнта кореляції за шкалою Чеддока (табл. 10).

З цієї таблиці випливає, що інформаційний коефіцієнт кореляції оцінює на якісному рівні зв'язок між випадковими величинами так само, як і традиційний коефіцієнт кореляції Пірсона. Для оцінки

Таблиця 13 – Чисельні значення статистичних критеріїв, які було використано при аналізі випадкових таблиць спряженості 2×2

№ таблиці	Критерій хі-квадрат (I) ^{••}		Критерій хі-квадрат незсунений(II) ^{••}		Критерій відношення гравдоподібності (III) ^{••}		Критерій Зелтермана (IV) ^{••}		Інформаційний коефіцієнт кореляції (V) ^{••}	Критерій Цоя-Самена
	Чисельне значення	Величина P _v	Чисельне значення	Величина P _v	Чисельне значення	Величина P _v	Чисельне значення	Величина P _v		
1	5,627	0,012	4,531	0,035	5,763	0,016	5,633	0,017	0,265	0,09
2	0,209	0,646	0,04	0,839	0,209	0,646	0,205	0,65	0,051	0,1
3	1,516	0,218	1,02	0,313	1,521	0,217	1,517	0,218	0,136	0,07
4	0,017	0,893	0,012	0,913	0,017	0,893	0,013	0,906	0,015	0,09
5	0,07	0,932	0,02	0,876	0,007	0,932	0,008	0,017	0,010	0,08
6	0,017	0,894	0,01	0,911	0,017	0,894	0,022	0,281	0,015	0,09
7	0,764	0,381	0,36	0,549	0,791	0,373	0,85	0,356	0,099	0,124
8	0,187	0,665	0,03	0,867	0,185	0,666	0,15	0,697	0,048	0,117
9	4,546	0,03	3,63	0,057	4,632	0,031	4,567	0,032	0,236	0,07
10	0,292	0,588	0,09	0,76	0,29	0,589	0,274	0,6	0,060	0,089

застосування інформаційного коефіцієнта кореляції при аналізі таблиць спряженості був проведений чисельний експеримент, в процесі якого для отриманих випадковим чином таблиць обчислювали критерії, визначені співвідношеннями (5), (7), (8), (9), (27), (28). Детальний виклад способу отримання цих таблиць у цій роботі не розглянуто. Для табл.11 з урахуванням позначень, які наведено в табл. 3, кількість інформації визначали, використовуючи співвідношення:

$$I = \frac{a}{N} \ln \frac{aN}{\alpha\lambda} + \frac{b}{N} \ln \frac{bN}{\alpha\delta} + \frac{c}{N} \ln \frac{cN}{\gamma\beta} + \frac{d}{N} \ln \frac{dN}{\beta\delta} \quad (45)$$

Вихідні дані для експерименту, враховуючи позначення, прийняті в табл. 3, наведені у табл. 12.

Таблиця 12 – Елементи таблиць спряженості 2×2, отримані випадковим методом

№ таблиці	Елементи таблиці				№ таблиці	Елементи таблиці			
	a	b	c	d		a	b	c	d
1	17	23	7	32	11	25	14	17	24
2	27	31	11	10	12	24	7	29	20
3	19	24	22	16	13	22	10	21	27
4	15	32	11	22	14	9	34	6	38
5	12	14	25	28	15	40	11	13	16
6	22	10	33	16	16	40	6	8	26
7	5	17	19	39	17	15	26	25	14
8	8	15	17	40	18	30	14	6	31
9	24	22	10	25	19	30	19	5	26
10	12	17	18	33	20	9	18	28	25

Чисельні значення критеріїв, які використовували при аналізі таблиць спряженості, наведено в табл. 13.

Оскільки значення критерію Цоя-Самена менше величини 0,25, то виконання подальшого аналізу критеріїв (I) ... (V) вважатимуться коректним. Для визначення зв'язку між цими критеріями побудована матриця їх взаємної кореляції, яка наведена в табл. 14.

Закінчення табл. 13

11	4,107	0,042	3,25	0,071	4,146	0,041	4,109	0,042	0,225	0,073
12	2,823	0,093	2,07	0,15	2,914	0,087	2,883	0,089	0,189	0,092
13	4,827	0,028	3,87	0,049	4,913	0,026	4,842	0,027	0,244	0,078
14	0,81	0,367	0,38	0,537	0,814	0,366	0,806	0,369	0,097	0,132
15	9,336	0,002	7,89	0,005	9,224	0,002	9,202	0,002	0,330	0,095
16	32,77	$<1 \cdot 10^{-3}$	30,18	$<1 \cdot 10^{-3}$	34,95	$<1 \cdot 10^{-3}$	32,69	$<1 \cdot 10^{-3}$	0,595	0,08
17	6,053	0,014	5	0,025	6,132	0,014	6,053	0,014	0,272	0,073
18	21,98	$<1 \cdot 10^{-3}$	19,93	$<1 \cdot 10^{-3}$	23,44	$<1 \cdot 10^{-3}$	22,001	$<1 \cdot 10^{-3}$	0,272	0,074
19	15,69	$<1 \cdot 10^{-3}$	14,01	0,002	16,82	$<1 \cdot 10^{-3}$	15,741	$<1 \cdot 10^{-3}$	0,435	0,08
20	2,735	0,098	2,01	0,1562	2,777	0,095	2,75	0,096	0,185	0,086

**) Умовні індекси критеріїв

Таблиця 14 – Коефіцієнти взаємної кореляції між критеріями, які використані при аналізі таблиць спряженості 2×2

Умовні індекси критеріїв	Умовні індекси критеріїв				
	I	II	III	IV	V
I	1	0,999	0,995	0,993	0,987
II		1	0,995	0,992	0,986
III			1	0,998	0,992
IV				1	0,993
V					1

З даних, наведених у цій таблиці, випливає, що інформаційний критерій кореляції рядків та стовпців таблиці спряженості може бути обґрунтовано використаний при аналізі цих таблиць. Перевага його перед іншими критеріями в тому, що він не тільки визначає наявність (відсутність) статистичного зв'язку між факторами рядків та стовпців, а й дає її кількісну оцінку. Скрін-копія вікна екрана з результатами обчислення інформаційного коефіцієнту кореляції згідно із (27), (28) показана на рис. 6.

Параметри таблиці спряженості				
	a	b	c	d
	4	16	1	21
Критерій χ^2 -квадрат	2.385848			
Величина P_v	0.1224			
Критерій χ^2 -квадрат нузсунений	1.139779			
Величина P_v	0.6236			
Інформаційний критерій кореляції	0.240851			

Рис. 6. Скрін-копія вікна екрана з результатами обчислення інформаційного коефіцієнту кореляції (Fig. 6. Screen – a copy of the screen window with the results of the information correlation coefficient calculation)

Приклади розв'язання задач визначення значущості розбіжностей між гістограмами згідно із співвідношеннями (36)...(39) наведено на рис. 7, 8.

Приклад оцінки значущості взаємодії фактора рядків та фактора стовпчиків для таблиці спряженості 2×2 згідно із співвідношеннями (40) та (41) пркааний на рис. 9.

Комірки, S	Фактична кількість спостережень, X	Кількість спостережень згідно з розрахунком
1	4	2
2	7	11
3	28	25
4	35	31
5	16	21
6	6	8
7	4	2

Кількість степеней свободи: 6

Мінімум розрізняючої інформації, I

χ^2 -квадрат: 3.7251 P_v : 0.71382

Оцінка розбіжності, J

χ^2 -квадрат: 7.3409 P_v : 0.2905

Рис. 7. Скрін-копія вікна екрана з результатами порівняння розбіжностей між фактичною гістограмою та теоретично можливою для прийнятого закону розподілу, Задача 1 (Fig. 7. A screenshot of the screen copy of the results of the difference between the actual histogram and the theoretically possible for the adopted law of the difference, Task 1)

Комірки, S	Кількість спостережень у комітках гістограми Г1	Кількість спостережень у комітках гістограми Г2
1	10	24
2	20	24
3	50	24
4	35	25
5	28	25
6	15	24
7	12	24

Кількість степеней свободи: 6

Мінімум розрізняючої інформації, I

χ^2 -квадрат: 23.1768 P_v : 0.0007

Оцінка розбіжності, J

χ^2 -квадрат: 77.582914 P_v : 0.000112

Рис. 8. Скрін-копія вікна екрана з результатами порівняння розбіжностей між гістограмами незалежно від їх законів розподілу, Задача 2 (Fig. 8. A screen copy of the screen window with the results of comparing the differences between histograms regardless of their distribution laws, Task 2)

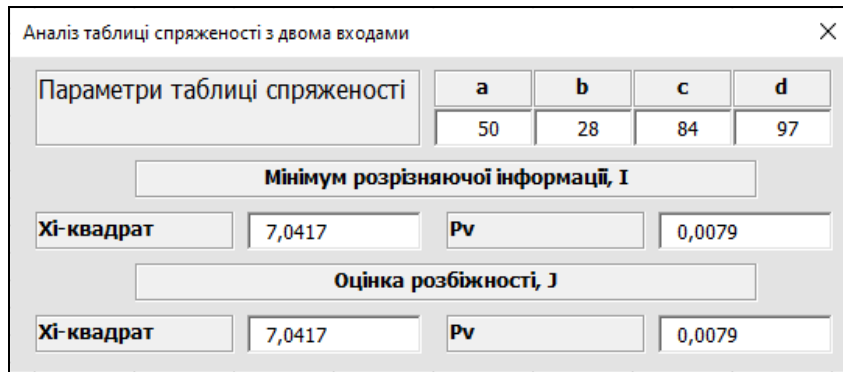


Рис. 9. Скрін-копія вікна екрану з результатами значущості взаємодії фактора рядків та фактора стовпчиків для таблиці спряженості 2×2
(Fig. 9. Screen copy of the screen window with the results of the significance of the interaction of the row factor and the term factor for the contingency table 2×2)

Приклад визначення умовних ентропій для таблиці спряженості 2×2 показаний на рис. 10. Прийняті в цьому прикладі позначення та співвідношення, за якими їх розраховували, наведено в табл. 15.

Приклад визначення умовних ентропій для таблиці спряженості 2×2 показаний на рис. 10. Прийняті в цьому прикладі позначення та співвідношення, за якими їх розраховували, наведено в табл. 15.

Таблиця 15 – Позначення умовних ентропій та співвідношення, прийняті для їх розрахунку

Показники ентропії таблиці спряженості 2×2	Визначення у вікні калькулятора	Розрахункове співвідношення
Безумовна ентропія, $H(X,Y)$	$H(X,Y)$	(21)
Ентропійна міра дисперсії, ϵ	ϵ	(22)
Умовна ентропія, H_y	$\{H(Y)\}$	(23)
Умовна ентропія, H_x	$\{H(X)\}$	(24)
Умовна ентропія $H[Y(y/x_1)]$	$H(y/x_1)$	(25)
Умовна ентропія $H[Y(y/x_2)]$	$H(y/x_2)$	(25)
Умовна ентропія $H[X(x/y_1)]$	$H(x/y_1)$	(26)
Умовна ентропія $H[X(x/y_2)]$	$H(x/y_2)$	(26)



Рис. 10. Скрін-копія вікна екрану з результатами визначення умовних ентропій для таблиці спряженості 2×2
(Fig. 10 Screen copy of the screen window with the results of determining the conditional entropies for the contingency table)

Висновки

1. Запропоновано використовувати методи ентропійного аналізу для аналізу даних, що розподілені по категоріях та наведено відомості про структуру Excel-орієнтованого калькулятора, призначеного для виконання таких розрахунків.

2. Калькулятор дає можливість обчислювати ентропійні характеристики гістограм, виконувати попарне порівняння ентропій гістограм, визначати відстань між гістограмами, обчислювати інформаційний коефіцієнт кореляції, порівнювати розбіжності між гістограмами.

3. Для таблиць спряженості 2×2 калькулятор дає можливість оцінювати значущість взаємодії фактору рядків та фактору стовпчиків.

4. Калькулятор визначає значення умовних ентропій для таблиць спряженості 2×2.

5. Запропонований калькулятор у деякій мірі заповнює прогалини в існуючих програмних продуктах та може бути використаний для обробки методами ентропійного аналізу даних, що розподілені по категоріях.

6. В роботі показано, що ентропійні методи аналізу гістограм доцільно використовувати у випадках, коли гістограми визначають довільні закони розподілу.

REFERENCES

- Motalo, V. (2015), "Analysis of measurement scales", *Measuring technique and metrology*, 2015, No. 76, pp. 21-35, available at: http://nbuv.gov.ua/UJRN/metrolog_2015_76_4.
- Kendall, Maurice G. and Stuart, Alan (1961). *The Advanced Theory of Statistics. Vol. 2, Inference and Relationship*. Charles Griffin, London, 676 p.
- Dubnitsky, V. Yu., Kobylin, A. M. and Kobylin, O. A. (2018), "Estimation of the lower bound of the reliability of a physically realizable system during its operation under arbitrary distribution laws of the generalized load and strength", *Information processing systems*, 2018, No. 1(152), pp. 53-60, doi: <https://doi.org/10.30748/soi.2018.152.08>.

4. Zhluktenko, V. I., Nakonechnyi, S. I. and Savina, S. S. (2001), “Probability Theory and Mathematical Statistics. Part II. *Mathematical statistics*, KNEU, Kyiv, 336 p., available at: https://www.studmed.ru/zhluktenko-v-nakonechniy-s-savna-ss-teorya-ymovnostey-matematichna-statistika-u-2-h-ch-ch-matematichna-statistika_3976c660ed4.html.
5. Kulldorff, Gunnar (1961), *Contributions to the Theory of Estimation from Grouped and Partially Grouped Samples*. Almqvist & Wiksell / John Wiley & Sons, Stockholm, 176 p., available at: <https://www.amazon.com/Contributions-Estimation-Grouped-Partially-Samples/dp/B0010VDVR26>.
6. Jun I. V. (1993), “On the number of gradations of histograms of errors in astronomical observations”, *Kinematics and physics of celestial bodies*, No. 1, vol. 9, pp. 88-92, available at: <https://www.mao.kiev.ua/biblio/jscans/kfnt/1993-09/kfnt-1993-09-1-11.pdf>.
7. Jun, I. V. “Mathematical processing of astronomical and space information with non-Gaussian observation errors: Abstract of the thesis for the competition uch. doctorate degrees. Phys.-Math. sciences: spec. 01.03.01 "Astrometry and Celestial Mechanics”, Kyiv, GAO NAS of Ukraine, 1992, 46 p., available at: <https://issuu.com/blindguardian/docs/asd>.
8. Paniotto, V. I., Maksymenko, V. S. and Kharchenko, N.M. (2004), *Statistical analysis of sociological data*, KM Academy, Kyiv, 2004, 270 p.
9. Tsvetkov, O. V. (2015), *Entropy analysis of data in physics, biology, and technology*, LETI, SPb, 202 p., available at: https://www.researchgate.net/profile/Oleg-Tsvetkov/publication/331686300_entropijnyj_analiz_dannyh_v_fizike_biologii_i_tehnike/links/5c87f3afa6fdcc38174f8a14/entropijnyj-analiz-dannyh-v-fizike-biologii-i-tehnike.pdf.
10. Dubnickij, V. Ju., Filatova, L. D. and Khodyrev, A. I. (2017), “The stability of the estimate of the entropy of the histogram of a continuous random variable with respect to the change in the number of its intervals”, *Control, Navigation and Communication Systems*, No 5 (45), pp. 42-46, available at: http://nbuv.gov.ua/UJRN/suntz_2017_5_12.
11. Agresti, A. (2002), *Categorical data analysis*, John Wiley & Sons Inc., New York, 742 p., available at: <https://onlinelibrary.wiley.com/doi/book/10.1002/0471249688>.
12. Joseph L., Fleiss, Bruce, Levin and Myunghee Cho, Paik (2003), *Statistical Methods for Rates and Proportions*, John Wiley & Sons, Inc. New York, 768 p., available at: <https://onlinelibrary.wiley.com/doi/book/10.1002/0471445428>.
13. Nobuoki, Eshima (2020), *Statistical Data Analysis and Entropy*, Springer Nature Singapore Pte Ltd, Singapore, 498 p., available at: <https://link.springer.com/book/10.1007/978-981-15-2552-0>.
14. Graham J.G., Upton (1978), *The Analysis of Cross-tabulated Data*, J. Wiley, New York, 160 p., available at: <https://www.amazon.com/Analysis-Cross-tabulated-Data-Graham-Upton/dp/0471996599>.
15. Duncan, Crammer (2003), *Advanced Quantitative Data Analysis*, Open University Press, Philadelphia, 272 p., available at: <https://www.amazon.com/Advanced-Quantitative-Analysis-Understanding-Research/dp/0335200591>.
16. Anne E., Magurran (1983), *Ecological Diversity and its Measurement*, London, Sydney, CROOM HELM Royal Society University Research Fellow University, 184 p., available at: <https://link.springer.com/book/10.1007/978-94-015-7358-0>.
17. Margalef, R. (1958), “Information theory in ecology”, *Gen. Syst.*, No 3, pp. 36-71, available at: [https://www.scirp.org/\(S\(351jmbntvnsjtl1aadkposzje\)\)/reference/ReferencesPapers.aspx?ReferenceID=1134401](https://www.scirp.org/(S(351jmbntvnsjtl1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=1134401).
18. Hutcheson, K. (1970), “A Test for Comparing Diversities Based on the Shannon Formula”, *Journal of Theoretical Biology*, vol. 29, pp. 151- 154, doi: [http://dx.doi.org/10.1016/0022-5193\(70\)90124-4](http://dx.doi.org/10.1016/0022-5193(70)90124-4).
19. Michalowicz, J.V., Nichols, J.M. and Bucholtz, F. (2014), *Handbook of differential entropy*, Taylor & Francis Group, LLC, London, 241 p., available at: <https://www.routledge.com/Handbook-of-Differential-Entropy/Michalowicz-Nichols-Bucholtz/p/book/9781138374799>.
20. Linfoot, E. and Linfoot, E. H. (1957), “An Informational Measure of Correlation”, *Information and Control*, vol. 1, No. 1, pp. 85-89, available at: <https://www.sciencedirect.com/science/article/pii/S001999585790116X>.
21. Vistelius, A. B. (1960), “The skew frequency distributions and the fundamental law of Geochemical processes”, *The Journal of Geology*, vol. 68, No. 1, pp. 1-22, available at: <https://www.jstor.org/stable/30058252>.
22. Marian, P. and Marian, T. A. (2015), “Hellinger distance as a measure of Gaussian discord”, *The Journal of Physics A: Math. Theor.*, 48:11 (2015), 115301, 21 p., arXiv: 1408.4477, doi: <http://dx.doi.org/10.1088/1751-8113/48/11/115301>.
23. Kullback, S. and Leibler, R.A. (1951), “On information and sufficiency”, *The Annals of Mathematical Statistics*, Vol. 22. No. 1. pp. 79-86, doi: <http://dx.doi.org/10.1214/aoms/1177729694>.
24. Dubnytskyi, V. Yu., Skorykova, I. G. and Khodyrev, O. I. (2017), “Optimal approximation of the distribution density function according to the minimum information loss criterion”, *Information processing systems*, No. 4, pp. 45-51, available at: <https://www.hups.mil.gov.ua/periodic-app/article/17653>.
25. Solomon, Kullback (1978), *Information Theory and Statistics*, Peter Smith, Gloucester, Mass, 399 p., available at: https://books.google.com.ua/books/about/Information_Theory_and_Statistics.html?id=XeRQAAAAMAAJ&redir_esc=y.
26. Hadetska, S. V., Dubnytskyi, V. Yu., Kushneruk, Yu. I. and Hodyrev, O. I. (2020), “A specialized software calculator for evaluating the clinical informativeness of laboratory tests”, *Advanced Information Systems*, No. 2, Vol. 4, pp. 80-84, doi: <https://doi.org/10.20998/2522-9052.2020.2.12>.
27. Soloshenko, O. M. (2014), “Study of the Kullback-Leibler distance in modeling problems in credit scoring”, *Development of information-resource support for education and science in the mining and metallurgical industry and in transport*, September 27-28, 2014, Dnepropetrovsk, pp. 328-333, available at: <https://ir.nmu.org.ua/handle/123456789/150310>.
28. Dubnytskyi, V. Yu., Krylenko, I. M., Fesenko, G. V. and Cherepnev, I. A. (2017), “The history of the development of means of eye protection for military personnel in combat conditions and modern requirements for controlling their impact resistance”, *Weapons and military equipment systems*, No. 1(49), pp. 23-37, available at: <https://www.hups.mil.gov.ua/periodic-app/article/17565>.
29. Brovko, D. V. (2020), “Construction of a system for monitoring the reliability of elements of the built and constructed surface complex of mines based on entropy estimation”, *Mining Bulletin: Scientific and Technical. coll.*, Kryvyi Rih, Issue 107, pp. 73–83, doi: <https://doi.org/10.31721/2306-5435-2020-1-107-73-83>.
30. Azarenkova, H. M., Zhuravel, T. M. and Mykhaylenko, R. M. (2009), “Enterprise finance: a study guide”, Knowledge-Press, Kyiv, 299 p., available at: http://www.library.univ.kiev.ua/ukr/elcat/new/detail.php3?doc_id=1247203.
31. Prohonov, D. O. (2018), “Theoretical and informational evaluations of container distortions during the formation of steganograms”, *Scientific and Technical Conference radioengineering fields, signals, devices and systems*. Conference Proceeding March 19-25, 2018, Kyiv, Ukraine, pp. 276-278, available at: <http://ptmip.ipt.kpi.ua/list/progonov17>.

32. Yehorshyn, O. O., Panova, N. V. and Polevych, V. V. (1955), *Regression analysis in examples and problems*, tutorial, Kharkiv State University of Economics, Kharkiv, 155 p.
33. Ulanova, E. S. and Zabelin, V. N. (1990), *Methods of correlation and regression analysis in agrometeorology*, Gidrometeoizdat, 207 p., available at: https://koha.lib.tsu.ru/cgi-bin/koha/opac-detail.pl?biblionumber=13364&shelfbrowse_itemnumber=40398.
34. Bondarenko V. N. (1970), "Statistical solutions of some problems of geology", NEDRA, Moscow, 244 p., available at: <https://www.libex.ru/detail/book482854.html>.
35. Chaddock, Robert Emmet (1925), *Principles and Methods of Statistics*, Hardcover, Houghton, Mifflin, 471 p., available at: https://books.google.com.sg/books/about/Principles_and_Methods_of_Statistics.html?id=-YxBTYcdnIoC&redir_esc=y.

СПИСОК ЛІТЕРАТУРИ

1. Мотало В. Аналіз шкал вимірювань. *Вимірювальна техніка та метрологія*. 2015. № 76. С. 21-35, URL: http://nbuv.gov.ua/UJRN/metrolog_2015_76_4
2. Maurice G. Kendall, Alan Stuart (1961). *The Advanced Theory of Statistics. Volume 2, Inference and Relationship*. Charles Griffin, London, 676 p.
3. Дубницький В. Ю., Кобылин А. М., Кобылин О. А. Оценка нижней границы надёжности физически реализуемой системы в процессе её эксплуатации при произвольных законах распределения обобщённой нагрузки и прочности. *Системи обробки інформації*. 2018. № 1(152). С. 53-60. DOI: <https://doi.org/10.30748/soi.2018.152.08>.
4. Жлуктенко В. І., Наконечний С. І., Савіна С. С. Теорія ймовірностей і математична статистика. У 2-х ч. – Ч. II. Математична статистика. Київ: КНЕУ, 2001. 336 с. URL: https://www.studmed.ru/zhluktenko-v-nakonechniy-s-savna-ss-teorya-ymovnostey-matematichna-statistika-u-2-h-ch-ch-matematichna-statistika_3976c660ed4.html.
5. Kulldorff Gunnar (1961), *Contributions to the Theory of Estimation from Grouped and Partially Grouped Samples*. Almqvist & Wiksell / John Wiley & Sons, Stockholm, 176 p. URL: <https://www.amazon.com/Contributions-Estimation-Grouped-Partially-Samples/dp/B0010VDR26>.
6. Джуль І. В. О числе градаций гистограмм ошибок астрономических наблюдений. *Кинематика и физика небесных тел*. 1993. №1, т. 9. С. 88-92. URL: <https://www.mao.kiev.ua/biblio/jscans/kfnt/1993-09/kfnt-1993-09-1-11.pdf>.
7. Джуль І. В. Математическая обработка астрономической и космической информации при негауссовых ошибках наблюдений: автореферат дис. на соиск. уч. степени докт. физ.-мат. наук: спец. 01.03.01 «Астрометрия и небесная механика». Киев, ГАО НАН Украины, 1992. 46 с. URL: <https://issuu.com/blindguardian/docs/asd>.
8. Паніотто В. І., Максименко В. С., Харченко Н. М. Статаналіз соціологічних даних. Київ: КМ Академія, 2004. 270 с.
9. Цветков О. В. Энтропийный анализ данных в физике, биологии и технике. СПб.: СПбГЭТУ ЛЭТИ, 2015. 202 с. URL: https://www.researchgate.net/profile/Oleg-Tsvetkov/publication/331686300_entropijnyj_analiz_dannyh_v_fizike_biologii_i_tehnike/links/5c87f3afa6fdcc38174f8a14/entropijnyj-analiz-dannyh-v-fizike-biologii-i-tehnikе.pdf.
10. Дубницький В. Ю., Філатова Л. Д., Ходырев А. И. Устойчивость оценки энтропии гистограммы непрерывной случайной величины по отношению к изменению количества её интервалов. *Системи управління, навігації та зв'язку*. 2017, вип. 5(45), С. 42-46. URL: http://nbuv.gov.ua/UJRN/suntz_2017_5_12.
11. Agresti A. *Categorical data analysis*. John Wiley & Sons Inc., New York, 2002, 742 p. URL: <https://onlinelibrary.wiley.com/doi/book/10.1002/0471249688>.
12. Joseph L. Fleiss, Bruce Levin, and Myunghee Cho Paik. *Statistical Methods for Rates and Proportions*. John Wiley & Sons, Inc. New York, 2003, 768 p. URL: <https://onlinelibrary.wiley.com/doi/book/10.1002/0471445428>.
13. Nobuoki Eshima. *Statistical Data Analysis and Entropy*. Springer Nature Singapore Pte Ltd, Singapore, 2020, 498 p. URL: <https://link.springer.com/book/10.1007/978-981-15-2552-0>.
14. Graham J. G. Upton (1978), *The Analysis of Cross-tabulated Data*. J. Wiley, New York, 160 p. URL: <https://www.amazon.com/Analysis-Cross-tabulated-Data-Graham-Upton/dp/0471996599>.
15. Duncan Crammer. *Advanced Quantitative Data Analysis*. Open University Press, Philadelphia, 2003. 272 p. URL: <https://www.amazon.com/Advanced-Quantitative-Analysis-Understanding-Research/dp/0335200591>.
16. Anne E. Magurran. *Ecological Diversity and its Measurement*. London, Sydney, CROOM HELM Royal Society University Research Fellow University, 1983, 184 p. URL: <https://link.springer.com/book/10.1007/978-94-015-7358-0>.
17. Margalef R. Information theory in ecology. *Gen. Syst.*, No 3, 1958. P. 36-71. URL: [https://www.scirp.org/\(S\(351jmbntvnsjt1aadkposzje\)\)/reference/ReferencesPapers.aspx?ReferenceID=1134401](https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/ReferencesPapers.aspx?ReferenceID=1134401).
18. Hutcheson K. A Test for Comparing Diversities Based on the Shannon Formula. *Journal of Theoretical Biology*. 29. 1970. P. 151- 154. URL: [http://dx.doi.org/10.1016/0022-5193\(70\)90124-4](http://dx.doi.org/10.1016/0022-5193(70)90124-4).
19. Michalowicz, J.V., Nichols, J.M. and Bucholtz, F. *Handbook of differential entropy*, T.&F Gr, LLC, London, 2014. 241 p. URL: <https://www.routledge.com/Handbook-of-Differential-Entropy/Michalowicz-Nichols-Bucholtz/p/book/9781138374799>.
20. Linfoot E., Linfoot E. H. An Informational Measure of Correlation. *Information and Control*. Vol. 1, No. 1. 1957. P. 85-89. URL: <https://www.sciencedirect.com/science/article/pii/S001999585790116X>.
21. Vistelius A. B. The skew frequency distributions and the fundamental law of Geochemical processes. *The Journal of Geology*. Vol. 68, No. 1. 1960. P. 1-22. URL: <https://www.jstor.org/stable/30058252>.
22. Marian P. and Marian T. A. (2015), [Hellinger distance as a measure of Gaussian discord], *The Journal of Physics A: Math. Theor.*, 48:11 (2015), 115301, 21 p., arXiv: 1408.4477. DOI: <http://dx.doi.org/10.1088/1751-8113/48/11/115301>.
23. Kullback S., Leibler R. A. On information and sufficiency. *The Annals of Mathematical Statistics*, Vol. 22, No 1. 1951. pp. 79-86. DOI: <http://dx.doi.org/10.1214/aoms/1177729694>.
24. Дубницький В. Ю., Скорикова І. Г., Ходырев О. І. Оптиміальна апроксимація функції щільності розподілу за критерієм мінімуму втрати інформації. *Системи обробки інформації*. 2017. Вип. 4. С. 45-51. URL: <https://www.hups.mil.gov.ua/periodic-app/article/17653>.
25. Solomon Kullback. *Information Theory and Statistics*. Peter Smith, Gloucester, Mass. 1978. 399 p. URL: https://books.google.com.ua/books/about/Information_Theory_and_Statistics.html?id=XeRQAAAAMAAJ&redir_esc=y.
26. Гадецька С. В., Дубницький В. Ю., Кушнерук Ю. І., Ходырев О. І. Спеціалізований програмний калькулятор для оцінки клінічної інформативності лабораторних тестів. *Сучасні інформаційні системи*. 2020. Т. 4, № 2. С. 80-84. DOI: <https://doi.org/10.20998/2522-9052.2020.2.12>.

27. Солошенко О. М. Дослідження відстані Кульбака-Лейблера у задачах моделювання у кредитному скорингу. *Розвиток інформаційно-ресурсного забезпечення освіти та науки в горно-металургічній галузі та на транспорті 2014*: Днепропетровск, 2014. С. 328-333. URL: <https://ir.nmu.org.ua/handle/123456789/150310>.
28. Дубницький В. Ю, Криленко І. М., Фесенко Г. В., Черепньов І. А. Історія розвитку засобів захисту очей військово-службовців в умовах бойових дій та сучасні вимоги до контролю їхньої стійкості до ударної дії. *Системи озброєння і військової техніки*. 2017, №1(49). С. 23-37. URL: <https://www.hups.mil.gov.ua/periodic-app/article/17565>.
29. Бровко Д. В. Построение системы мониторинга надежности элементов зданий и сооружений поверхностного комплекса шахт на базе оценки энтропии. *Гірничий вісник, Кривий Ріг*, 2020. Вип. 107. С. 73-83. DOI: <https://doi.org/10.31721/2306-5435-2020-1-107-73-83>.
30. Азаренкова Г. М., Журавель Т. М., Михайленко Р. М. Фінанси підприємств : навчальний посібник. Київ, Знання-Прес, 2009. 299 с. URL: http://www.library.univ.kiev.ua/ukr/elcat/new/detail.php?doc_id=1247203.
31. Прогонов Д. О. Теоретико-інформаційні оцінки спотворень контейнерів при формуванні стеганограм. *Міжнародна науково-технічна конференція «Радіотехнічні поля, сигнали, апарати та системи»*. Київ, 19-25 березня 2018 р. Київ, 2018. С. 276-278. URL: <http://ptmip.ipt.kpi.ua/list/progonov17>.
32. Сгоршин О. О., Панова Н. В., Полевич В. В. Регресійний аналіз у прикладах і задачах: навчальний посібник. Харків: ХНЕУ, 1999. 155 с.
33. Уланова Е. С., Забелин В. Н. Методы корреляционного и регрессионного анализа в агрометеорологии. Гидрометеоиздат, 1990. 207 с. URL: https://koha.lib.tsu.ru/cgi-bin/koha/opac-detail.pl?biblionumber=13364&shelfbrowse_itemnumber=40398.
34. Бондаренко В. Н. Статистические решения некоторых задач геологии. М.: НЕДРА, 1970. 244 с. URL: <https://www.libex.ru/detail/book482854.html>.
35. Chaddock Robert Emmet. *Principles and Methods of Statistics*. Hardcover, Houghton, Mifflin. 1925. 471 p. URL: https://books.google.com.sg/books/about/Principles_and_Methods_of_Statistics.html?id=-YxBTYcdnIoC&redir_esc=y.

Received (Надійшла) 22.01.2023

Accepted for publication (Прийнята до друку) 12.04.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

- Гадецька Світлана Вікторівна** – кандидат фізико-математичних наук, доцент, доцент кафедри вищої математики, Харківський національний автомобільно-дорожній університет, Харків, Україна;
Svitlana Gadetska – PhD in Physics and Mathematics, Associate Professor, Associate Professor of Department of Higher Mathematics of Kharkiv National Automobile and Highway University, Kharkiv, Ukraine;
 e-mail: svgadetska@ukr.net; ORCID ID: <https://orcid.org/0000-0002-9125-2363>.
- Дубницький Валерій Юрійович** – кандидат технічних наук, старший науковий співробітник, старший науковий співробітник Харківського навчально-наукового інституту “Каразінський банківський інститут” Харківського національного університету ім. В. Н. Каразіна, Харків, Україна;
Valeriy Dubnitskiy – PhD in Engineering Senior Researcher Senior Researcher of “Karazin Banking Institute” of V.N. Karazin Kharkiv National University, Kharkiv, Ukraine;
 e-mail: dubnitskiy@gmail.com; ORCID ID: <https://orcid.org/0000-0003-1924-4104>.
- Кушнерук Юрій Іонович** – кандидат технічних наук, доцент, доцент Інституту цивільної авіації Харківського національного університету Повітряних Сил імені Івана Кожедуба, Харків, Україна;
Yuri Kushneruk – Candidate of Technical Sciences, Associate Professor, Associate professor of Civil Aviation Institute of Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine;
 e-mail: kyshneryk_ui@ukr.net; ORCID ID: <https://orcid.org/0000-0001-5844-7137>.
- Ходирєв Олександр Іванович** – старший викладач Харківського навчально-наукового інституту “Каразінський банківський інститут” Харківського національного університету ім. В. Н. Каразіна, Харків, Україна;
Alexander Khodyrev – Senior Lecturer of “Karazin Banking Institute” of V.N. Karazin Kharkiv National University, Kharkiv, Ukraine;
 e-mail: khodyrevmjk3758@gmail.com; ORCID ID: <https://orcid.org/0000-0001-9871-9440>.

Excel-oriented calculator for calculating results of entropy analysis of data distributed by categories

Svitlana Gadetska, Valeriy Dubnitskiy, Yuri Kushneruk, Alexander Khodyrev

Abstract. The goal of the work. Development of EXCEL-oriented calculator for calculating the results of entropy analysis of data, which are distributed by categories. The subject of research is histograms of arbitrary distribution laws and conjugation tables 2×2 . **Research methods:** Entropy and information analysis of histograms of arbitrary distribution laws and conjugation tables. **The obtained results.** It is proposed to use methods of entropy analysis for the analysis of data distributed by categories; information on the structure of the EXCEL-oriented calculator designed for this purpose is given. The calculator makes it possible to calculate entropy characteristics of histograms, namely: histogram entropy, histogram dispersion, histogram confidence intervals, diversity information index. The calculator performs a pairwise comparison of entropies of histograms using the Hutcheson method, determines Hellinger and Kullback-Leibler distances between histograms of arbitrary distribution laws and thus complements the chi-square criterion, determines the informational correlation coefficient. The correspondence between the Pearson correlation coefficient and the information correlation coefficient is established by the method of statistical modeling. For 2×2 conjugation tables, the calculator makes it possible to estimate the significance of the interaction between the row factor and the column factor. The calculator determines the values of conditional entropies for 2×2 conjugation tables. The proposed calculator fills the gaps in existing software products and can be used to process data distributed by categories using entropy analysis methods. It is shown that entropy methods of analysis are appropriate to use in cases where histograms determine arbitrary distribution laws.

Keywords: entropy; entropy analysis; information correlation coefficient; Hellinger distance; Kullback-Leibler distance.

Serhii Koshman¹, Victor Krasnobayev¹, Serhii Nikolsky², Dmytro Kovalchuk¹

¹ V. N. Karazin Kharkiv National University, Kharkiv, Ukraine

² Kharkiv humanitarian-pedagogical academy, Kharkiv, Ukraine

THE STRUCTURE OF THE COMPUTER SYSTEM IN THE RESIDUAL CLASSES

Abstract. The **subject** of the article is the formulation and solution of the inverse problem of optimal redundancy in the system of residual classes (RNS) based on the use of the dynamic programming method. The solution of this problem makes it possible to improve the reliability of the operation of computer systems and components (CSC) in the RNS. The **purpose** of the article is to increase the reliability of the functioning of CSC, which are built on the basis of the use of RNS, without reducing the speed of calculations, as well as to calculate and compare the reliability, in terms of the probability of failure-free operation, of CSC in RNS and a tripled computing system that operates in a positional binary number system (PNS). **Tasks:** to analyze the influence of the number system used on the reliability of the CSC, taking into account the primary and secondary redundancy; to synthesize a computing system in RNS for a l -byte bit grid based on the use of the passive fault tolerance method (constant structural redundancy); formulate and solve the inverse problem of optimal redundancy in RNS based on the use of the dynamic programming method; to check the correctness of the results obtained, calculate the conditional amount of computer system equipment in the residual classes; evaluate the efficiency of using RNS to improve reliability when building a redundant CSC in relation to a redundant CSC in the PNS. **Research methods:** methods of analysis and synthesis of computer systems, number theory, coding theory in RNS, reliability theory. The following **results** are obtained. The paper shows that the use of PNS as a number system does not allow a radical increase in the performance and reliability of CSC. In this regard, the article developed the concept of using RNS as a number system for constructing a CSC. Based on this, the inverse problem of optimal redundancy in RNS is formulated and solved. **Conclusions.** As shown by the results of calculations and comparative analysis, the use of RNS provides a higher reliability of the CSC than the majority three-channel computing system in the PNS. The obtained research results can be used for the synthesis of fault-tolerant computer structures in RNS.

Keywords: non-positional code structure; system in residue classes; positional binary number system; reliability of computer systems and components.

Introduction

Modern trends in the development of computer technology cause its widespread introduction into various spheres of human activity. At the same time, large volumes of tasks to be solved require an increase in productivity and ensuring a given level of reliability of computing systems (CS). In addition, the content and complexity of such requests outstrip the pace of increasing power of existing computer systems and components (CSC) of general and special purpose, functioning in positional binary number system (PNS). In this aspect, the main directions of improving computer systems and components in positional number system are increasing user productivity and reliability (primarily reliability) of their functioning [1-3]. Ensuring the presence of a fault tolerance property in a CS can increase the reliability of its operation.

The fault tolerance property provides an ability to perform specified computational functions after failures, both by reducing, within acceptable limits, any indicators of the functioning quality (for example, by gradual degradation), and without deteriorating the functioning quality of computer system. Thus, considering the above, research in the field of developing methods for improving fault tolerance in the process of functioning of computer systems and components are relevant.

Reserves for increasing the speed of computing in positional number system are the use of computer system and components, created on the principle of problem (algorithm) parallelization at the level of microoperations. The concept of parallelism has long

attracted the attention of specialists with its potential to increase the performance of computing systems. The theoretical, experimental and industrial developments in this direction have made it possible to substantiate the basic principles for constructing parallel computing systems. The prospect of further increasing a computing power of devices is currently associated with such systems [5-7].

The main methods that are widely used in the construction of fault tolerant computing devices and systems in positional number system are structural redundancy. There are a large number of different backup methods, but some of them is characterized by significant structural redundancy. Even with the correction of single errors, most often it is necessary to increase the volume of the computer system equipment at least three times. Such a high structural redundancy is explained by the fact that when applying redundancy, all the specific properties of specific types of computing systems and components are almost completely ignored [8-10].

In the modern literature, it is shown, that, firstly, the use of positional binary number system as the number system does not allow a radical increasing in the performance and fault tolerance of computer systems and components. Secondly, there are results of basic research and specific technical developments that show the possibility of significantly increasing the speed of implementing integer arithmetic operations of addition, multiplication and subtraction by using a non-positional number system in residue classes (RNS). This is achieved through the use of the following properties of residue system: independence, equality and low bit

depth of the residues that determine a non-position code structure (NCS), which allows the following: parallelize arithmetic calculations at the level of decomposition of the remainders of numbers; realize spatial diversity of data elements with the possibility of their subsequent asynchronous independent processing; perform tabular execution of arithmetic operations of the basic set and polynomial functions with a single-cycle sample of the result of modular operation. The above significantly improves the computer system and components performance.

However, the lack of the results of basic research on the use of systems in a residue classes to increase fault tolerance hinders the solution of the problem of a significant increase in the reliability of the operation of the computer systems and components.

The purpose of the article is to increase the reliability of the functioning of CSC, which are built on the basis of the use of non-positional number system in residue classes, without reducing the speed of calculations, as well as to calculate and compare the reliability, in terms of the probability of failure-free operation, of CSC in RNS and a tripled computing system, which operates in a positional binary number system.

Problem statement

Suppose that at the design stage it is necessary to provide the necessary (predetermined) level of reliability of the computing system. It is possible to increase (ensure) reliability if the computer system will have a certain property, the use of which will allow it to be done. Such a property is defined and called fault tolerance [11-13]. With respect to the computer system, the concept of fault tolerance can be understood as the property of the computer system to ensure its operational state in case of failures of the elements included in their composition.

In the definition of the term fault tolerance there are three main aspects of its use: the fault tolerance property is laid down by the developers during the design of the computer system in order to increase its reliability; at the same time, the necessary level of fault tolerance is achieved mainly when using redundant (additional) technical means (introducing artificial structural and (or) other redundancy) in comparison with the necessary minimum to perform all the required functions of the computer system and components in full; the use of fault-tolerance properties allows to save the full or partial performance of the computer system; it is believed that the failure of the elements of the computer system is not associated with exposure not provided for by the operating conditions.

In the most cases, developers are interested in the fact of ensuring fault tolerance only while maintaining full operability, i.e. without reducing the quality of the computer system functioning. In the future, when considering the concept of fault tolerance, we will be interested only in such option for the operation of computer systems and components.

To provide the computer system with the fault tolerance property, at the design stage, it is necessary to provide not only an introduction and use of artificial

redundancy (AR), i.e. use of various types of redundancy: structural, informational, functional, temporary and load, but also to identify and use the possible natural ("natural" available redundancy) redundancy (NR). In this regard, the main designer's task to ensure the necessary level of fault tolerant operation is to identify (determine) and use the existing internal reserves (IR) of computer systems and components for fault tolerance at the pre-design stage, due to the number system used and, with this in mind, in the future, select and apply the necessary reservation methods (introduction of IR). Accounting and use of NR will increase the reliability of computer systems and components.

In [14-16], as applied to computing devices, particular definitions of primary and secondary redundancy were introduced. In this aspect, it is believed that the primary redundancy is due to used number system in the computer systems and components. Obviously, secondary redundancy is redundancy due to the application of traditional backup methods widely used in various information systems to improve their individual characteristics. Primary redundancy for computer system coincides with the concept of natural redundancy of information processing systems, and secondary redundancy - with the concept of artificial redundancy. The need for the addition and uses of secondary redundancy is due to the requirements for the characteristics at the design stage of the computer system. Note that the selected and used number system significantly affects the following characteristics: structure (architecture); principles of information processing (to a greater extent on methods and algorithms for performing arithmetic operations); requirements for the use of the new element base; system and user performance; reliability, survivability, fault tolerance, operational characteristics and indicators of computer system, etc.

Quantitatively, a volume V_{PR} of computer systems and components equipment due to the presence of primary redundancy (the presence of redundancy only due to the used number system) is slightly less than a volume of equipment V_{NR} in the presence of natural redundancy (redundancy due not only to the used number system). The volume of additional equipment V_{SR} determined by the presence of secondary redundancy fully coincides with the volume of equipment V_{AR} due to the presence of artificial redundancy. An analysis of the number system influence on the structure and individual characteristics of computer system showed that it is completely correct to assume that $V_{PR} \approx V_{NR}$. In the traditional approach to the choice of the positional number system base, it is first necessary to ensure the following condition:

$$V_{PR} = \min . \quad (1)$$

However, the fulfillment of the condition (1) is not always valid when developing computational structures, when a priori the problem arises of improving some characteristics of the computer system. It is possible that

the option of constructing a computer system based on the fulfillment of condition (1), when solving the problem of increasing reliability, is not at all advisable. This feature is clearly manifested when used as number system, for example, residue number system. It is known [17-19] that redundant computer system with residue number system contains (15-20)% more equipment V_{PR} than a computer system in a positional binary number system with the same given bit grid without taking into account the addition of secondary redundancy. As preliminary studies have shown, in order to achieve a given level of fault tolerance of the computer system in residue classes, 50% less volume of equipment is required than for system in a positional binary number system. However, the lack of practical results of the synthesis of fault-tolerance computer system in residue classes does not make it possible to show the efficiency of using the non-position code structure to increase the reliability of the computer system functioning without reducing the performance of solving problems.

Statement and solution of the inverse problem of optimal redundancy

For the purpose of calculating and comparative analysis of the operational reliability of the computer system in positional number system and in residue classes, we will conduct the synthesis of the fault tolerant computer system with residue number system. Let it be necessary to synthesize a computer system in residue number system for a l -byte bit grid. Obviously, the results of solving the synthesis problem of computer system in residue classes will substantially depend on the type of structural reservation used: constant or dynamic.

Therefore, it is advisable to separately solve the synthesis problem in the case of constant or dynamic structural redundancy.

This article solves the problem of synthesizing the structure of the CS, based on using the method of passive fault tolerance (permanent structural redundancy).

In order to solve the problem of synthesis of computer system in residue classes in the case of constant structural redundancy with a loaded reserve without restoring a failed element, we introduce a concept of the state vector

$$X^{(n)}_{RNS} = (x_1, x_2, \dots, x_i, \dots, x_n)$$

of a redundant computer system in residue classes.

In this case, the role of the elements of the redundant system is played by computing paths (CP) for each module $m_i (i = \overline{1, n})$ of the residue number system, and the values $x_i = 0, 1, 2, \dots$ indicate the multiplicity of the reservation of a separate CP of computer system in residue classes by the corresponding module (when the value of the main CP equals to $x_i = 0$ by the module m_i , there are no redundant computing paths). Verbally, the task of synthesizing computer system in residue

classes is formulated as follows: from the whole set $X^{(n)}_{RNS}$ of possible values of the state vector, it is necessary to determine the only reserve composition vector at which the reliability of the computer system in residue classes $P^{(n)}_{RNS}[X^{(n)}_{RNS}]$ would reach the maximum possible value. Obviously, the solution to this synthesis problem is directly related to the formulation and solution of the inverse optimal reservation problem in residue classes.

The inverse problem of optimal reservation in residue classes is formulated as follows: it is necessary to determine a vector

$$X^{(n)}_{RNS} = (x_1, x_2, \dots, x_i, \dots, x_n)$$

of reserve composition for which $V^{(l)}_{add}$, at acceptable costs, the maximum probability of fault tolerant operation $P^{(n)}_{RNS}[X^{(n)}_{RNS}]$ would be achieved:

$$\begin{cases} P^{(n)}_{RNS} [X^{(n)}_{RNS} = X^{(n)}_{RNS} (x_1, x_2, \dots, x_i, \dots)] = \\ = P^{(n)}_{RNS} \left[\prod_{i=1}^n P_{x_i}(t) \right] \rightarrow \max, \\ \left(V^{(n)}_{RNS} \leq V^{(l)}_{add} \right), \end{cases} \quad (2)$$

Mathematically, this problem can be represented as follows where:

the x_i is an its component of the vector $X^{(n)}_{RNS}$ of the redundant computing system with modulo m_i of residue classes, which numerically characterizes a number of redundant computing paths connected to the main (working) computing path for this module (base) of residue classes;

n is a number of bases m_i in residue number system; $P_{x_i}(t) = 1 - (1 - e^{-\alpha_i \cdot \lambda_{FR} \cdot t})^{x_i + 1}$ is a probability of fault tolerant operation of the redundant system with modulo m_i in residue number system;

λ_{FR} is a failure rate of a conventional unit of equipment of the computer system, assigned to one binary digit of the bit grid of the computer system;

$\alpha_i = \lceil \log_2(m_i - 1) \rceil + 1$ is a number of binary digits needed to represent a module (the relative "cost" of one computing path in absolute value, expressed in binary digits; a value $V^{(n)}_0$ is a set limit on the cost of the system when solving the inverse optimal reservation problem in residue number system) [20, 21].

As the computer system in positional number system, used for comparison with the computer system in residue classes, we take the majority three-channel computing system, consisting of three same type l -discharge computing systems. This is the most widely used at present to increase the reliability of the computer system. In this case, without considering the reliability of the majority part, the amount of equipment is equal to $V_{add} = 3 \cdot 8 \cdot l$ conventional units. We note

that the probability of fault tolerant operation of a three-channel computing system in positional number system (without taking into account the reliability of the majority part) is equal to

$$P^{(l)}_{PNS}(t) = 1 - [1 - P^{(l)}_0(t)]^3,$$

where $P^{(l)}_0(t) = e^{-8 \cdot l \cdot \lambda_{FR} \cdot t}$ is a probability of the fault tolerant operation of a l -byte computer system in residue classes.

Note that $V^{(n)}_0 = V^{(l)}_{add} - \sum_{i=1}^n \alpha_i$ is a difference

between the permissible costs in positional number system and the costs necessary to build a fault tolerant computer system in residue classes (the $V^{(n)}_0$ values of the permissible restrictions on the creation of a redundant computer system in residue classes).

To solve the inverse problem (2) of optimal reservation formulated in the article in residue number system, a dynamic programming method is recommended in the literature [6]. The approach using this method is very flexible for solving problems associated with multi-stage selection. In addition, the dynamic programming method due to the fact that the solutions are recurrence relations is very convenient for performing numerical calculations on a computer. To solve the inverse optimal reservation problem, when using dynamic programming, it is necessary to leave the main functional equation in the form:

$$\left\{ \begin{array}{l} \max P^{(n)}_{RNS} [x_1, x_2, \dots, x_i, \dots, x_n] = \max \prod_{i=1}^n P_{x_i}(t), \\ \left(0 \leq \sum_{i=0}^n a_i \cdot x_i \leq V^{(n)}_0, \quad x_n = 0, 1, 2, \dots \right). \end{array} \right. \quad (3)$$

We introduce into consideration some function $F_n(V^{(n)}_0)$, index n at which means the dimension of the maximized function $\prod_{i=1}^n P_{x_i}(t)$, and its argument is a permissible restriction imposed on the arguments of the function $V^{(n)}_0$. In this case, functional equation (3) can be written as

$$\left\{ \begin{array}{l} F_n(V^{(n)}_0) = \max P_{x_n} \cdot F_{n-1}(V^{(n)}_0 - a_n \cdot x_n), \\ \left(0 \leq a_n \cdot x_n \leq V^{(n)}_0, \quad x_n = 0, 1, 2, \dots \right). \end{array} \right. \quad (4)$$

In view of the foregoing, a functional equation giving a recursive solution for the inverse optimal reservation problem in residue classes will be presented as follows

$$\left\{ \begin{array}{l} F_n \left[3 \cdot 8 \cdot l - \sum_{i=1}^n a_i \right] = \\ = \max P_{x_n} \cdot F_{n-1} \left[V^{(n)}_0 - a_n \cdot x_n \right], \\ \left(0 \leq a_n \cdot x_n \leq \left(3 \cdot 8 \cdot l - \sum_{i=1}^n a_i \right), \quad x_n = 0, 1, 2, \dots \right). \end{array} \right. \quad (5)$$

The algorithm for solving the inverse optimal reservation problem in residue number system is as follows.

The optimal two-dimensional vectors of the reserve composition are determined for the first and second computing path of computer system in residue classes, corresponding to the modules m_1 and m_2 , for all values of the cost indicator, not exceeding the value $V^{(n)}_0$.

The optimal three-dimensional reserve composition vectors for the third computing path are determined by modulo m_3 and the corresponding vectors (x_1, x_2) for all values of the cost indicator not exceeding the value $V^{(n)}_0$. A similar process continues until the optimal $(n-1)$ -measured vector

$$X^{(n)}_{RNS} = (x_1, x_2, \dots, x_i, \dots, x_{n-1})$$

of the reserve composition and the corresponding optimal vector for the value of the conditional cost indicator equal to $V^{(n)}_0$ are found.

An optimal value x_n is determined, which, together with the value of the vector $(x_1, x_2, \dots, x_i, \dots, x_{n-1})$, gives the optimal solution to the problem.

Let's consider an example of solving the inverse optimal reservation problem in residue classes for a single-byte ($l=1$) bit network of computer system. Redundant computing paths of computer system in residue classes, in relation to the main (working) computing path, are in the load reserve, and failed computing paths are not restored. For the case when $l=1$ ($n=4$), the RNS consists of four modules (bases). Table 1 presents a data for solving the inverse problem of optimal reservation in residue number system of various l values of the bit ($l=1, 4, 8$) grids in the computer system.

For the value $l=1$, the inverse optimal reservation problem in residue classes is formulated as follows: it is necessary to determine a reserve composition vector

$$X^{(n)}_{RNS} = X^{(4)}_{RNS} = (x_1, x_2, x_3, x_4)$$

for which at acceptable costs

$$V^{(l)}_{add} = 3 \cdot 8 \cdot l = 3 \cdot 8 \cdot 1 = 24$$

of conventional units, and the fault tolerant operation (probability of fault-tolerance operation) $P^{(4)}_{RNS}[X^{(4)}_{RNS}]$ of the computer system in residue classes would reach the maximum possible value. Note that the cost of one element of the i -type of the redundant system (the conditional volume V_{m_i} of computing paths equipment operating by modulo m_i) is determined by the number of binary digits $\alpha_i = [\log_2(m_i - 1)] + 1$ required to represent the number m_i .

Table 1 – Initial data for solving the inverse optimal reservation problem in residue number system ($l = \overline{1,4,8}$)

The value of the l -bit grid of computer system	The number of bases of residue number system	Set of bases $\{m_i\}$ of residue number system $i = \overline{1, n}$	$V^{(l)}_{add}$	$\sum_{i=1}^n a_i$	$V^{(n)}_0$
1	4	$m_1 = 3, m_2 = 4, m_3 = 5, m_4 = 7$	24	10	14
2	6	$m_1 = 3, m_2 = 4, m_3 = 5, m_4 = 7, m_5 = 11, m_6 = 13,$	48	19	29
3	8	$m_1 = 3, m_2 = 4, m_3 = 5, m_4 = 7, m_5 = 11, m_6 = 13,$ $m_7 = 17, m_8 = 19$	72	28	44
4	10	$m_1 = 3, m_2 = 4, m_3 = 5, m_4 = 7, m_5 = 11, m_6 = 13,$ $m_7 = 17, m_8 = 19, m_9 = 23, m_{10} = 29$	96	37	59
8	16	$m_1 = 3, m_2 = 4, m_3 = 5, m_4 = 7, m_5 = 11, m_6 = 13,$ $m_7 = 17, m_8 = 19, m_9 = 23, m_{10} = 29, m_{11} = 31,$ $m_{12} = 37, m_{13} = 41, m_{14} = 43, m_{15} = 47, m_{16} = 53$	192	72	120

The indicator of the necessary conditional costs $V^{(n)}_0$ (specified restrictions on the cost of the reserved system) is defined as the difference between the allowable costs $V^{(l)}_{add}$ in positional number system and the costs $\sum_{i=1}^n a_i$ necessary to build breakeven computer system in residue classes, i.e.

$$V^{(n)}_0 = V^{(l)}_{add} - \sum_{i=1}^n \alpha_i.$$

For the value $l=1$, we can write (Table 1) the following:

$$V^{(n)}_0 = 14 = 24 - 10,$$

where $V^{(l)}_{add} = 3 \cdot 8 \cdot 1 = 24$, $\sum_{i=1}^n a_i = 2 + 2 + 3 + 3 = 10$.

In order to obtain a solution to the inverse problem of optimal redundancy in residue number system while $l=1$ and conducting a comparative analysis of the reliability of the computer system in positional number system and in residue classes, we give an example of the calculation of reliability (expression (5)).

As initial data, we take, for example, that the value λ_{FR} is equal $\lambda_{FR} = 0,1[1/\text{hour}]$, and the given operating time of the computer system, assigned to one binary digit of the bit grid of the computer system, is equal to $t_N = 0,1$ an hour. In this case, we obtain that $P_E(t_N) = e^{-\lambda_{FR} \cdot t_N} = e^{-0,01} = 0,99$ is a probability of fault tolerant operation of the equipment, assigned to one binary discharge of the bit grid of the computer system. The probability of failure $P_i(t_N)(i = \overline{1,4})$ of one computing path by modulo m_i equal to:

$$P_1(t_N) = e^{-a_1 \cdot \lambda_{FR} \cdot t_N} = e^{-2 \cdot 0,01} \approx 0,98;$$

$$P_2(t_N) = e^{-a_2 \cdot \lambda_{FR} \cdot t_N} = e^{-2 \cdot 0,01} \approx 0,98;$$

$$P_3(t_N) = e^{-a_3 \cdot \lambda_{FR} \cdot t_N} = e^{-3 \cdot 0,01} \approx 0,97;$$

$$P_4(t_N) = e^{-a_4 \cdot \lambda_{FR} \cdot t_N} = e^{-2 \cdot 0,01} \approx 0,98.$$

Let's pre-calculate the values of the unreliability indicator, i.e. we calculate the values of the failure probability $q_{x_i} = 1 - P_{x_i}$ of each of the four subsystems of the computer system (computing path) in residue number system by modulo m_i for the number of backup elements x_i not exceeding five. For a large value of the values of the failure probability $q_{x_i} = 1 - P_{x_i}$ exceeding five, it is impractical to calculate q_{x_i} , since they will not be used. The results of the calculation of the values q_{x_i} are presented in Table 2.

Table 2 – Initial data for solving the optimal reservation problem for the value $l=1$

x_i	q_{x_1}	q_{x_2}	q_{x_3}	q_{x_4}
0	$2 \cdot 10^{-2}$	$2 \cdot 10^{-2}$	$3 \cdot 10^{-2}$	$3 \cdot 10^{-2}$
1	$4 \cdot 10^{-4}$	$4 \cdot 10^{-4}$	$9 \cdot 10^{-4}$	$9 \cdot 10^{-4}$
2	$8 \cdot 10^{-6}$	$8 \cdot 10^{-6}$	$27 \cdot 10^{-6}$	$27 \cdot 10^{-6}$
3	$16 \cdot 10^{-8}$	$16 \cdot 10^{-8}$	$81 \cdot 10^{-8}$	$81 \cdot 10^{-8}$
4	$32 \cdot 10^{-10}$	$32 \cdot 10^{-10}$	$243 \cdot 10^{-10}$	$243 \cdot 10^{-10}$
5	$64 \cdot 10^{-12}$	$64 \cdot 10^{-12}$	$729 \cdot 10^{-12}$	$729 \cdot 10^{-12}$

In the further calculations of reliability, we use an approximate expression of the form:

$$P_{x_i} \cdot P_{x_{n-1}} \approx 1 - q_{x_n} - q_{x_{n-1}}.$$

In accordance with the above algorithm for solving the inverse optimal reservation problem in residue classes and based on the initial data shown in Tables 1 and 2, for the value $l=1$ we obtain the desired optimal vector $X^{(4)}_{RNS} = (1,1,1,2)$, the value of the i coordinate of which ($i = \overline{1,4}$) is equal to the number of backup computing paths connected to the operating computing path for this base m_i of the residue classes (Table 3).

Table 3 also presents the results of solving the inverse optimal reservation problem in residue classes for the values l_j while $j = 2, 3, 4, 8$. In order to verify the correctness of the results of solving the inverse

optimal reservation problem in residue classes, the article presents calculated values $V_{calc}^{(n)}_0$ of the conditional amount of the equipment of the computer system in residue classes, presented in Tables 4-8.

Table 3 – The result of solving the optimal reservation problem $j = 1, 2, 3, 4, 8$

$l(n)$	$X^{(n)}_{OPT} = (x_1, x_2, \dots, x_i, \dots, x_n)$	$P_{RNS}^{(n)}(t_H)$	$P_{PNS}^{(l)}(t_H)$	$K_H(t_H)$
1 (4)	(1, 1, 1, 2)	0,9983	0,9995	–
2 (6)	(1, 2, 2, 1, 1, 2)	0,9966	0,9963	1,0983
3 (8)	(1, 2, 1, 1, 1, 2, 2, 2)	0,9959	0,9902	2,3809
4 (10)	(1, 2, 1, 1, 1, 1, 2, 2, 2, 2)	0,9944	0,9787	3,7973
8 (16)	(2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1)	0,9800	0,9101	4,5101

Table 4 – Data to verify the solution of the inverse optimal reservation problem in residue classes for $l = 1$

m_i	3	4	5	7	$V_{calc}^{(4)}_0$
α_i	2	2	3	3	10
x_i	1	1	1	2	–
$\alpha_i \cdot x_i$	2	2	3	6	13

Table 5 – Data to verify the solution of the inverse optimal reservation problem in residue classes for $l = 2$

m_i	2	5	7	9	11	13	$V_{calc}^{(6)}_0$
α_i	1	3	3	4	4	4	19
x_i	1	2	2	1	1	2	–
$\alpha_i \cdot x_i$	1	6	6	4	4	8	29

Table 6 – Data to verify the solution of the inverse optimal reservation problem in the residue classes for $l = 3$

m_i	3	4	5	7	11	13	17	19	$V_{calc}^{(8)}_0$
α_i	2	2	3	3	4	4	5	5	28
x_i	1	2	1	1	1	2	2	2	–
$\alpha_i \cdot x_i$	2	4	3	3	4	8	10	10	44

Table 7 – Data to verify the solution of the inverse optimal reservation problem in residue classes for $l = 4$

m_i	2	3	5	7	11	13	17	19	23	29	$V_{calc}^{(10)}_0$
α_i	1	2	3	3	4	4	5	5	5	5	37
x_i	1	2	1	1	1	1	2	2	2	2	–
$\alpha_i \cdot x_i$	1	4	3	3	4	4	10	10	10	10	59

Table 8 – Data to verify the solution of the inverse optimal reservation problem in residue classes for $l = 8$

m_i	2	3	5	7	11	13	17	19	23	29	31	37	41	43	47	53	$V_{calc}^{(16)}_0$
α_i	1	2	3	3	4	4	5	5	5	5	5	6	6	6	6	6	72
x_i	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	–
$\alpha_i \cdot x_i$	2	4	6	6	8	8	10	10	10	10	10	12	6	6	6	6	120

In accordance with the necessary condition $V_{calc}^{(n)}_0 \leq V^{(n)}_0$ of the inverse optimal reservation problem in residue classes, a comparative analysis of the initial values $V^{(n)}_0$ (Table 1) and the calculated values $V_{calc}^{(n)}_0$ (Table 4-8) was performed.

The results of the comparative analysis showed the correctness of the obtained results of solving the inverse optimal reservation problem in residue classes.

We evaluate an effectiveness of the computer system in residue classes and in positional number system as a ratio of the reliability of two redundant

computing systems. In reliability theory, there is a criterion for evaluating the effectiveness of redundancy.

This criterion is a coefficient $K_N(t_N)$ of reliability increase, and it is defined as the ratio of the failure probabilities of two redundant computer systems at a given operating time t_N , i.e.

$$K_N(t_N) = \frac{1 - P^{(l)}_{PNS}(t_N)}{1 - P^{(n)}_{RNS}(t_N)}.$$

The coefficient $K_N(t_N)$ characterizes the decrease in the failure probability of the computer system in residue classes compared to the computer system in positional number systems.

The results of the calculation of values $K_N(t_N)$ are summarized in the table 3, which contains the results of solving the inverse optimal reservation problem in residue classes for l – byte ($l = 1, 4, 8$) bit grids.

The results of solving this problem showed that the use of residue number system for $l \geq 2$ provides a higher value of the probability of fault tolerant operation $P^{(l)}_{RNS}(t)$ than the method of tripling of the computer system widely used in positional binary number system.

Note that with an increase in the value of the computer system l – bit grid, the efficiency of using residue classes increases.

Conclusions

A new concept is proposed to increase the reliability of the computer system by using the available redundancy of the number system. The concept assumes that in the process of designing computer systems and components, accounting and possibility of using natural redundancy (account of used number system) and artificial redundancy (reservation methods) are made. The basis of these methods is PFT and AFT, which are based on the joint use of natural and artificial redundancy. This fact allows to set and solve the problem of achieving the required level of reliability at the design stage of the computer system for any applicable number system. When implementing PFT or AFT, the essence of which is to identify (determine) a natural redundancy of the computer system through the use of the applicable number system. With the combined use of natural and artificial redundancy, on the basis of well-known methods for increasing reliability, the maximum value of the operational reliability of the computer system due to the total redundancy can be achieved. Note that in residue number system, primary structural redundancy is significantly manifested only in the presence of secondary structural redundancy. As an example of the use of the proposed concept, the computer system is considered in residue classes. For this, the inverse problem of optimal redundancy in residue classes for l -byte bit grids is formulated and solved.

REFERENCES

1. Kamaraj, A. and Marichamy, P. (2019), "Design of integrated reversible fault-tolerant arithmetic and logic unit", *Microprocess Microsyst* 69, pp. 16-23, doi: <https://doi.org/10.1016/j.micpro.2019.05.009>.
2. Jain, S. and Gajjar, S. (2023), "Design and Implementation of Fault Tolerance and Diagnosis Technique for Arithmetic Logic Unit (ALU) in Soft-Core Processor", Darji, A.D., Joshi, D., Joshi, A., Sheriff, R. (eds) *Advances in VLSI and Embedded Systems. Lecture Notes in Electrical Engineering*, vol. 962, Springer, doi: https://doi.org/10.1007/978-981-19-6780-1_10.
3. Krasnobaev, V., Kuznetsov, A., Kiian, A. and Kuznetsova, K. (2021), "Fault Tolerance Computer System Structures Functioning in Residue Classes", *11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, pp. 471-474, doi: [10.1109/IDAACS53288.2021.9660919](https://doi.org/10.1109/IDAACS53288.2021.9660919).
4. Ankit, K.V., Narasimham, S.M. and Prakash, V. (2015), "Design approach for fault recoverable ALU with improved fault tolerance", *Int J VLSI Des Commun Syst* 6(4), doi: <https://doi.org/10.5121/vlsic.2015.6402>.
5. Vallero, A. (2016), "Cross-layer system reliability assessment framework for hardware faults", *2016 IEEE International Test Conference (ITC)*, Fort Worth, TX, USA, pp. 1-10, doi: [10.1109/TEST.2016.7805863](https://doi.org/10.1109/TEST.2016.7805863).
6. Krasnobaev, V., Kuznetsov, A., Popenko, V. and Kuznetsova, T. (2021), "Mathematical Model of the Reliability of a Computer System which is Functioning in the Residual Class System, Taking into Account the Reliability of Switching Devices", *IEEE 4th International Conference on Advanced Information and Communication Technologies (AICT)*, pp. 225-229, doi: [10.1109/AICT52120.2021.9628929](https://doi.org/10.1109/AICT52120.2021.9628929).
7. Egwuotuoha, I P, Levy, D, Selic, B. and Chen, S. (2013), "A survey of fault tolerance mechanisms and checkpoint/restart implementations for high performance computing systems", *The Journal of Supercomputing*, 65(3), pp. 1302–1326, available at: <https://link.springer.com/article/10.1007/s11227-013-0884-0>.
8. Krasnobayev, V. A., Koshman, S. A. and Mavrina, M. A. (2014), "A method for increasing the reliability of verification of data represented in a residue number system", *Cybernetics and Systems Analysis*, vol. 50, Issue 6, pp. 969-976, available at: <https://link.springer.com/article/10.1007/s10559-014-9688-3>.
9. Bravo-Montes, J.A., Martín-Toledano, A. and Sánchez-Macián, A. (2022), "Design and implementation of efficient QCA full-adders using fault-tolerant majority gates", *The Journal of Supercomputing* 78, pp. 8056–8080, doi: <https://doi.org/10.1007/s11227-021-04247-9>.
10. Moroz, S. A. and Krasnobayev, V. A. (2011), "A data verification method in a non-positional residue number system", *Control, Navigation, and Communication Systems* No. 2(18), pp. 134–138.
11. Ahmadpour, S.-S., Mosleh, M. and Heikalabad, S.R. (2020), "An efficient fault-tolerant arithmetic logic unit using a novel fault-tolerant 5-input majority gate in quantum-dot cellular automata", *Computers & Electrical Engineering*, Vol. 82, 106548, doi: <https://doi.org/10.1016/j.compeleceng.2020.106548>.
12. Krasnobayev, V. A., Kuznetsov, A. A., Koshman, S. A. and Kuznetsova, K. O. (2020), "A method for implementing the operation of modulo addition of the residues of two numbers in the residue number system", *Cybernetics and Systems Analysis*, Vol. 56, No. 6, pp. 1029-1038, doi: <https://doi.org/10.1007/s10559-020-00323-9>.
13. Valvano, J. (2017), "Embedded Systems: Real-Time Operating Systems for Arm Cortex M Microcontrollers", *CreateSpace Independent Publishing Platform*, 486 p., available at: <https://faculty.uobasrah.edu.iq/uploads/1631127575.pdf>.

14. Shahid, M.A. (2021), "Towards resilient method: an exhaustive survey of fault tolerance methods in the cloud computing environment", *Computer Science Review*, Vol. 40, 100398, doi: <https://doi.org/10.1016/j.cosrev.2021.100398>.
15. Zhang, Y., Mandal, A., Koelbel, C. and Cooper, K. (2009), "Combined fault tolerance and scheduling techniques for workflow applications on computational grids", *Proceedings of the 9th IEEE/ACM International Symposium on Cluster Computing and the Grid*. CCGRID'09, pp 244–251, doi: <https://doi.org/10.1109/CCGRID.2009.59>.
16. Pascucci, F. (2023), "Fault Tolerance in Cyber-Physical Systems", Jajodia, S., Samarati, P., Yung, M. (eds) *Encyclopedia of Cryptography, Security and Privacy*. Springer, Berlin, Heidelberg, doi: https://doi.org/10.1007/978-3-642-27739-9_1724-2.
17. José Rodrigo, Azambuja, Fernanda, Kastensmidt and Jürgen, Becker, (2014), *Hybrid Fault Tolerance Techniques to Detect Transient Faults in Embedded Processors*. Springer Cham, doi: <https://doi.org/10.1007/978-3-319-06340-9>.
18. Ozturk, Z., Topcuoglu, H.R. & Kandemir, M.T. (2022), "Studying error propagation on application data structure and hardware", *The Journal of Supercomputing*, 78, pp. 18691–18724, doi: <https://doi.org/10.1007/s11227-022-04625-x>.
19. Jia, J., Liu, Y. and Zhang, G. (2023), "Software approaches for resilience of high performance computing systems: a survey", *Frontiers of Computer Science*. Sci. 17, 174105, doi: <https://doi.org/10.1007/s11704-022-2096-3>.
20. Amiri, Z., Heidari, A. and Navimipour, N.J. (2022), "Resilient and dependability management in distributed environments: a systematic and comprehensive literature review", *Cluster Computing*, doi: <https://doi.org/10.1007/s10586-022-03738-5>.
21. Thakral, S. and Bansal, D. (2020), "Novel high functionality fault tolerant ALU", *TELKOMNIKA (Telecommunication, Computing, Electronics and Control) Formal Verification of a Fully IEEE Compliant Floating-Point Unit*, Christian Jacobi, Vol. 18, pp. 234-239, doi: <https://doi.org/10.12928/TELKOMNIKA.v18i1.12645>.

Received (Надійшла) 15.03.2023

Accepted for publication (Прийнята до друку) 17.05.2023

Кошман Сергій Олександрович – доктор технічних наук, доцент, професор кафедри безпеки інформаційних систем і технологій, Харківський національний університет імені В. Н. Каразіна, Харків, Україна;

Serhii Koshman – Doctor of Technical Sciences, Assistant Professor, Assistant Professor of Information Systems and Technologies Security Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine;

e-mail: s.koshman@karazin.ua; ORCID ID: <http://orcid.org/0000-0001-8934-2274>.

Краснобаєв Віктор Анатолійович – доктор технічних наук, професор, професор кафедри електроніки і управляючих систем, Харківський національний університет імені В. Н. Каразіна, Харків, Україна;

Victor Krasnobayev – Doctor of Technical Sciences, Professor, Professor of Electronics and Control Systems Department, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine;

e-mail: v.a.krasnobayev@karazin.ua; ORCID ID: <http://orcid.org/0000-0001-5192-9918>.

Нікольський Сергій Борисович – кандидат технічних наук, доцент, доцент кафедри інформатики Комунального закладу "Харківська гуманітарно-педагогічна академія" Харківської обласної ради, Харків, Україна;

Serhii Nikolsky – Candidate of Technical Sciences, Associate Professor, Associate Professor of Informatics Department, Municipal establishment «Kharkiv humanitarian-pedagogical academy» of Kharkiv regional council, Kharkiv, Ukraine;

e-mail: s.nikoloss@gmail.com; ORCID ID: <http://orcid.org/000-0002-3279-6459>.

Ковальчук Дмитро Миколайович – аспірант, Харківський національний університет ім. В. Н. Каразіна, Харків, Україна;

Dmytro Kovalchuk – PhD student, V. N. Karazin Kharkiv National University, Kharkiv, Ukraine

e-mail: kovalchuk.d.n@ukr.net; ORCID ID: <https://orcid.org/0000-0002-8229-836X>.

Структура комп'ютерної системи у залишкових класах

С. О. Кошман, В. А. Краснобаєв, С. Б. Нікольський, Д. М. Ковальчук

Анотація. Предметом статті є постановка та вирішення зворотної задачі оптимального резервування в системі залишкових класів (СЗК) на основі використання методу динамічного програмування. Вирішення поставленого завдання дає можливість підвищити надійність роботи комп'ютерних систем та компонентів (КСК) у СЗК. Метою статті є підвищення надійності функціонування КСК, які будуються на базі використання СЗК, без зниження швидкості обчислень, а також проведення розрахунку та порівняльного аналізу надійності, за ймовірністю безвідмовної роботи, КСК у СЗК та трійованої обчислювальної системи, яка функціонує у позиційній двійковій системі числення (ПСЧ). **Завдання:** провести аналіз впливу системи числення, що використовується, на надійність КСК з урахуванням первинної та вторинної надмірності; синтезувати обчислювальну систему в СЗК для *l*-байтової бітової сітки на основі використання методу пасивної відмовостійкості (постійної структурної надмірності); сформулювати та вирішити обернену задачу оптимального резервування в СЗК на основі використання методу динамічного програмування; для перевірки коректності отриманих результатів провести розрахунок умовної кількості обладнання обчислювальної системи у залишкових класах; оцінити ефективність використання СЗК для підвищення надійності при побудові резервованої КСК по відношенню до резервованої КСК у ПСЧ. **Методи дослідження:** методи аналізу та синтезу комп'ютерних систем, теорія чисел, теорія кодування в СЗК, теорія надійності. Отримано такі **результати.** У роботі показано, що використання ПСЧ, як системи числення, не дозволяє радикально підвищити продуктивність і надійність КСК. У зв'язку з цим у статті розроблено концепцію використання СЗК, як систему числення для побудови КСК. На підставі цього сформульована та вирішена зворотна задача оптимального резервування в СЗК. **Висновки.** Як показали результати розрахунків та порівняльного аналізу використання СЗК забезпечує більш високу надійність КСК, ніж мажоритарна триканальна обчислювальна система в ПСЧ. Отримані результати досліджень можуть бути використані для синтезу відмовостійких комп'ютерних структур у СЗК.

Ключові слова: позиційна кодова структура; система залишкових класів; позиційна двійкова система числення; надійність комп'ютерних систем та компонентів.

Intelligent information systems

УДК 621.39 (045)

DOI: <https://doi.org/10.20998/2522-9052.2023.2.07>Аль-Мудхафар Акіл Абдулхуссейн М.¹, Т. В. Смірнова², К. О. Буравченко², О. А. Смірнов²¹ Національний авіаційний університет, Київ, Україна² Центральнотрапнський національний технічний університет, Кропівницький, Україна

МЕТОД ОЦІНКИ ТА ПІДВИЩЕННЯ КОРИСТУВАЛЬНИЦЬКОГО ДОСВІДУ АБОНЕНТІВ В ПРОГРАМНО-КОНФІГУРОВАНИХ МЕРЕЖАХ НА ОСНОВІ ВИКОРИСТАННЯ МАШИННОГО НАВЧАННЯ

Анотація. Еволюційні процеси, які в першу чергу торкнулися комп'ютерних технологій, призвели до появи кількох типів обчислювальних мереж, що представляють сукупність комп'ютерних пристроїв, об'єднаних в одну систему. Основним призначенням такої системи є доступ користувачів до спільних ресурсів та можливість обміну даними між абонентами у процесі роботи. Такі мережі називаються програмно-конфігурованими – SDN. Мережі SDN вже давно стали основою побудови телекомунікаційних мереж операторського класу. Проте, в них є певна кількість недоліків, які необхідно усунути. **Об'єктом дослідження** є процес оцінки та підвищення користувацького досвіду абонентів в програмно-конфігурованих мережах. **Предметом дослідження** є метод оцінки та підвищення користувацького досвіду абонентів в програмно-конфігурованих мережах на основі використання машинного навчання. **Мета роботи** полягає у розробці моделі та відповідного методу оцінки якості користувацького досвіду абонентів мереж SDN. **У результаті дослідження** вперше було розроблено метод оцінки та підвищення користувацького досвіду абонентів мереж SDN на основі використання машинного навчання. Метод полягає у послідовному проведенні автоматизованого опитування користувачів, вимірюванні показників якості обслуговування абонентів, виборі й побудові регресійної моделі із множини визначених моделей та керування користувацьким досвідом за вимірними параметрами якості обслуговування абонентів мережі SDN. Розроблений метод на відміну від відомих, надає змогу підвищувати якість користувацького досвіду у режимі реального часу. **Висновки.** Проведене дослідження існуючих механізмів керування користувацьким досвідом абонентів та аналіз регресійних моделей на можливість їх використання для встановлення взаємозв'язку між параметрами мережі та користувацьким досвідом, дозволило розробити узагальнену модель оцінки та підвищення користувацького досвіду абонентів мереж SDN, на основі використання машинного навчання, та розробити алгоритм роботи методу. Розроблений метод дозволяє будувати точні моделі взаємозв'язку параметрів QoE та QoS та підвищує на величину до 10% якість користувацького досвіду абонентів мереж SDN.

Ключові слова: програмно-конфігуровані мережі; машинне навчання.

Вступ

Еволюційні процеси, які в першу чергу торкнулися комп'ютерних технологій, призвели до появи кількох типів обчислювальних мереж, що представляють сукупність комп'ютерних пристроїв, об'єднаних в одну систему. Основним призначенням такої системи є доступ користувачів до спільних ресурсів та можливість обміну даними між абонентами у процесі роботи.

Сучасний стан та тенденції розвитку комп'ютерних мереж показали, що потенціал зростання продуктивності та пропускної спроможності мереж на основі традиційних технологій практично вичерпаний [1]. В даний час телекомунікаційні мережі будуються на базі безлічі комутаційних пристроїв, кожен з яких може працювати незалежно від інших пристроїв, самостійно визначаючи правила, за якими воно оброблятиме пакети, що надходять на нього. Також сучасні комутаційні пристрої виробляють свої правила комутації пакетів за закладеними в них алгоритмами. Правила обміну службовою інформацією між пристроями та способи її застосування для вироблення узгоджених правил обробки пакетів визначаються безліччю різноманітних службових протоколів, кожен із яких вирішує деяку підмножину завдань адміністрування мережі. Використання безлічі службових протоколів аж ніяк не робить мережне адміністрування

простим. Вирішення цієї проблеми значно спростилося разом із порівняно недавнім появою концепції Програмно конфігурованих мереж (SDN). Такі мережі дозволяють прискорити маршрутизацію, підвищити зручність конфігурування, віртуалізації, налаштування якості обслуговування мереж зв'язку.

Головними завданнями SDN є відокремлення функцій передачі трафіку від функцій управління (включаючи контроль як самого трафіку, так і пристроїв, що здійснюють його передачу). Це відбувається за рахунок створення спеціального програмного забезпечення, яке може працювати на окремому сервері (комп'ютері) та яке знаходиться під контролем адміністратора мережі. Всі маршрутизатори та комутатори об'єднуються під керуванням контролера SDN або Мережевої Операційної Системи (МОС), яка забезпечує додатком доступ до управління мережею та постійно відстежує конфігурацію засобів мережі.

Аналіз останніх досліджень і публікацій. Поняття, практичні особливості та застосування QoE значно розвинулися з моменту його створення в телекомунікаційному контексті [2]. QoE визначається як «загальна прийнятність програми або послуги, яка суб'єктивно сприймається кінцевим користувачем». QoE вважалось таким, що включає «наскрізні системні ефекти» і що це може бути змінено під впливом очікувань користувачів і контексту.

Еволюція методів оцінки QoE та включала вдосконалення агрегованих вручну оцінок, наприклад, середньої оцінки думки (MOS); стандартизацію відображення QoE для мережевих вимірювань надання послуг; дослідження асоціації/кореляції з спостереженнями та відповідями кінцевого користувача, тощо [3].

QoE можна визначити як процес вимірювання або оцінки якості для набору програм або послуг користувачів за допомогою спеціальних процедур та з урахуванням факторів впливу (можливо контрольованих, вимірних або просто зібраних і звітних).

Наприклад, оцінка якості на основі методів SDN/NFV [4-6] є важливим кроком до моніторингу та управління на основі якості. Залежно від мети та напрямку дослідження в рекомендаціях ITU-T описані різні методи та рекомендації для суб'єктивного оцінювання [7-9].

У роботі [10] пропонується динамічний підхід до розподілу ресурсів за допомогою VBS і RRH, спрямований на підвищення ефективності використання ресурсів та енергії, забезпечуючи при цьому високий рівень QoE.

Постановка задач дослідження. Хоча SDN є відповідним рішенням для ІТ та хмарних провайдерів та підприємств, SDN стикається з деякими проблемами, які перешкоджають його продуктивності та впровадженню.

Однією із таких проблем є відсутність дієвого механізму управління користувальницьким досвідом кінцевих користувачів.

Таким чином, **метою даної роботи** є розробка методу оцінки та підвищення користувальницького досвіду абонентів мереж SDN на основі використання машинного навчання.

Для досягнення поставленої мети було вирішено наступні задачі:

1. Проведено аналіз існуючих механізмів керування користувальницьким досвідом абонентів.
2. Розроблено узагальнену модель оцінки та підвищення користувальницького досвіду абонентів мереж SDN.
3. Проведено аналіз регресійних моделей на можливість їх використання для встановлення взаємозв'язку між параметрами мережі та користувальницьким досвідом.
4. Розроблено алгоритм роботи методу.

Розвиток сучасних мережевих технологій на базі програмно-кофігурованих мереж

Розглянемо архітектуру програмно-конфігурованих мереж. В них функція комутатора з управління переносяться на окремий центральний пристрій – контролер SDN. Такий підхід дозволяє керувати та контролювати стан мережі на логічно централізованому контролері. З іншого боку, з'являється можливість рівню управління відокремитися від фізичної складової, використовуючи логічне уявлення мережі загалом. Взаємодія між рівнем передачі здійснюється за допомогою єдиного уніфікованого відкритого інтерфейсу. В архітектурі SDN згідно [11] виділяють 3 рівні (рис. 1):

– рівень інфраструктури мережі, що представляє набір мережевих структур (комутатори та канали зв'язку);

– рівень управління, що складається з операційної системи, що забезпечує додаткам мережеві сервіси та програмний інтерфейс для управління мережевими пристроями та мережею;

– рівень додатків.

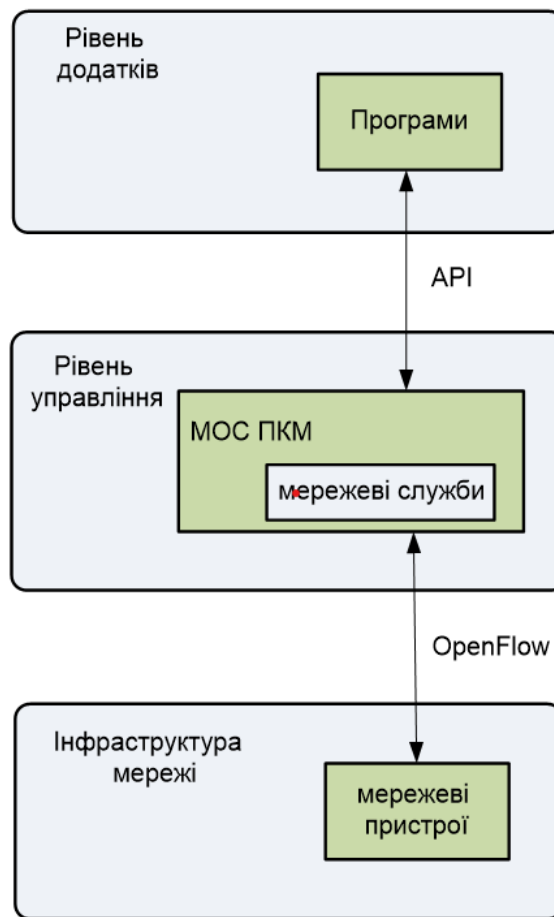


Рис 1. Архітектура SDN (Fig. 1. SDN architecture)

Одне з центральних місць в архітектурі SDN займає мережева операційна система, яка є операційною системою, що забезпечує обробку, зберігання та передачу даних в інформаційній мережі. Мережева операційна система або контролер SDN визначає взаємопов'язану групу протоколів верхніх рівнів, що забезпечують основні функції мережі: адресацію об'єктів, функціонування служб, безпеку даних, керування мережею. Контролер SDN формує дані про стан всіх ресурсів мережі та забезпечує доступ до них для програм управління мережею. Ці програми керують різними аспектами функціонування мережі, включаючи побудову топології, прийняття маршрутизуючих рішень, балансування навантаження. За допомогою цього протоколи фахівці самі можуть визначати та контролювати, які вузли, за яких умов та з якістю можуть взаємодіяти в мережі. Таким чином, на контролер ще можуть бути покладені задачі по оцінці та підвищенні користувальницького досвіду абонентів, що обслуговуються мережею.

Узагальнена модель забезпечення користувальницького досвіду в мережах SDN

В останні роки технічна спільнота переключила певну увагу з одного пов'язаного показника, якості обслуговування (QoS), на більш орієнтовану на споживача метрику, якість досвіду (QoE). У той час як QoS стоїть між мережею та додатком, QoE зосереджено на абоненті. Зокрема, QoE фокусується на людині як користувачеві, який взаємодіє з додатком, і людині як клієнту, який має справу з постачальником послуг.

Різниця між QoE і QoS підкреслена нижче [12]:

QoS – якість обслуговування:

- характеристики/поведінка мережі;
- гарантії продуктивності, надані провайдером мережі на основі вимірювань;

QoE – якість досвіду:

- вплив поведінки мережі на кінцевого користувача;
- деякі недоліки можуть залишитися непоміченими;
- деякі недоліки можуть зробити додаток марним;
- не фіксується мережевими вимірюваннями.
- QoE враховує очікування користувача, QoS є більш раціональним на основі технічних вимірювань (рис. 2).

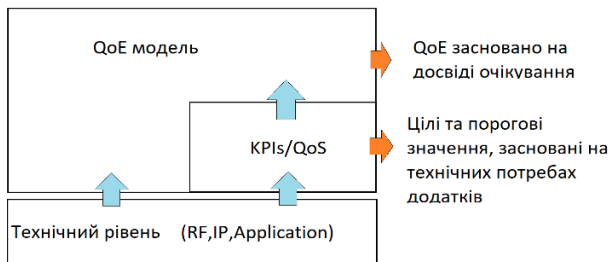


Рис 2. Зв'язки між QoE, QoS і KPI
(Fig. 2. Relationships between QoE, QoS and KPI)

На підставі вищезазначеного пропонується модель аналізу оцінки загального QoE за допомогою показників QoS, які можна оцінити більш об'єктивним способом.

$$\left\{ \bigcup_{i=1}^n S_i \right\} = \{S_1, S_2, \dots, S_n\},$$

де $S_i \subseteq S, (i = \overline{1, n})$ – кількість послуг;

$$S_i = \left\{ \bigcup_{j=1}^{m_i} S_{ij} \right\} = \{S_{i1}, S_{i2}, \dots, S_{jm_i}\};$$

$S_{ij} (j = \overline{1, m_i})$ – підмножина елементів системи забезпечення якості.

Підмножини QoE метрики $S_{ij} \subseteq S_i$ можуть бути представлені як:

$$S_{ij} = \left\{ \bigcup_{p=1}^{r_{ij}} S_{ijp} \right\} = \{S_{ij1}, S_{ij2}, \dots, S_{ijr_{ij}}\},$$

де $S_{ijp} (p = \overline{1, r_{ij}})$ показники QoE, що характеризують – QoE для S_{ij} ;

r_{ij} – кількість таких показників.

На другому етапі вибирають показники QoS і QoE, використовуючи багатофакторний кореляційно-регресійний аналіз. Для побудови багатофакторної регресійної моделі необхідно виконати наступні кроки:

1. Вибираються всі можливі фактори QoS, що впливають на досліджуваний показник QoE (або процес). Для кожного фактора визначаються його чисельні характеристики.

2. Вибирається форма регресійної або багатомірної моделі, тобто знаходження аналітичного виразу, що найбільш повно відображає зв'язок факторних характеристик з рівнодійною, тобто відбувається вибір відповідних функцій (табл. 1).

На наступному етапі необхідно розрахувати підмножини показників QoS з використанням відповідних алгоритмів та формул їх розрахунку [13]. QoE необхідно розраховувати з використанням, наприклад, MOS, DSCQR, ACR [7] або інших відповідних методів/технік.

На останньому етапі отримані значення порівнюються з максимально допустимими, можливими для забезпечення нормального функціонування мережі та досягнутих KPIs.

Для порівняння отриманих у результаті розрахунків значень, із максимально допустимими, вводиться логічна функція еквівалентності:

$$E(x, y) = \begin{cases} 1, & \text{if } x > y, \\ 0, & \text{if } x \leq y. \end{cases}$$

QoE – чи не найважливіший параметр, за оцінкою якого можна визначити досвід користувача і порівняти його з очікуваннями інших користувачів.

Саме тому було розроблено метод оцінки QoE, щоб оцінити рівень задоволеності кінцевих користувачів та гравців вертикалей розгорнутими варіантами використання. Це включає QoE користувачів, а також відгуки вертикальних гравців про те, як технологія, що надається, може поліпшити їх бізнес-операції [13].

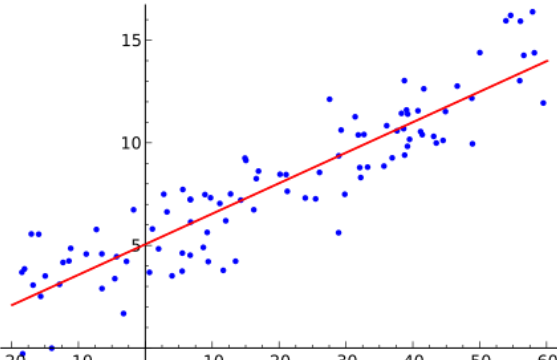
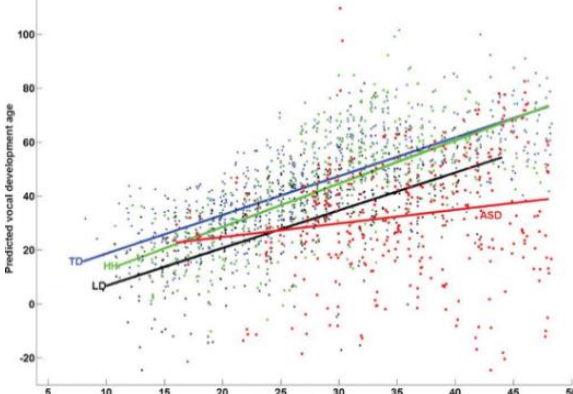
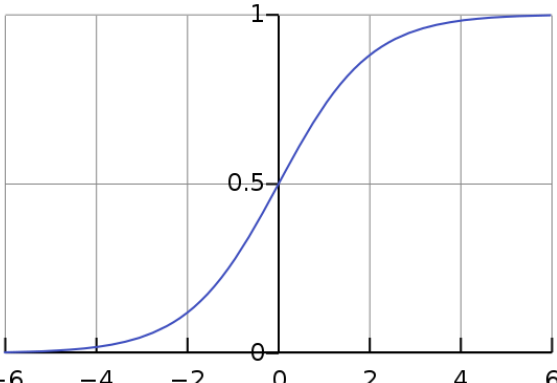
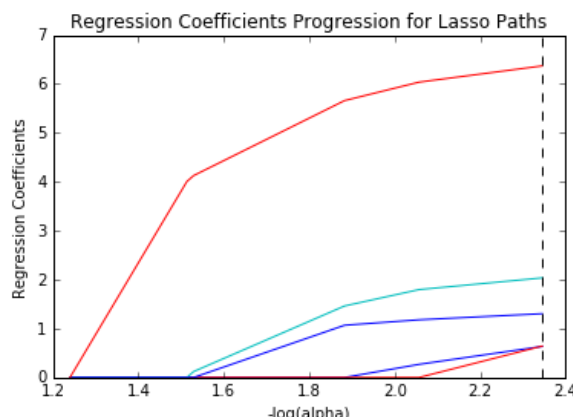
На додаток до перевірки результатів QoS, які в основному ілюструють продуктивність ключових показників ефективності мережі, надзвичайно важливо підтвердити фактичну задоволеність кінцевих користувачів та вертикальних гравців. (або як постачальників послуг, або користувачів вторинних послуг).

У цьому напрямі було розроблено метод оцінки QoE, високорівнева архітектура якого представлена на рис. 3 [13]. Метод реалізується у два етапи:

– Перший етап реалізується під час виконання випробувань як метрики QoS, що автоматично збираються з інфраструктури, так і метрики QoE (і вертикальну задоволеність), зібрані за допомогою відповідних анкет.

– Другий етап реалізується після виконання випробувань та з використанням кореляційно-регресійного аналізу, з якого створюється модель для кореляції QoS-QoE.

Таблиця 1 – Набір досліджуваних моделей (Table 1 – Set of studied models)

Назва моделі /Графічна модель	Математична модель (вираз)
<p>Лінійна регресія [12]</p> 	$Y = a + b \cdot X + \varepsilon ,$ <p>де X – пояснювальна змінна; Y – залежна змінна; b – нахил прямої; c – перехоплення; ε – помилка моделі.</p>
<p>Множинна лінійна регресія [12]</p> 	$Y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_p \cdot x_p + \varepsilon ,$ <p>де Y – прогнозоване значення; β_0 – точка перетину Y (значення y, коли всі інші параметри встановлені на 0); β_p – коефіцієнт регресії незалежної змінної; x_p – кількість незалежних змінних; ε – помилка моделі.</p>
<p>Логістична регресія [12]</p> 	$Y = \frac{1}{1 + e^{-(\beta_0 + \beta_1 \cdot x)}} ,$ <p>де Y – прогнозоване значення; β_0 – є точкою перетину Y; β_1 – коефіцієнт регресії змінної x.</p>
<p>Регресія Ласо [12]</p> 	$\sum_{i=1}^n \left(y_i - \sum_j x_{ij} \cdot \beta_j \right)^2 + \lambda \cdot \sum_{j=1}^p \beta_j ,$ <p>де y_i – прогнозоване значення; β_j – коефіцієнт регресії незалежної змінної x; λ – позначає величину усадки.</p>

Закінчення табл. 1

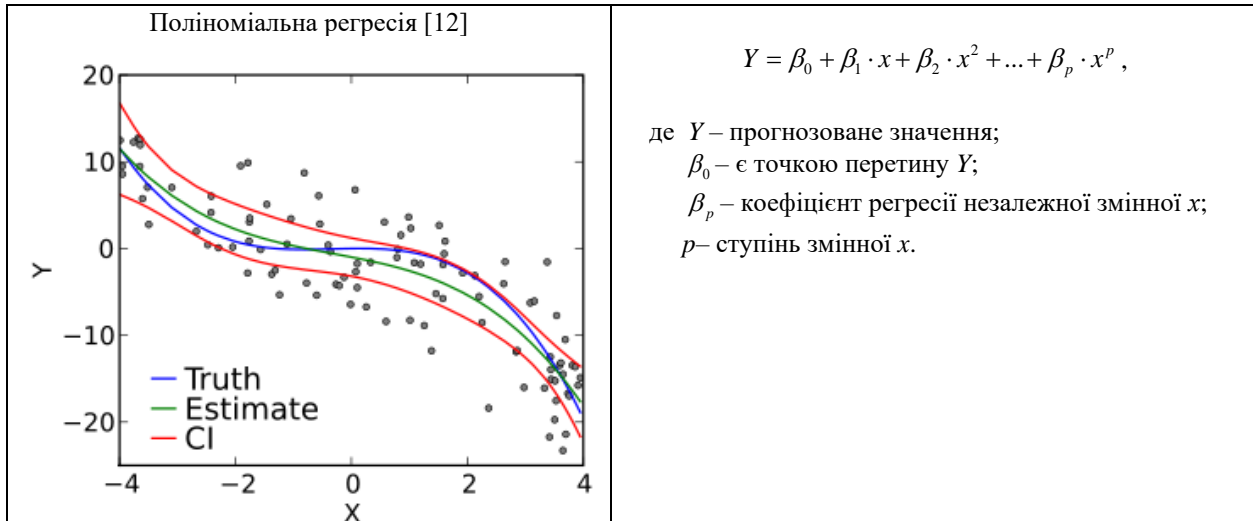


Рис. 3. Метод оцінки QoE в мережі SDN
 (Fig. 3. QoE evaluation method in the SDN network)

Кожен підхід, використаний у попередніх дослідженнях для вимірювання QoS та QoE, має свій потенціал. Очікується, що цей потенціал забезпечить баланс між QoS та QoE, простий, швидкий та точний процес, а також гнучкість між ідеальними умовами та реальністю.

Вимірювання трафіку та відгуку користувача у групі об'єктних підходів має потенціал завдяки балансу QoS та QoE, простоті та гнучкості, як у підматеріалі, який стає метою вимірювання. Це пов'язано з тим, що трафік має різні підматеріали, а також підматеріали у відповідях користувачів, які дуже гнучкі. Користувальницький трафік та відповідь також відповідні ціннішим вимогам високотехнологічних та систематичних QoS та QoE, які дуже прив'язані до користувачів.

Окремо треба виділити етап вибору регресійної моделі. Цей вибір може бути доволі складним.

Спроба змоделювати його лише за допомогою зразка не полегшить цей вибір.

Тому виникає необхідність математично описати взаємозв'язок між деякими предикторами та змінною відгуком. На практиці, зазвичай вимірюється безліч змінних параметрів, але в модель включаються лише деякі з них. Тоді, намагаються виключити змінні, які не пов'язані між собою, і включають лише ті, що мають справжній взаємозв'язок. В цьому випадку розглядають множину можливих моделей. При цьому, прагнуть досягти балансу щодо кількості предикторів, які вони включають.

– **Замало:** недостатньо конкретизована модель має тенденцію давати необ'єктивні оцінки.

– **Занадто багато:** завищена модель зазвичай дає менш точні оцінки.

– **Оптимальна:** модель з правильними умовами не має зміщення та дає найточніші оцінки.

Щоб вибрати найбільш відповідну модель регресії, краще включити змінні, які спеціально тестуються, разом з іншими змінними, які впливають на відповідь, щоб уникнути упереджених результатів. Існують статистичні заходи та процедури, що допомагають конкретизувати регресійну модель [14].

Скоригований R-квадрат та Прогнозований R-квадрат: моделі з вищими скоригованими та прогнозованими значеннями R-квадрату. Ці статистичні дані призначені для того, щоб уникнути ключової проблеми зі звичайним R-квадратом – він збільшується щоразу, коли ви додаєте предиктор, і може обдурити вас, змусивши вказати занадто складну модель.

– Скоригований квадрат R збільшується тільки в тому випадку, якщо новий член покращує модель більше, ніж можна було б очікувати випадково, і він також може зменшуватись з предикторами низької якості.

– R-квадрат, що прогнозується, є формою перехресної перевірки, і він також може зменшуватись. Перехресна перевірка визначає, наскільки добре ваша модель узагальнюється інші набори даних шляхом поділу ваших даних.

R-значення для предикторів: у регресії низькі p -значення вказують терміни, які є статистично значущими. "Скорочення моделі" відноситься до практики включення в модель всіх предикторів-кандидатів, а потім систематичного видалення члена з найвищим значенням p один за одним доти, доки не залишаться тільки значущі предиктори.

Покрокова регресія та регресія найкращих підмножин: це дві автоматизовані процедури, які можуть ідентифікувати корисні предиктори на етапах побудови моделі. На рис. 4 нижче показано схему вибору методу та подальшого створення моделі.

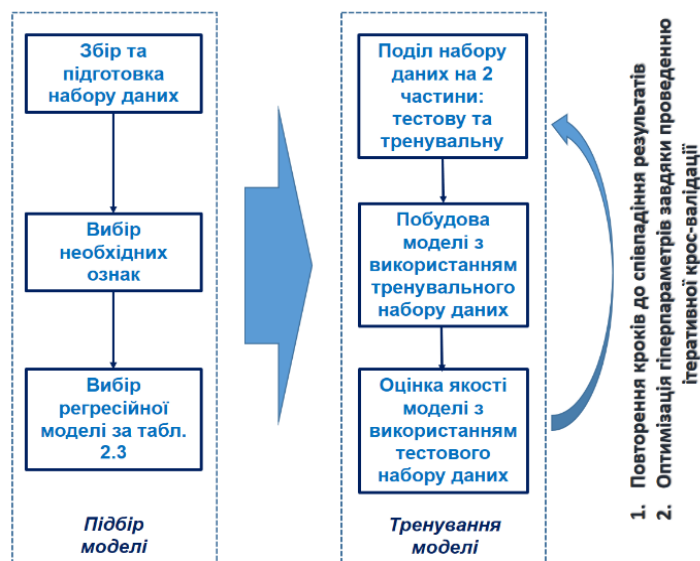


Рис. 4. Вибір методу регресії та подальше створення моделі QoE-to-QoS
(Fig. 4. Selection of the regression method and subsequent creation of the QoE-to-QoS model)

Вибір функції вже відбувається до побудови моделі та визначає вхідні атрибути пізнішої моделі регресії. Набори даних уже були структуровані під час створення таким чином, що вони містять лише відповідні атрибути.

Методи регресії підходять для різних проблем, по-різному добре. Для оцінки набір даних розбивається на навчальний і тестовий набір даних перед створенням моделі. Цей крок автоматично виконується ітеративно під час перехресної перевірки.

Оптимізація гіперпараметрів шляхом повторюваної перехресної перевірки з різними налаштуваннями гіперпараметрів

Результатом перехресної перевірки є список зі значеннями вибраних параметрів скорингу. Оскільки оцінка виконується після кожного запуску, якщо набір даних поділено на п'ять розділів, також існує список із п'ятьма значеннями оцінки. Усереднення цих значень дозволяє оцінити ефективність процедури регресії. Оскільки більшість методів регресії дозволяють регулювати складність моделі за допомогою одного або кількох гіперпараметрів, коригування

гіперпараметрів є необхідним для значущого порівняння методів регресії. Знаходження цих оптимальних налаштувань гіперпараметрів виконується шляхом ітеративної побудови моделі. Перехресна перевірка виконується неодноразово для різних налаштувань гіперпараметрів. Нарешті, вибираються налаштування параметрів, які показали найкращу точність моделі під час оцінювання. Цей процес виконується за допомогою циклів, які автоматично змінюють гіперпараметри в певних межах і зберігають оцінні значення. Потім вибір оптимальних налаштувань виконується шляхом ручного або автоматичного пошуку найкращих результатів оцінювання.

Висновки

Мережі SDN вже давно стали основою побудови телекомунікаційних мереж операторського класу. Проте, в них є певна кількість недоліків, які необхідно усунути. Відповідно, метою даної статті є розробка моделі та відповідного методу оцінки якості користувальницького досвіду абонентів мереж SDN.

Таким чином, вперше було розроблено метод оцінки та підвищення користувальницького досвіду

абонентів мереж SDN на основі використання машинного навчання, що полягає у послідовному проведенні автоматизованого опитування користувачів, вимірюванні показників якості обслуговування абонентів, виборі та побудові регресійної моделі із множини визначених моделей та керування користувальницьким досвідом за вимірними параметрами якості

обслуговування абонентів мережі SDN, що на відміну від відомих, надає змогу підвищувати якість користувальницького досвіду у режимі реального часу.

Даний метод дозволяє будувати точні моделі взаємозв'язку параметрів QoE та QoS та підвищує на величину до 10% якість користувальницького досвіду абонентів мереж SDN.

REFERENCES

- Oleshchenko L.M. (2018), *Organization of computer networks*, summary of lectures, education. manual for students specialty 121 "Software engineering", specialization "Computer and information search system software", KPI named after Igor Sikorsky, Kyiv, 225 p., available at: <https://ela.kpi.ua/handle/123456789/22890>.
- Porcu, S., Floris, A., Voigt-Antons, J.-N., Atzori, L. and Moller, S. (2020), "Estimation of the Quality of Experience During Video Streaming From Facial Expression and Gaze Direction", *IEEE Trans. Netw. Serv. Manag.*, Vol. 17, Is. 4, pp. 2702–2716, doi: <https://doi.org/10.1109/TNSM.2020.3018303>.
- Kourtis, M.-A., Liberal, F., Koumaras, H., Xilouris, G. and Trouva, E. (2017), "Exploiting NFV techniques towards future VQA methods", *IEEE 22nd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, Lund, Sweden, June 2017, doi: <https://doi.org/10.1109/CAMAD.2017.8031638>.
- Kourtis, M.-A., Koumaras, H., Xilouris, G. and Liberal, F. (2017), "An NFVBased Video Quality Assessment Method over 5G Small Cell Networks", *IEEE MultiMedia*, vol. 24, pp. 68–78, doi: <https://doi.org/10.1109/MMUL.2017.265091534>.
- Koumaras, H., Kourtis, M., Sakkas, C., Xilouris, G. and Kolometsos, S. (2016), "In-service Video Quality assessment based on SDN/NFV techniques", *2016 23rd International Conference on Telecommunications (ICT)*, Thessaloniki, Greece, pp. 1–5, doi: <https://doi.org/10.1109/ICT.2016.7500385>.
- International Telecommunication Union (2012), *ITU-R, BT.500.11: Methodology for the subjective assessment of the quality of television pictures*, Jan 2012, available at: <https://www.itu.int/rec/R-REC-BT.500>.
- International Telecommunication Union (2008), ITU-T Rec. P.910: Subjective video quality assessment methods for multimedia applications, April 2008, available at: <https://www.itu.int/rec/T-REC-P.910>.
- International Telecommunication Union (2016), "ITU-T Rec. P.913: Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment," March 2016, available at: <https://www.itu.int/rec/T-REC-P.913>.
- Ouali, K., Kassar, M., Nguyen, T. M. T., Sethom, K. and Kervella, B. (2017), "Modeling D2D handover management in 5G cellular networks", *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, available at: <https://doi.org/10.1109/IWCMC.2017.7986285>.
- Makarenko Yu. V. (2019), "The Analysis of the Possibilities of Using 5G Technology in Internet Systems of Things", *Electronic and Acoustic Engineering*, vol. 2, no. 2, pp. 45–51, doi: <https://doi.org/10.20535/2617-0965.2019.2.2.163355>.
- Thirupathi, V., Sandeep, C. H., Kumar, N., & Kumar, P. P. (2019), "A comprehensive review on sdn architecture, applications and major benefits of SDN", *International Journal of Advanced Science and Technology*, 28(20), pp. 607-614, doi: <http://sersc.org/journals/index.php/IJAST/article/view/2854>.
- Osman, H., Kosmatos, E., Giannopoulou, N. and Odarchenko, R. (2021), *First Integrated 5G-TOURS Ecosystem*, D7.2, Technical Report. 5G-TOURS - ICT-19-2019 – G.A:856950, available at: https://www.researchgate.net/publication/350725143_First_Integrated_5G-TOURS_Ecosystem_D72.
https://www.researchgate.net/publication/350725143_First_Integrated_5G-TOURS_Ecosystem_D72#fullTextFileContent.
- Mouhouche, B. (2020), *5G-TOURS, D7.1*, 5G Experimentation Facilities and Vertical Trials, available at: https://5g-ppp.eu/wp-content/uploads/2020/09/5GTOURS_Vertical_workshop_14102020.pdf.
- Tikhvinskiy, V., Bohechka, G., Gryazev, A. and Aitmagambetov, A. (2018), "Comparative Analysis of QoS Management and Technical Requirements in 3GPP Standards for Cellular IoT Technologies", *Journal of Telecommunications and Information Technology*, vol. 2(2), pp. 41–47, doi: <https://doi.org/10.26636/jtit.2018.122717>.

СПИСОК ЛІТЕРАТУРИ

- Олещенко Л.М. Організація комп'ютерних мереж, конспект лекцій: навч. посіб. для студ. спеціальності 121 «Інженерія програмного забезпечення» спеціалізації «Програмне забезпечення комп'ютерних та інформаційно-пошукових систем». Київ : КПІ ім. Ігоря Сікорського, 2018. 225 с. URL: <https://ela.kpi.ua/handle/123456789/22890>.
- Porcu, S., Floris, A., Voigt-Antons, J.-N., Atzori, L., Moller, S. Estimation of the Quality of Experience During Video Streaming From Facial Expression and Gaze Direction. *IEEE Trans. Netw. Serv. Manag.* 2020. Vol. 17, Is. 4. P. 2702–2716. DOI: <https://doi.org/10.1109/TNSM.2020.3018303>.
- Kourtis M.-A., Liberal F., Koumaras H., Xilouris G., Trouva E. Exploiting NFV techniques towards future VQA methods. *IEEE 22nd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*. Lund, Sweden. June 2017. DOI: <https://doi.org/10.1109/CAMAD.2017.8031638>.
- Kourtis M.-A., Koumaras H., Xilouris G., Liberal F. An NFVBased Video Quality Assessment Method over 5G Small Cell Networks. *IEEE MultiMedia*. Dec 2017. Vol. 24. P. 68–78. DOI: <https://doi.org/10.1109/MMUL.2017.265091534>.
- Koumaras H., Kourtis M., Sakkas C., Xilouris G., Kolometsos S. In-service Video Quality assessment based on SDN/NFV techniques. *2016 23rd International Conference on Telecommunications (ICT)*. May 2016. Thessaloniki, Greece. P. 1–5. DOI: <https://doi.org/10.1109/ICT.2016.7500385>.
- International Telecommunication Union. ITU-R, BT.500.11: Methodology for the subjective assessment of the quality of television pictures. Jan 2012. URL: <https://www.itu.int/rec/R-REC-BT.500>.
- International Telecommunication Union. ITU-T Rec. P.910: Subjective video quality assessment methods for multimedia applications. April 2008. URL: <https://www.itu.int/rec/T-REC-P.910>.

8. International Telecommunication Union. ITU-T Rec. P.913: Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment. March 2016. URL: <https://www.itu.int/rec/T-REC-P.913>.
9. Ouali K., Kassar M., Nguyen T. M. T., Sethom K., Kervella B. Modeling D2D handover management in 5G cellular networks. *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*. Jun. 2017. DOI: <https://doi.org/10.1109/IWCMC.2017.7986285>.
10. Макаренко Ю.В. Аналіз можливостей використання технологій 5G в системах Інтернету речей. *Електронна та Акустична Інженерія*. 2019. Том 2, № 2. С. 45–51. DOI: <https://doi.org/10.20535/2617-0965.2019.2.2.163355>.
11. Thirupathi, V., Sandeep, C. H., Kumar, N., Kumar, P. P. A comprehensive review on sdn architecture, applications and major benefits of SDN. *International Journal of Advanced Science and Technology*. 2019. Vol. 28(20). P. 607-614. DOI: <http://sersc.org/journals/index.php/IJAST/article/view/2854>.
12. Osman H., Kosmatos E., Giannopoulou N. and Odarchenko R. (2021), First Integrated 5G-TOURS Ecosystem, D7.2, Technical Report. 5G-TOURS - ICT-19-2019 – G.A:856950, URL: https://www.researchgate.net/publication/350725143_First_Integrated_5G-TOURS_Ecosystem_D7.2. https://www.researchgate.net/publication/350725143_First_Integrated_5G-TOURS_Ecosystem_D7.2#fullTextFileContent.
13. Mouhouche, B. 5G-TOURS, D7.1, 5G Experimentation Facilities and Vertical Trials, 2020, URL: https://5g-ppp.eu/wp-content/uploads/2020/09/5GTOURS_Vertical_workshop_14102020.pdf.
14. Tikhvinskiy V., Bochechka G., Gryazev A., Aitmagambetov A. Comparative Analysis of QoS Management and Technical Requirements in 3GPP Standards for Cellular IoT Technologies. *Journal of Telecommunications and Information Technology*. Vol. 2. Jun. 2018. P. 41–47, DOI: <https://doi.org/10.26636/jtit.2018.122717>.

Received (Надійшла) 28.01.2023

Accepted for publication (Прийнята до друку) 12.04.2023

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Аль-Мудхадфар Акіл Абдулхуссейн М. – PhD-докторант, Національний авіаційний університет, Київ, Україна;

Al-Mudhafar Aqeel Abdulhussein M – PhD student, National Aviation University, Kyiv, Ukraine;

e-mail: almudhaffar2004@gmail.com; ORCID ID: <https://orcid.org/0000-0001-5306-4081>.

Смірнова Тетяна Віталіївна – кандидат технічних наук, доцент, доцент кафедри кібербезпеки та програмного забезпечення, Центральноукраїнський національний технічний університет, Кропивницький, Україна;

Tetiana Smirnova – candidate of technical sciences (PhD), associate professor, associate professor of cybersecurity and software academic department, Central Ukrainian National Technical University, Kropivnitskiy, Ukraine;

e-mail: sm.tetyana@gmail.com; ORCID ID: <http://orcid.org/0000-0001-6896-0612>.

Буравченко Костянтин Олегович – кандидат технічних наук, доцент, доцент кафедри кібербезпеки та програмного забезпечення, Центральноукраїнський національний технічний університет, Кропивницький, Україна;

Kostiantyn Buravchenko – candidate of technical sciences, associate professor, associate professor of cybersecurity and software academic department, Central Ukrainian National Technical University, Kropivnitskiy, Ukraine;

e-mail: buravchenkok@gmail.com; ORCID ID: <http://orcid.org/0000-0001-6195-7533>.

Смірнов Олексій Анатолійович – доктор технічних наук, професор, завідувач кафедри кібербезпеки та програмного забезпечення, Центральноукраїнський національний технічний університет, Кропивницький, Україна;

Oleksii Smirnov – doctor of engineering, professor, head of cybersecurity and software academic department, Central Ukrainian National Technical University, Kropivnitskiy, Ukraine;

e-mail: dr.SmirnovOA@gmail.com; ORCID ID: <http://orcid.org/0000-0001-9543-874X>.

The method of assessing and improving the user experience of subscribers in software-configured networks based on the use of machine learning

Al-Mudhafar Aqeel Abdulhussein M., Tetiana Smirnova, Kostiantyn Buravchenko, Oleksii Smirnov

Abstract. Evolutionary processes, which primarily affected computer technologies, led to the appearance of several types of computer networks, representing a set of computer devices combined into one system. The main purpose of such a system is user access to shared resources and the ability to exchange data between subscribers during work. Such networks are called software-configurable networks - SDN. SDN networks have long been the basis for building carrier-class telecommunication networks. However, they have a certain number of shortcomings that must be eliminated. **The object of the study** is the process of evaluating and improving the user experience of subscribers in software-configured networks. **The subject of the study** is a method of evaluating and improving the user experience of subscribers in software-configured networks based on the use of machine learning. **The purpose of the work** is to develop a model and a corresponding method for assessing the quality of the user experience of subscribers of SDN networks. **As a result of the research**, for the first time, a method of evaluating and improving the user experience of subscribers of SDN networks was developed based on the use of machine learning. The method consists in sequentially conducting an automated survey of users, measuring indicators of subscriber service quality, selecting and building a regression model from a set of defined models, and managing the user experience according to the measured parameters of SDN subscriber service quality. The developed method, in contrast to the known ones, makes it possible to improve the quality of the user experience in real time. **Conclusions.** The research of existing mechanisms for managing the user experience of subscribers and the analysis of regression models for the possibility of using them to establish the relationship between network parameters and user experience allowed to develop a generalized model for evaluating and improving the user experience of subscribers of SDN networks, based on the use of machine learning, and to develop an algorithm works of the method. The developed method makes it possible to build accurate models of the relationship of QoE and QoS parameters and increases the quality of the user experience of subscribers of SDN networks by up to 10%.

Keywords: software-configured networks; machine learning.

Dmytro Dashkov, Oleksii Liashenko

Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

MOTION CAPTURE WITH MEMS SENSORS

Abstract. The object of this article is the registration and analysis of human movements based on sensors. This paper presents a comparison of the basic methods of data processing from inertial micromechanical sensors to collect data a device was implemented that captures movements. As result the device uses the motion data from accelerometer and gyroscope to calculate the motion trajectory: the angle of rotation and acceleration. The data is read by the microcontroller, after which it is filtered and processed by one of the filters (Complementary, Kalman), and finally transferred to a computer for further analysis and display. **The purpose of the article** is to compare several methods of data processing from microelectromechanical. **The results obtained:** device was developed, obtained data that can be used to characterize the methods and analyze their work in the system. **Conclusions:** In the course of the study, a device was developed for collecting and processing data from MEMS sensors, which showed the effectiveness of the complementary filter in comparison with the Kalman filter in real-time systems with limited computing power. Real results confirmed that the results of the complementary method using less computational resources are not far behind the more costly Kalman filter without the use of auxiliary sensors, like a digital compass.

Keywords: MEMS; MCU; gyroscope; accelerometer; motion detecting; Complementary filter; Kalman filter.

Introduction

Problem statement. The problem of recording the movement of human limbs includes the task of accurately tracking and analyzing the movements of various parts of the body during various actions. There are various approaches to capture human limb motion, including optical computer vision (CV), inertial measurement sensors (IMU), and electromyography (EMG). Each of these methods has advantages and disadvantages, and the choice of technique depends on the specific application and the accuracy and precision required.

Optical motion capture systems use cameras to track the position of a person's limbs using computer vision algorithms. Additionally, this technology can use markers attached to the object's limbs for greater stability. These systems are highly accurate and can capture detailed motion data, but are typically expensive and require a controlled environment with minimal interference.

They can also be affected by changes in lighting conditions or occlusion in the scene. Computer vision methods can also be more susceptible to noise and errors than gyroscope sensors, especially if the camera is not stabilized or if there are moving objects in the scene.

IMUs are sensors that are attached to limbs and measure acceleration, angular velocity, and magnetic field strength to detect changes in an object's orientation and rotation. These sensors can be used in a wide range of environments and are relatively inexpensive, but they can suffer from drift and noise, which can affect measurement accuracy. Gyro sensors are generally more accurate than computer vision methods for detecting rotation and are less susceptible to noise or errors. However, gyroscope sensors are less flexible than computer vision methods because they are designed specifically to detect rotation and cannot detect other types of motion.

The EMG approach measures electrical activity in limb muscles and can be used to determine limb

movement. This method is noninvasive and can be used in a variety of situations, but it may have interference and may not provide as accurate or detailed information about limb motion as optical or IMU-based approaches.

So computer vision techniques can be more flexible and can detect a wider range of motion, but they can also be computationally expensive and more sensitive to noise and errors. Gyro sensors are more accurate for detecting rotation, but are less flexible and subject to drift over time. The choice of method will depend on the specific requirements of the application and available resources.

IMU sensors are most suitable in case of tracing separate human body parts, the final result can be small device, that connects to different limbs to precisely track motions [1].

Main material

IMU sensors. One of the types of IMU sensors is a digital gyroscope, it is a device that measures the angular velocity, or the rate at which an object's orientation changes over time. It is used in various applications such as navigation, robotics and virtual reality. The last one could help in areas of human-computer interaction. The main principle of operation of a digital gyroscope is the Coriolis effect. When an object rotates, a force acts on it that is perpendicular to the direction of its motion. This force is known as the Coriolis force. This force arises due to the fact that different points of the rotating surface move at different speeds, in other words, the Coriolis effect can be considered as a force felt by all objects that move along the surface of a rotating object. In a digital gyroscope, this force is measured using a vibrating structure known as a MEMS (Micro Electro-Mechanical System) sensor. The MEMS sensor consists of a test mass suspended from a spring. As the gyroscope rotates, the Coriolis force causes the test mass to vibrate in a direction perpendicular to the plane of rotation. Capacitive sensors located at the edges of the reference mass detect vibration.

As the test mass moves (Fig 1, 2), the capacitance between the test mass and the sensors changes, producing an electrical signal proportional to the speed of rotation.

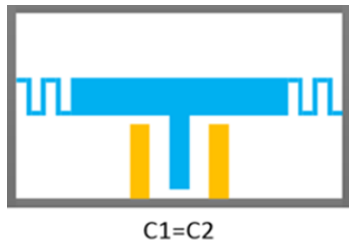


Fig. 1. Normal sensor status

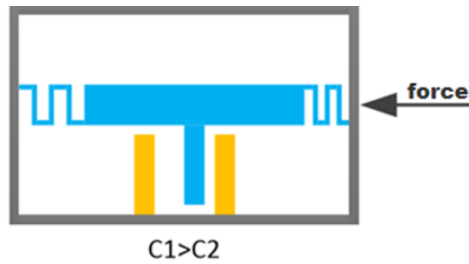


Fig. 2. Sensor status under force

Hardware. Comparison of algorithms requires data from sensors, and to obtain data, a device was developed that tracks the movements of human limbs, which includes boards with a microprocessor and inertial microelectromechanical sensors.

The STM32F401CCU6 microcontroller was chosen as the central computing module, and the MPU-6050 microelectromechanical sensor as a gyroscope and accelerometer.

The STM32F401CCU microcontroller is a microcontroller from STMicroelectronics' STM32F4 series, which is based on the ARM Cortex-M4 processor core. It has a clock speed about 84 MHz, 512 KB of Flash memory, and 96 KB of SRAM. But mostly STM32F401CCU is suitable for our purpose, because it provides several features:

- I2C bus for MEMS sensor connection;
- USB bus to transfer data to the computer;
- FPU(floating point unit) to speed up calculating mathematical formulas.

Additional advantages are its low cost and high availability.

The MPU-6050 is a 6-axis sensor that conveniently combines a gyroscope and an accelerometer in a small package, which makes it possible to use single physical connection and collect data over common I2C bus. MPU-6050 is also suitable to this project due to its low cost and ease of use [2].

As already mentioned, the sensor is connected to the microcontroller via the I2C bus (Fig 3), since the SDA and SCL lines are open drain, which means that they can only sink current, but not give it away, so 4.7 kΩ pull-up resistors are used.

The USB bus is used to transfer prepared data for further processing and visualization using virtual serial port technology (Fig 4), which represents UART (universal asynchronous receiver-transmitter) protocol over USB.

The circuit also receives power from the USB port when connected to a computer, the linear regulator lowers it from 5 volts to 3.3 volts, which is acceptable for the microcontroller and sensor.

Complementary filter. Another problem in the way of motion capture is the methods of processing data received from sensors, since in fact these data are noisy digitalized signals that are not related to each other in any way. To calculate the rotation angle, was used a complementary filter, which combines the gyroscope and accelerometer data together to obtain a more accurate estimate of the object's orientation. The gyroscope provides high frequency angular velocity measurements and the accelerometer provides low frequency gravity measurements [3]. To implement the complementary filter, the algorithm needs to calculate an orientation using the gyroscope data. This can be done using the trapezoid integration algorithm; the trapezoidal rule is a numerical integration method that can be used to estimate the value of an integral by approximating the area under a curve. In the context of integrating raw data from a gyroscope, the trapezoidal rule can be used to estimate the object's orientation over time:

$$\theta[N-1] = \theta[0] + \sum_{i=1}^{N-1} \left((w[i] + w[i-1]) \cdot \frac{\Delta t}{2} \right), \quad (1)$$

where $\theta[0]$ is the initial value of the angle (usually zero); $\theta[N-1]$ is the calculated end value of the angle; $w[0]$, $w[1]$, ..., $w[N-1]$ – sequence of N samples of angular velocity; Δt – time delta.

A simple representation of a given formula in a C programming language is shown in Listing 1.

```
#define SAMPLE_TIME 0.01
// Function to integrate the gyroscope data using
the trapezoidal rule
double integrate_data(double *gyro_data, int
num_samples)
{
    double area = 0.0;
    for (int i = 1; i < num_samples; i++) {
        double avg = (gyro_data[i] + gyro_data[i-
1]) / 2.0;
        area += avg * SAMPLE_TIME;
    }
    return area;
}
```

Listing 1. Estimation the object's orientation over time

Now it is important to process raw data from the accelerometer and get its orientation, it can be done with atan2 function. This function is a mathematical function that calculates the arctangent of two arguments, y and x, given as atan2(y, x). It returns the angle between the positive x-axis and the point (x, y) in the Cartesian plane, measured in radians (Fig 5):

$$\theta = \arctg(y/x). \quad (2)$$

Its representation in code is shown in Listing 2.

```
#define RAD_TO_DEG 57.295779513082
double pitch = atan2(accel_x_raw, accel_z_raw)
* RAD_TO_DEG;
```

Listing 2. Representation in formula code (2)

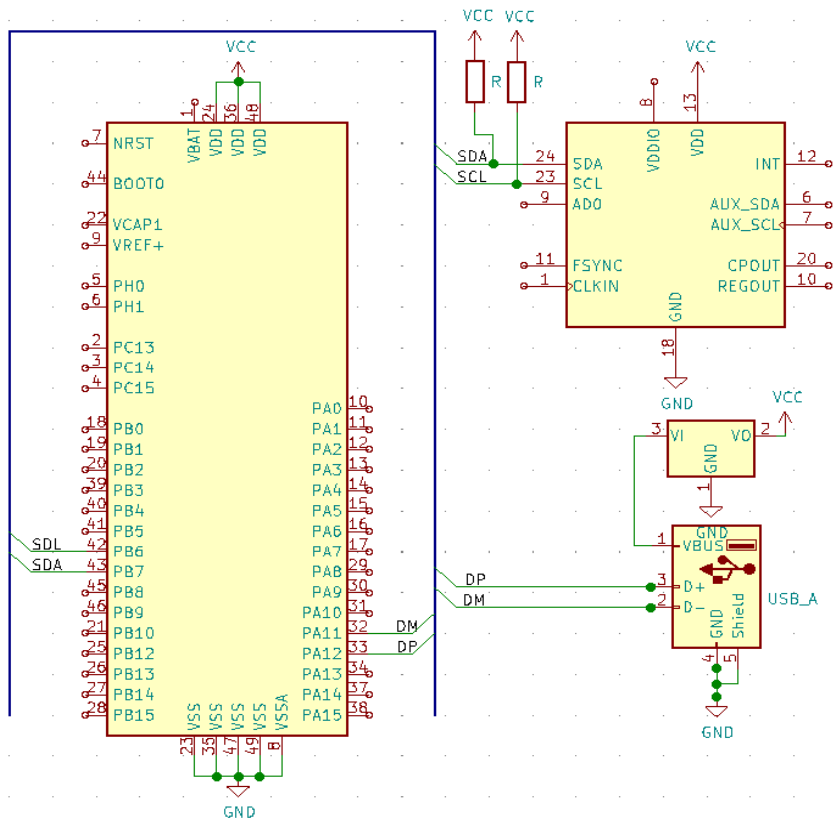


Fig. 3. Connection circuit of the microcontroller, sensor, and USB port

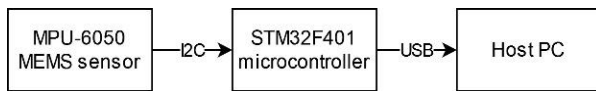


Fig. 4. Diagram of the data pipeline

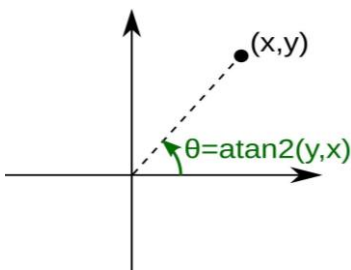


Fig. 5. Angle θ between the ray to the point (x, y) and the positive x -axis

Then it is possible to proceed to the calculation of the filter itself with this value. The filter is a weighted average of the two scores, where the weights are chosen to balance the accuracy and stability of the filter (Fig 6, $G(s)$ represents the transfer function for the low-pass filter, whereas $G^{-}(s)$ is the transfer function of the high-pass filter, such that $G(s) + G^{-}(s) = 1.$). The basic idea is to use the gyroscope value to correct for drift in the accelerometer estimate and use the accelerometer estimate to correct high frequency noise in the gyroscope estimate (Fig 7).

In the result of the filters work the estimate of the object's orientation will be more accurate [4].

On Fig 6 and 7:

$$orientation = (orientation_{gyroscope} * coef) + (orientation_{accelerometer} * (1 - coef)),$$

where $orientation_{gyroscope}$ is a gyroscope orientation, $orientation_{accelerometer}$ is an accelerometer orientation, $coef$ is a weighting factor from 0 to 1 that determines the balance between gyroscope and accelerometer estimates.

Red line is gyroscope values, blue is accelerometer values, green is a result.

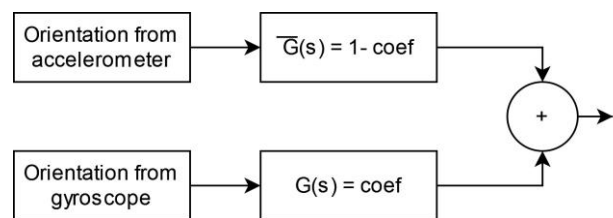


Fig. 6. Basic structure of Complementary Filter.

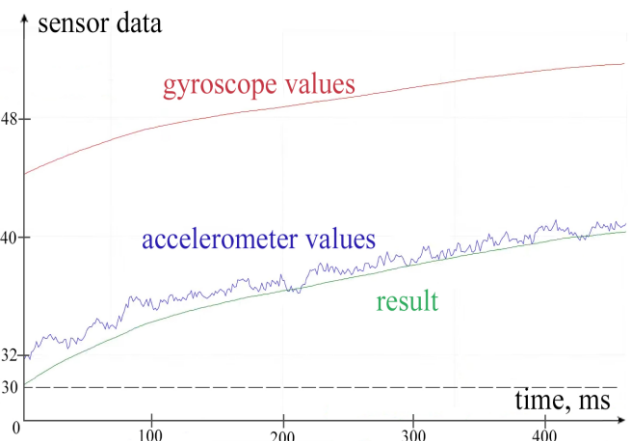


Fig. 7. Complementary filter result

Over time, the gyroscope values will begin to drift due to noise and bias in the sensor data. The complementary filter will detect this drift and slowly put more weight on the accelerometer estimate, resulting in a more accurate estimate of the object's orientation. In addition, this approach shows a good result when hitting the device (Fig. 8).

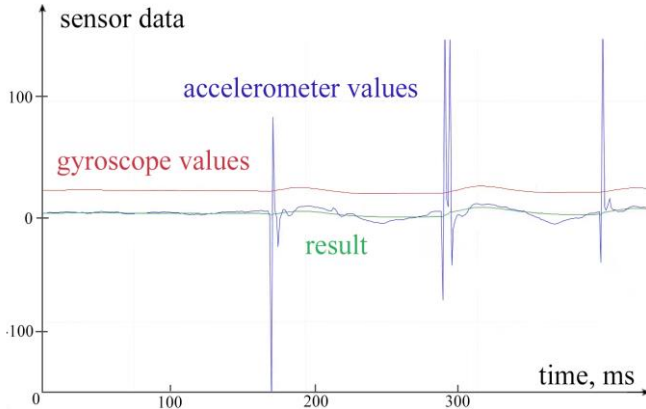


Fig. 8. Result of Complementary filter with hits

Main disadvantage. This method of calculation is poorly applicable for calculating the angle of rotation along the Z-axis, since its calculation will require the projections $accel_angle_x$ and $accel_angle_y$, and in this case, they are almost equal to zero.

When x and y are both zero, the point (x, y) is located at the origin of the Cartesian plane. At the origin, the direction of the vector (x, y) is undefined since it has no direction. Therefore, it is not possible to determine the angle between the positive x -axis and the point (x, y) using $atan2$ when x and y are both zero.

$$atan2(y, x) = \begin{cases} arctg(y/x), & \text{if } x > 0; \\ arctg(y/x) + \pi, & \text{if } x < 0 \ \& \ y \geq 0; \\ arctg(y/x) - \pi, & \text{if } x < 0 \ \& \ y < 0; \\ +\pi/2, & \text{if } x = 0 \ \& \ y > 0; \\ -\pi/2, & \text{if } x = 0 \ \& \ y < 0; \\ \text{undefined,} & \text{if } x = 0 \ \& \ y = 0; \end{cases} \quad (3)$$

So that is why, the calculation of this angle is not possible or will be performed with a large error. Calculation of the desired angle should be made only by the value of the gyroscope, without filtering (Fig. 9).

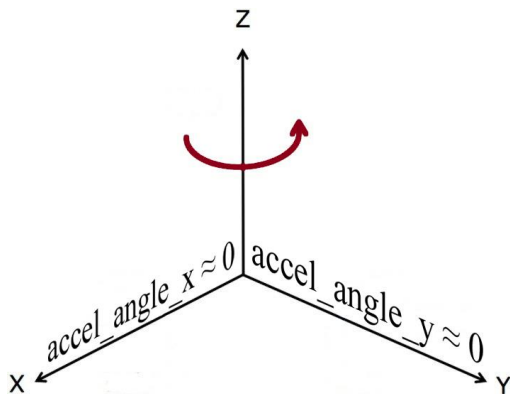


Fig. 9. Axes projection demonstration

In addition, to obtain a more accurate angle of rotation along the Z-axis, an alternative is to use an auxiliary magnetometer module - a digital compass.

Magnetometer module can provide the raw magnetic field data, which are measurements in the X, Y, and Z directions. By calculation and subtraction the magnetic declination angle from the inclination angle, the Z-axis rotation angle can be obtained.

Kalman filter. Another filter that can be reviewed and compared to complementary filter is Kalman filter. It is a mathematical algorithm that estimates the state of a system by combining noisy measurements and a model of the system dynamics. It is commonly used in many applications, including navigation, control systems, and signal processing.

The main idea of the Kalman filter is to recursively update an estimate of the state of the system based on two sources of information: measurements from sensors and predictions from a mathematical model of the system. The filter maintains two key components: a state estimate and a covariance matrix.

The state estimate is a vector that represents the best guess of the current state of the system, based on all available information up to the current time. The covariance matrix is a measure of the uncertainty or error in the state estimate [5].

The Kalman filter operates in two stages: the prediction stage and the update stage. In the prediction stage, the filter uses the mathematical model to predict the state of the system at the next time step, based on the current state estimate. This prediction is then used to calculate the covariance matrix for the predicted state [6].

In the update stage, the filter uses the actual measurement data to correct the predicted state estimate and covariance matrix. The Kalman filter calculates the Kalman gain, which is a weighting factor that determines how much to trust the predicted state estimate versus the measurement. The measurement is then used to update the state estimate and covariance matrix, which are then used in the next prediction stage.

The Kalman filter is designed to handle noisy or incomplete measurements, and it is able to incorporate new measurements as they become available, continuously refining its estimate of the system state. By using both the mathematical model and the actual measurements, the Kalman filter is able to provide a more accurate and robust estimate of the system state than either of these sources alone (Fig 10).

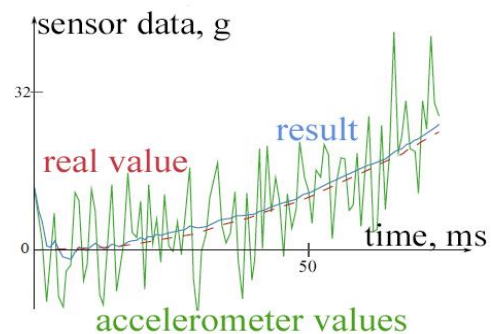


Fig. 10. Kalman filter result

Red line is original coordinate, green is sensor values, and blue Kalman filter is result.

Complementary filter and Kalman filter are both used for sensor fusion in order to improve the accuracy of the sensor readings. However, they differ in their underlying algorithms, complexity, and computational requirements (Table 1).

Complementary filter is a simple algorithm that combines the readings from an accelerometer and a gyroscope to estimate the orientation of an object. It is easy to implement and requires less computational power than a Kalman filter. However, it may not be as accurate as a Kalman filter, especially in the presence of noise or other sources of error.

Table 1. Comparison of method

	Complementary filter	Kalman filter
Algorithm complexity	Simple, easy to implement	More complex, requires more computational resources and expertise to implement
Accuracy	Low accuracy, particularly in noisy environments	High accuracy
Sensor fusion ability	Limited to combining data from two sensors, typically an accelerometer and a gyroscope	Can fuse data from multiple sensors, such as accelerometer, gyroscope, magnetometer, and GPS, allowing for more accurate tracking and compensation for various external factors
Real-time performance	Real-time processing is possible	Real-time processing is possible but can be limited by the complexity of the algorithm and the available computational resources
Robustness	Less robust to sudden changes and noise	More robust to sudden changes and noise due to its ability to estimate the noise and uncertainty in the system

Kalman filter, on the other hand, is a more complex algorithm that uses a mathematical model of the system and the sensor measurements to estimate the state of the system.

It is able to handle more complex systems and can incorporate multiple sensors and other sources of information to improve the accuracy of the estimate. However, it requires more computational power than a complementary filter and can be more challenging to implement.

In general, a complementary filter may be suitable for simple applications where accuracy is not so critical as performance, while a Kalman filter may be more appropriate for more complex systems where accuracy is critical and more computational power is available. It is also possible to use both filters in combination to take advantage of their strengths and improve overall accuracy.

DSP and DMP. It is also necessary to note several approaches that can speed up work with data and even increase the accuracy of the calculation result.

The first approach is to use a digital signal processor (DSP), which is a specialized microprocessor designed to process digital signals in real time. Unlike general-purpose microprocessors, DSPs are optimized for signal processing, which typically involves performing a lot of mathematical operations on data samples such as processing vectors or matrix.

Their special hardware functions can significantly speed up the processing of raw data from MEMS sensors, e.g. multiply floating point numbers in one cycle.

Also, DSPs usually have a specialized memory architecture that allows them to access data quickly and

efficiently, and moreover DSPs often have multiple processor cores that allow them to process different signals from gyroscope, accelerometer and magnetometer in parallel.

Another type of hardware accelerator is a digital motion processor (DMP), besides of DSP, motion processor are embedded in MEMS sensors and are able to fuse the data from these sensors to provide more accurate measurements of the device's orientation, motion, and position in space. Such component is often integrated into microcontrollers or microprocessors and used in motion detection applications such as smartphones, smart watches, and game controllers [7].

DMP is designed to offload some processing tasks from the main processor, such as sensor data combining and motion processing, and provides more accurate and reliable motion detection capabilities. In other words, DMP is already a full-fledged replacement for motion detection algorithms for MEMS sensors, which is implemented in hardware as a coprocessor.

Conclusions

For the tasks of processing "raw" data for tracking, the complementary filter and the Kalman filter are effective tools that provide a means to accurately estimate the state of the system based on noise measurements.

They are particularly useful for motion tracking systems that include multiple sensors, such as gyroscopes, accelerometers, and magnetometers, because they can efficiently combine measurements from these sensors to obtain a more accurate representation of the system's motion.

They can be used for a wide range of motion tracking applications, including gait analysis, sports performance monitoring, and rehabilitation [8].

However, the effectiveness of filters in motion tracking systems depends on careful parameter setting and model selection. The parameters of the Kalman filter must be chosen to balance accuracy and sensitivity, and the model used must accurately represent the dynamics of the monitored system. Overall, the choice between Complementary filter and Kalman filter depends on the

specific application and the available resources. Complementary filter is a simple and efficient algorithm that can be useful for applications where real-time performance [9] and simplicity are important, such as in simple robotic systems or basic motion tracking.

On the other hand, Kalman filter provides higher accuracy and more robustness to noise and sudden changes, making it more suitable for complex and high-precision applications, such as aerospace and autonomous vehicles.

REFERENCES

- Xu, J.Y., Nan, X., Ebken, V., Wang, Y., Pottie, G.J. and Kaiser, W.J. (2015), "Integrated inertial sensors and mobile computing for real-time cycling performance guidance via pedaling profile classification", *IEEE journal of biomedical and health informatics*, Vol. 19(2), pp. 440–445, doi: <https://doi.org/10.1109/JBHI.2014.2322871>.
- Chan, Y.J. and Huang, J.-W. (2015), "Multiple-point vibration testing with micro-electromechanical accelerometers and micro-controller unit", *Mechatronics*, Vol. 44, pp. 84-93, doi: <https://doi.org/10.1016/j.mechatronics.2017.04.006>.
- Shao X. and Si, H. (2022), "Low-frequency learning quantized control for MEMS gyroscopes accounting for full-state constraints", *Engineering Appl. of Artificial Intelligence*, 2022. No. 115, doi: <https://doi.org/10.1016/j.engappai.2022.104724>.
- Parag, N., Shashi, P. and Rahee, W. (2015), "Cascaded Complementary Filter Architecture for Sensor Fusion in Attitude Estimation", *Modern. technol. Honey*, No. 21, doi: <https://doi.org/10.3390/s21061937>.
- HanSung, K., Jeong, Y.P. and Chungkuk, J. (2023), "Real-time inverse estimation of multi-directional random waves from vessel-motion sensors using Kalman filter", *Ocean Engineering*, No. 280, doi: <https://doi.org/10.1016/j.oceaneng.2023.114501>.
- Strid, I. and Walentin, K. (2009), "Block Kalman Filtering for Large-Scale DSGE Models, *Computational Economics*, (Springer), available at: http://archive.riksbank.se/Upload/Dokument_riksbank/Kat_publicerat/WorkingPapers/2008/wp224ny.pdf.
- Sharma, M., Srivastava, R., Anand A., Prakash, D. and Kaligounder, L. (2017), "Wearable motion sensor based phasic analysis of tennis serve for performance feedback", *Acoustics, Speech and Signal Processing (ICASSP)*, 2017 IEEE International Conference, pp. 5945-5949, doi: <https://10.1109/ICASSP.2017.7953297>.
- Niznikowski, T., Sadowski, J. and Starosta, W. (2016), "Coordination Abilities in Physical Education, Sports and Rehabilitation, *Jozef Pilsudski University of Physical Education*, Warsaw. Faculty of Ph. and Sport, 2016. 323 p., available at: https://www.researchgate.net/publication/304581438coordination_abilities_in_physical_education_sports_and_rehabilitation.
- Chakravorti, N., Le Sage, T. and Slawson, S. E. (2013), "Design and implementation of an integrated performance monitoring tool for swimming to extract stroke information at real time", *IEEE Transactions on Human-Machine Systems*, Vol. 43(2), pp. 199–213, doi: <https://10.1109/TSMC.2012.2235428>.

Received (Надійшла) 16.03.2023

Accepted for publication (Прийнята до друку) 24.05.2023

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Дашков Дмитро Євгенович – студент, Харківський національний університет радіоелектроніки, Харків, Україна;

Dmytro Dashkov – student, Kharkiv National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: dmytro.dashkov@nure.ua; ORCID ID: <http://orcid.org/0009-0008-4137-5083>.

Ляшенко Олексій Сергійович – кандидат технічних наук, доцент, доцент кафедри електронних обчислювальних машин, Харківський національний університет радіоелектроніки, Харків, Україна;

Oleksii Liashenko – candidate of technical sciences (PhD), associate professor, associate professor of Electronic Computers department, National University of Radio Electronics, Kharkiv, Ukraine;

e-mail: oleksii.liashenko@nure.ua; ORCID ID: <http://orcid.org/0000-0002-0146-3934>.

Захоплення руху за допомогою датчиків MEMS

Д. Є. Дашков, О. С. Ляшенко

Анотація. Предметом дослідження даної статті є реєстрація та аналіз рухів людини за допомогою датчиків. У даній роботі представлено порівняння основних методів обробки даних інерційних мікромеханічних датчиків. Для збору даних було реалізовано пристрій, що фіксує рухи. У результаті пристрій використовує дані руху з акселерометра та гіроскопа для розрахунку траєкторії руху: кута повороту та прискорення. Дані зчитуються мікроконтролером, після чого фільтруються та обробляються одним із фільтрів (додатковим, Калмана) і, нарешті, передаються на комп'ютер для подальшого аналізу та відображення. **Метою статті** є порівняння кількох методів обробки даних з мікроелектромеханічних. **Отримані результати:** розроблено пристрій, отримані дані, які можна використовувати для характеристики методів та аналізу їх роботи в системі. **Висновки:** У ході дослідження було розроблено пристрій для збору та обробки даних від датчиків MEMS, який показав ефективність комплементарного фільтра порівняно з фільтром Калмана в системах реального часу з обмеженою обчислювальною потужністю. Реальні результати підтвердили, що результати додаткового методу з використанням менших обчислювальних ресурсів ненабагато поступаються дорожчому фільтру Калмана без використання додаткових датчиків, таких як цифровий компас.

Ключові слова: мікромеханічні датчики; мікроконтролер; гіроскоп; акселерометр; детектування руху; додатковий фільтр; фільтр Калмана.

Lev Raskin¹, Larysa Sukhomlyn², Dmytro Sokolov¹, Vitalii Vlasenko¹

¹ National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine

² Kremenchuk Mykhailo Ostrohradskyi National University, Kremenchuk, Ukraine

MULTI-CRITERIA EVALUATION OF THE MULTIFACTOR STOCHASTIC SYSTEMS EFFECTIVENESS

Annotation. Subject of the research in the article is the evaluation of complex multifactor stochastic systems functioning effectiveness according to a variety of correlated criteria. The problem actuality is determined by the fact that an independent evaluation of system effectiveness for each of the mutually correlated criteria for the system under consideration is not informative. In well-known works in the direction of multiple correlation, a relatively simple problem of estimating the correlation between one resulting parameter and a set of influencing factors is considered, which is not enough for the analysis and management of multicriteria systems. In addition, the known results do not take into account possible significant differences in influencing factors mutual correlation values. **Purpose of Work** is to develop a methodology for a comprehensive assessment of system effectiveness according to a variety of interrelated criteria. **Tasks to be Solved:** splitting the set of system parameters into two subsets (parameters determining the effectiveness of system and parameters affecting it), forming additive convolutions of parameters included in subsets, developing a methodology for calculating the multiple correlation coefficient between the components of the selected subsets, developing a method for differentiating a scalar function from a vector argument by this argument. **Applied Methods:** nonlinear programming, multidimensional correlation analysis, method of differentiation of scalar functions by vector argument. These methods are used for forming and calculating a multiple correlation coefficient between the set of system effectiveness complex criterion components values and its control parameters set values. **Results Obtained:** proposed methodology provides the possibility of solving the problems of system management, taking into account the revealed relationship between the multi-criteria evaluation of system effectiveness and values of its controlled parameters. At the same time, an important advantage of the obtained result lies in the possibility that arises when using it to take into account the joint (group) influence of control variables on the complex criterion of system efficiency. The developed technology of scalar functions differentiation by vector argument has great practical importance which expands the arsenal of computational mathematics.

Keywords: multifactorial stochastic system; complex efficiency criterion; vector differentiation procedure.

Introduction, publications analysis

One of the methods widely used in the study of multidimensional systems with interrelated parameters is the calculation of multiple correlations [1-3]. This method allows to estimate the measure of the relationship between one of the random variables and many others. The method is useful in evaluating a system of controlled parameters (indicators), as well as in solving numerous problems of evaluating the effectiveness of complex systems, optimization, management, rational allocation of resources, etc. At the same time, in particular, the task of establishing a link between the performance indicators of the system and the values of factors affecting it is of considerable interest. Typical similar problems are considered in [4-9]. Multiple correlation techniques significantly extend the arsenal of computational methods for systems analysis. However, the well-known methodological works [10-12] do not consider some fundamental features of a significant number of real stochastic systems. In these works, the possibilities of applying multiple correlation are limited to considering the special case when the system is single-criteria. The canonical works [13-15] in this direction do not take into account the possible significant differences between the correlation relationships of the factors that differently affect the efficiency of the system. Finding the analytical relationships that define these relationships is usually a difficult task. A possible direction of its solution is to identify statistical descriptions of correlations obtained using appropriate technologies for processing real observations of the system functioning process.

Materials and methods

We introduce $x_i = (x_{i1}, x_{i2}, \dots, x_{in_i})$ – a set of values of the system's performance indicators, and $x_2 = (x_{21}, x_{22}, \dots, x_{2n_2})$ – a set of values of factors that presumably affect the system's efficiency. We assume that the random values of the vector components x_1 and x_2 are pre-centered, that is

$$M[x_{1i_1}] = M[x_{2i_2}] = 0, \quad i_1 = \overline{1, n_1}, \quad i_2 = \overline{1, n_2}. \quad (1)$$

We introduce a covariance matrix K for the entire set of observed random variables, which we represent blockwise as follows:

$$K = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix},$$

where k_{11} is the covariance matrix for the random components of the vector x_1 ,

$$k_{11} = M\left[\left(x_{1s_1} - M[x_{1s_1}]\right)\left(x_{1s_2} - M[x_{1s_2}]\right)\right],$$

$$\dim k_{11} = n_1 \times n_1;$$

k_{12} – covariance matrix for components of vectors x_1 and x_2 ,

$$k_{12} = M\left[\left(x_{1s_1} - M[x_{1s_1}]\right)\left(x_{2s_2} - M[x_{2s_2}]\right)\right],$$

$$\dim k_{12} = n_1 \times n_2;$$

$$k_{21} = k_{12}^T, \dim k_{21} = n_2 \times n_1;$$

k_{22} – covariance matrix for vector components x_2 ,

$$k_{22} = M\left[\left(x_{2s_1} - M[x_{2s_1}]\right)\left(x_{2s_2} - M[x_{2s_2}]\right)\right],$$

$$\dim k_{22} = n_2 \times n_2.$$

Now, using column vectors of weighting coefficients

$$A_1 = \begin{pmatrix} a_{11} \\ a_{12} \\ \dots \\ a_{1n_1} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{21} \\ a_{22} \\ \dots \\ a_{2n_2} \end{pmatrix}$$

for vectors x_1 and x_2 we introduce additive convolution

$$U = A_1^T x_1 = \begin{pmatrix} a_{11}a_{12}\dots a_{1n_1} \\ x_{11} \\ x_{12} \\ \dots \\ x_{1n_1} \end{pmatrix},$$

$$V = A_2^T x_2 = \begin{pmatrix} a_{21}a_{22}\dots a_{2n_2} \\ x_{21} \\ x_{22} \\ \dots \\ x_{2n_2} \end{pmatrix}. \quad (2)$$

At the same time, taking into account (1)

$$M[U] = M[V] = 0,$$

so $M[A_1^T x_1] = M[A_2^T x_2] = 0.$

Next, we determine the correlation coefficient between U and V :

$$K_{UV} = M[A_1^T x_1 x_2^T A_2] = A_1^T k_{12} A_2. \quad (3)$$

We will choose the vectors A_1 and A_2 so that the sums of the weighting coefficients used in the formation of random vectors U and V are equal to one, that is

$$A_1^T I_1 = A_2^T I_2 = 1, \quad (4)$$

$$I_1^T = (11\dots 1), \quad \dim I_1^T = 1 \times n_1,$$

$$I_2^T = (11\dots 1), \quad \dim I_2^T = 1 \times n_2.$$

Now set the task of finding vectors A_1 and A_2 maximizing (3) and satisfying constraints (4).

The resulting mathematical programming problem will be solved by the method of indefinite Lagrange multipliers [8]. The Lagrange function has the form

$$F(A_1, A_2) = A_1^T k_{12} A_2 - \lambda_1 (A_1^T I_1 - 1) - \lambda_2 (A_2^T I_2 - 1). \quad (5)$$

We differentiate the scalar function (5) by vector arguments A_1, A_2 and equate the results to zero. In this case, the operation of differentiating a scalar function from a vector argument by this argument is performed as follows.

Let $F(x)$ be a scalar function of the vector x . We introduce an auxiliary function of the scalar argument t

$$F(x, t) = F(x + ht), \quad \dim h = \dim x.$$

It is clear that

$$F(x + ht) \Big|_{t=0} = F(x).$$

Now we calculate

$$\begin{aligned} \frac{dF(x, t)}{dt} \Big|_{t=0} &= \frac{dF(x + ht)}{dt} \Big|_{t=0} = \\ &= \frac{dF(x + ht)}{d(x + ht)} \cdot \frac{d(x + ht)}{dt} \Big|_{t=0} = \frac{dF(x)}{dx} \cdot h. \end{aligned} \quad (6)$$

Thus, it turned out that the desired derivative $\frac{dF(x)}{dx}$ is a linear operator acting on a vector h in relation (6).

Let's illustrate the technique of vector differentiation with extremely simple example.

Example

Find a set $x = (x_1, x_2)$ that minimizes the function

$$F(x_1, x_2) = x_1^2 + 2x_2^2$$

and satisfies the constraint $x_1 + x_2 = 1.$

We can solve this problem in the usual way using the method of indefinite Lagrange multipliers. The Lagrange function has the form of

$$F(x_1, x_2) = x_1^2 + 2x_2^2 - \lambda(x_1 + x_2 - 1).$$

We differentiate $F(x_1, x_2)$ by x_1, x_2 and equate the results to zero.

$$\frac{dF(x_1, x_2)}{dx_1} = 2x_1 - \lambda = 0,$$

$$\frac{dF(x_1, x_2)}{dx_2} = 4x_2 - \lambda = 0.$$

From here we express x_1 and x_2 through λ . We have an

$$x_1 = \frac{\lambda}{2}, \quad x_2 = \frac{\lambda}{4}.$$

We will find indefinite multiplier λ using the constraint.

$$x_1 + x_2 = \frac{\lambda}{2} + \frac{\lambda}{4} = \frac{3}{4}\lambda = 1,$$

where from $\lambda = 4/3.$

Then

$$x_1 = \frac{\lambda}{2} = \frac{2}{3}, \quad x_2 = \frac{\lambda}{4} = \frac{1}{3}.$$

Solution has been received. We will now solve this problem using the vector differentiation procedure. According to the task condition, let's introduce a vector $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ and a unit vector $I = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Now the formulation of the problem in terms of matrix theory has the form: find a vector x that minimizes the function $F(x) = x^T A x$ and satisfies the constraint $x^T I = 1$. Use the method of indefinite Lagrange multipliers.

The Lagrange function has the form

$$\phi(x) = x^T A x - \lambda (x^T I - 1).$$

In accordance with the introduced technology of vector differentiation, we form an auxiliary function

$$\begin{aligned} \phi(x+ht) &= \\ &= (x+ht)^T A (x+ht) - \lambda [(x+ht)^T I - 1] = \\ &= (x^T + h^T t) A (x+ht) - \lambda [(x^T + h^T t) \cdot I - 1] = \\ &= (x^T A + h^T A t)(x+ht) - \lambda (x^T I + h^T I t - 1) = \\ &= x^T A x + x^T A h t + h^T A x t + h^T A h t^2 - \\ &\quad - \lambda (x^T I - h^T I t - 1). \end{aligned}$$

We will find

$$\begin{aligned} \left. \frac{d\phi(x+ht)}{dt} \right|_{t=0} &= x^T A h + h^T A x - \lambda h^T I = \\ &= x^T A h + x^T A h - \lambda I^T h = (2x^T A - \lambda I^T) h. \end{aligned}$$

Hence, according to (6),

$$\frac{dF(x)}{dx} = 2x^T A - \lambda I^T = 0, \quad x^T A = \frac{1}{2} \lambda I^T.$$

We will find the unknown multiplier λ using the constraint. We have

$$x^T = \frac{1}{2} \lambda I^T A^{-1},$$

$$x^T I = \frac{1}{2} \lambda I^T A^{-1} I = 1,$$

where from

$$\frac{1}{2} \lambda = \frac{1}{I^T A^{-1} I}.$$

Then

$$x^T = \frac{I^T A^{-1}}{I^T A^{-1} I}. \quad (7)$$

Since

$$A^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix},$$

that

$$\begin{aligned} x^T &= \frac{(1 \ 1) \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}}{(1 \ 1) \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}} = \frac{\begin{pmatrix} 1 & \frac{1}{2} \end{pmatrix}}{\begin{pmatrix} 1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}} = \\ &= \frac{2}{3} \begin{pmatrix} 1 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \end{pmatrix}, \end{aligned}$$

The resulting solution, of course, coincides with the previous one.

The obvious usefulness of the introduced vector differentiation operation is as follows. The computational scheme of the resulting matrix representation using vector differentiation technology, the final ratio (7) does not depend on the problem dimension, while the complexity of the standard solution increases with increasing vector x dimension.

Let's go back to the original problem. Let's perform the introduced operation (6) on the function (5) sequentially to calculate derivatives by vector arguments A_1 and A_2 .

In accordance with (5) we introduce

$$\begin{aligned} F [(A_1 + ht, A_2)] &= \\ &= (A_1 + ht)^T k_{12} A_2 - \lambda_1 [(A_1 + ht)^T I_1 - 1] - \\ &\quad - \lambda_2 (A_2^T I_2 - 1) = \\ &= (A_1^T + h^T t) k_{12} A_2 - \lambda_1 [(A_1^T I_1 + h^T I_1) - 1] - \\ &\quad - \lambda_2 (A_2^T I_2 - 1) = \\ &= (A_1^T k_{12} A_2 + h^T k_{12} A_2 t) - \\ &\quad - \lambda_1 (A_1^T I_1 + h^T I_1 - 1) - \lambda_2 (A_2^T I_2 - 1). \end{aligned}$$

Now

$$\left. \frac{dF [(A_1 + ht)^T]}{dt} \right|_{t=0} = h^T k_{12} A_2 - \lambda_1 h^T I_1. \quad (7)$$

Relation (7) describes a scalar. Therefore

$$\begin{aligned} h^T k_{12} A_2 - \lambda_1 h^T I_1 &= (h^T k_{12} A_2 - \lambda_1 h^T I_1)^T = \\ &= A_2^T k_{12}^T h - \lambda_1 I_1^T h = (A_2^T k_{12}^T - \lambda_1 I_1^T) h. \end{aligned}$$

Hence, in accordance with (6), we have

$$\frac{dF(A_1, A_2)}{dA_1} = A_2^T k_{12}^T - \lambda_1 I_1^T. \quad (8)$$

Equate (8) to zero and express A_2^T by λ_1 . We have

$$A_2^T k_{12}^T - \lambda_1 I_1^T = 0. \tag{9}$$

To find, A_2^T multiply (9) on the right by K_{12} . We will get

$$A_2^T k_{12}^T k_{12} - \lambda_1 I_1^T k_{12} = 0.$$

From here

$$A_2^T = \lambda_1 I_1^T k_{12} (k_{12}^T k_{12})^{-1}.$$

We will find an unknown indefinite multiplier λ_1 using constraint (4). At the same time

$$A_2^T I_2 = \lambda_1 I_1^T k_{12} (k_{12}^T k_{12})^{-1} I_2 = 1.$$

From here

$$\lambda_1 = \frac{1}{I_1 k_{12} (k_{12}^T k_{12})^{-1} I_2}.$$

Then

$$A_2^T = \frac{1}{I_1 k_{12} (k_{12}^T k_{12})^{-1} I_2} \cdot I_1^T k_{12} (k_{12}^T k_{12})^{-1}. \tag{10}$$

Similarly to the previous one, we define A_1 . We will perform the necessary operations without detailed additions.

$$\begin{aligned} F &= \left[(A_2 + ht)^T, A_1 \right] = (A_2 + ht)^T k_{21} A_1 - \\ &- \lambda_1 (A_1^T I_1 - 1) - \lambda_2 \left[(A_2 + ht)^T I_2 - 1 \right] = \\ &= (A_2^T + h^T t) k_{21} A_1 - \lambda_1 (A_1^T I_1 - 1) - \\ &- \lambda_2 (A_2^T I_2 + h^T I_2 t - 1) = A_2^T k_{21} A_1 + h^T k_{21} A_1 t - \\ &- \lambda_1 (A_1^T I_1 - 1) - \lambda_2 (A_2^T I_2 + h^T I_2 t - 1). \end{aligned} \tag{11}$$

Let's perform differentiation (11) by A_2 .

$$\left. \frac{dF \left[(A_2^T + ht), A_1 \right]}{dt} \right|_{t=0} = h^T k_{21} A_1 - \lambda_2 h^T I_2 =$$

$$\begin{aligned} &= (h^T k_{21} A_1 - \lambda_2 h^T I_2)^T = \\ &= A_1^T k_{21}^T h - \lambda_2 I_2^T h = \\ &= (A_1^T k_{21}^T - \lambda_2 I_2^T) h. \end{aligned}$$

Then

$$\frac{dF(A_1, A_2)}{dA_2} = A_1^T k_{21}^T - \lambda_2 I_2^T = 0. \tag{12}$$

Further

$$A_1^T k_{21}^T k_{21} - \lambda_2 I_2^T k_{21} = 0.$$

From here

$$A_1^T = \lambda_2 I_2^T k_{21} (k_{21}^T k_{21})^{-1}. \tag{13}$$

We will find λ_2

$$A_1^T I_1 = \lambda_2 I_2^T k_{21} (k_{21}^T k_{21})^{-1} I_1 = 1.$$

Then

$$\lambda_2 = \frac{1}{I_2^T k_{21} (k_{21}^T k_{21})^{-1} I_1}. \tag{14}$$

From where, substituting (14) into (13), we get

$$\begin{aligned} A_1^T &= \frac{1}{I_2^T k_{21} (k_{21}^T k_{21})^{-1} I_1} \times \\ &\times I_2^T k_{21} (k_{21}^T k_{21})^{-1}. \end{aligned} \tag{15}$$

Problem solution is completed.

Conclusions

1. The obtained ratios (10) and (15) enable rational management in a complex system. The proposed approach allows us to determine to what extent the values of a set of control variables affect the values of controlled variables that determine the system efficiency.

2. These relations determine a new quality of assessment of the joint (group) influence of some variables on others.

REFERENCES

1. Cramér H. (1999), *Mathematical Methods of Statistics*, Princeton Mathematical Series, Vol. 9, 575 p., available at: <https://archive.org/details/in.ernet.dli.2015.149716>.
2. Gmurman V.E. (1977), *Fundamentals of Probability Theory and Mathematical Statistics*, First Thus Edition, 479 p., available at: <https://www.amazon.com/Fundamentals-Probability-Theory-Mathematical-Statistics/dp/0592039315>.
3. Raskin, L., Sira, O., & Ivanchykhin, Y. (2017), "Models and methods of regression analysis under conditions of fuzzy initial data", *Eastern-European Journal of Enterprise Technologies*, Vol. 4(4 (88)), pp. 12–19, doi: <https://doi.org/10.15587/1729-4061.2017.107536>
4. Gill, F., Murray, W. and Wright, M. (1997), *Practical optimization*, Academic Press, San Diego, CA 92101, 402 p., available at: <http://gaznevada.iq.usp.br/wp-content/uploads/livros/gill.pdf>.
5. Kullback, S. and Leibler, R.A. (1951), "On information and Sufficiency", *Ann. Math. Statist*, Vol. 22, No. 1, pp. 79–86, available at: <https://www.jstor.org/stable/2236703>
6. Hemdy, A. (2007), *Operations Research*, Williams, 912 p., available at: <https://pdfgoes.com/download/2835726-Operations%20Research%20Hamdy%20Taha%20Solutions%20Manual.pdf>.

7. Raskin, L.G. and Pustovoitov, P.E. (2002), "Solution of a multi-product inventory management problem by a probabilistic criterion", *Bulletin of NTU KhPI, Series "System Analysis, Management, Information Technologies"*, No. 13, pp. 49-53.
8. Raskin, L., Sira, O. and Ivanchikhin, Y. (2017), "Modes and methods of regression analysis under conditions of fuzzy initial data", *Eastern-European Journal of Enterprise Technologies*, Vol. 1, No. 4 (75), pp. 12-19, doi: <https://doi.org/10.15587/1729-4061.2017.127034>.
9. Feller, W. (1970), *An introduction to probability theory and applications*, J. Wiley, 527 p., available at: <https://bitcoinwords.github.io/assets/papers/an-introduction-to-probability-theory-and-its-applications.pdf>.
10. Semenov, S., Sira, O., Gavrylenko, S. and Kuchuk, N. (2019), "Identification of the state of an object under conditions of fuzzy input data", *Eastern-European Journal of Enterprise Technologies*, Vol 1, No 4 (97), pp. 22-30, doi: <https://doi.org/10.15587/1729-4061.2019.157085>.
11. Reilly, O. (2017), *Practical Statistics for Data Scientist*, O. Reilly Media Inc., 286 p., available at: <https://www.oreilly.com/library/view/practical-statistics-for/9781491952955>.
12. Cramer, H. (1999), *Mathematical methods of statistics*, Princeton University Press, 575 p., Stockholm, 647 p., available at: https://books.google.com.ua/books/about/Mathematical_Methods_of_Statistics.html?id=CRKKaJ00DYC&redir_esc=y.

Received (Надійшла) 16.02.2023

Accepted for publication (Прийнята до друку) 26.04.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Раскін Лев Григорович – доктор технічних наук, професор, професор кафедри мультимедійних інформаційних технологій і систем, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Lev Raskin – Doctor of Technical Sciences, Professor, Professor in the Department of Multimedia Information Technology and Systems, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: topology@ukr.net; ORCID: <http://orcid.org/0000-0002-9015-4016>.

Сухомлин Лариса Вадимівна – кандидат технічних наук, доцент, доцент кафедри менеджменту, Кременчуцький національний університет імені Михайла Остроградського, Кременчук, Україна;
Larysa Sukhomlyn – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Management, Kremenchuk Mikhail Ostrogradskiy National University, Kremenchuk, Ukraine;
e-mail: lar.sukhomlyn@gmail.com; ORCID: <https://orcid.org/0000-0001-9511-5932>.

Соколов Дмитро Дмитрович – аспірант кафедри мультимедійних інформаційних технологій і систем, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Dmytro Sokolov – postgraduate student in the Department of Multimedia Information Technology and Systems, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: sokolovddd@gmail.com; ORCID: <https://orcid.org/0000-0002-4558-9598>.

Власенко Віталій Владиславович – магістр кафедри мультимедійних інформаційних технологій і систем, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Vitalii Vlasenko – master of the Department of Multimedia Information Technologies and Systems, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine;
e-mail: vitalik.vlasenko.000@gmail.com; ORCID: <https://orcid.org/0000-0001-5427-0223>.

Багатокритеріальна оцінка ефективності багатofакторних стохастичних систем

Л. Г. Раскін, Л. В. Сухомлин, Д. Д. Соколов, В. В. Власенко

Анотація. Предметом дослідження у статті є оцінка ефективності функціонування складних багатofакторних стохастичних систем за багатьма корельованими критеріями. **Актуальність проблеми** визначається тим, що незалежна оцінка ефективності системи по кожному з взаємно корельованих критеріїв для системи, що розглядається, малоінформативна. У відомих роботах за напрямом множинна кореляція розглядається відносно просте завдання оцінки кореляційного зв'язку між одним результуючим параметром і набором факторів, що впливають, чого задачам аналізу та управління багатокритеріальних систем недостатньо. Крім того, відомі результати не враховують можливі суттєві відмінності значень взаємної кореляції факторів, що впливають. **Мета роботи** – розробка методики комплексної оцінки ефективності системи з безлічі взаємозалежних критеріїв. Розв'язування задач: розбиття безлічі параметрів системи на два підмножини (параметри, що визначають ефективність системи, і параметри, що впливають на неї), формування адитивних згорток параметрів, що входять до підмножини, розробка методики розрахунку коефіцієнта множинної кореляції між компонентами виділених підмножин, розробка методу від векторного аргументу з цього аргументу. **Методи, що застосовуються:** нелінійне програмування, багатомірний кореляційний аналіз, метод диференціювання скалярних функцій за векторним аргументом. Ці методи використовуються для формування та розрахунку множинного коефіцієнта кореляції між сукупністю значень компонентів комплексного критерію ефективності функціонування системи та значеннями набору її керуючих параметрів. **Отримані результати:** запропонована методика забезпечує можливість вирішення завдань управління системою з урахуванням взаємозв'язку, що виявляється між багатокритеріальною оцінкою ефективності системи і значеннями керування її параметрів. При цьому важлива перевага отриманого результату полягає в можливості, що виникає при його використанні, обліку спільного (групового) впливу управляючих змінних на комплексний критерій ефективності системи. Велике практичне значення має розроблена технологія диференціювання скалярних функцій за векторним аргументом, що розширює арсенал обчислювальної математики.

Ключові слова: багатofакторна стохастична система; комплексний критерій ефективності; процедура векторного диференціювання.

Intelligent information systems

UDC 004.7

doi: <https://doi.org/10.20998/2522-9052.2023.2.10>Dun Bao^{1,2}, Oleksandr Zakovorotnyi¹, Nina Kuchuk¹¹National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine²Bank of China, Beijing, China

GENERATING CURRENCY EXCHANGE RATE DATA BASED ON QUANT-GAN MODEL

Abstract. The aim of the research. This paper discusses the use of machine learning algorithms to generate data that meets the demands of academia and industry in the context of exchange rate fluctuations. **Research results.** The paper builds a Quant-GAN model using temporal convolutional neural networks (CNN) and trains it on end-of-day and intraday high-frequency rates of currency pairs in the global market. The generated data is evaluated using various statistical methods and is found to effectively simulate the real dataset. Experimental results show that data generated by the model effectively fits statistical characteristics and typical facts of real training datasets with good overall fit. The results provide effective means for global FX market participants to carry out various tasks such as stress tests and scenario simulations. **Future work** includes accumulating data and increasing computing power, optimizing and improving GAN models, and establishing evaluation standards for generating exchange rate price data. As computing power continues to grow, the GAN model's ability to process ultra-large-scale datasets is expected to improve.

Keywords: convolutional neural network; generative adversarial nets, foreign exchange rate.

Introduction

The global foreign exchange market is one of the most active financial markets. Exchange rate fluctuations have a direct impact on market participants and can cause potential risks to their economic status. However, relying solely on historical data for research is not enough. It has become necessary to use methods to generate data that simulates market fluctuations for theoretical and practical verification. In this context, economists and financial traders have adopted various methods to generate financial transaction simulation data. The best results are obtained using artificial intelligence methods [1–6]. The Quant-GAN studied in this paper is an emerging artificial intelligence tool that uses machine learning to generate time-series data of trading prices in financial markets.

1 Generative Adversarial Networks

Generative Adversarial Networks (GAN) is a deep generative model machine learning algorithm that generates realistic data by training a neural network. Goodfellow et al. designed a game-like competition where the generator learns to simulate data that is similar to the real data distribution [7]. The specific algorithm of GAN consists of the following components:

Generator G: Simulates data (images, texts, sounds, time series, etc.) based on certain rules.

Discriminator D: The discriminator D is a classifier that judges (or provides a probability) whether the input data is from the real dataset.

Network training: First, the generator is fixed, and the discriminator is trained using a batch of mixed data that contains both real samples and data generated by the generator. The binary cross-entropy function is used as the loss function, and the discriminator can distinguish between real and fake data after updating the gradient through backpropagation. Second, the discriminator is fixed, and the generator is trained by

generating data from random noise inputs. As the discriminator has already been trained in the previous stage, it can identify the authenticity of the input data, and the generator's ability is improved by updating its weight through backpropagation.

Repeat this process so that the generator and the discriminator compete with each other until they reach a Nash equilibrium, completing the training of the entire GAN algorithm. At this point, the data generated by the generator is realistic enough in its distribution, and the discriminator cannot determine whether the input data is real or fake (output a probability of 50%). The optimization problem can be expressed as the solution of Equation:

$$\begin{aligned} \min_G \max_D V(D, G) &= \\ &= \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \\ &+ \mathbb{E}_{z \sim p_z(z)} \left[\log \left(1 - D(G(z)) \right) \right]. \end{aligned}$$

It can be proved that in the function space $D(x; \theta_d)$ and $G(z; \theta_g)$, there is a unique solution that makes G reproduce the training data distribution, at this time $D(x) = 0.5$.

2 Quant-GAN Method

Wiese et al. (2020) proposed the Quantitative Generative Adversarial Network (Quant-GAN) model [8], using the S&P 500 index from 2008 to 2018 as a training set to simulate and generate stock index sequence data. Compared to traditional GARCH, ARIMA, and other machine learning methods, it achieved better results.

Quant-GAN uses seven layers of the Temporal Convolutional Networks (CNN) module, based on the Dilated Causal Convolutional Networks architecture [9] proposed by Bai et al. (2018). Each temporal block consists of two one-dimensional expansion causal convolutional layers and two PReLU layers as the

activation function. The main feature of CNN is the use of dilated convolution, where the distance moved by the convolution kernel is adjusted by a hyperparameter in each iteration, thus increasing the receptive field without increasing the size of the convolution kernel. At the same time, a skip-connection method for both the generator and discriminator's is also introduced to avoid the gradient disappearance problem. Quant-GAN uses the Wasserstein distance as the loss function, enabling the generated logarithmic return random process to achieve a risk-neutral distribution.

Network training: In the first stage, the generator is fixed, and the discriminator is trained. A batch of real samples and the mixed data generated by the generator are used as inputs. The Wasserstein distance is used as the loss function, and the discriminator updates the gradient through backpropagation to identify real and virtual data. In the second stage, the discriminator trains the generator once every 5 iterations. The discriminator discriminates the generator's randomly inputted data and updates the generator's weights through backpropagation. Repeat the above process for a certain number of times, and the generator and the discriminator compete with each other until they reach the Nash equilibrium, thus completing the training of the entire Quant-GAN. At this point, the generator can generate realistic financial time series data, and the discriminator cannot distinguish between the input data's source whether it is from a real dataset or generated by a generator (with an output probability of 50%).

Data preprocessing: Step 1: Convert the absolute value of the series into logarithmic rate of return; Steps 2 and 4: Standardize (normalize) the log return to a standard normal distribution with mean 0 and variance 1; Step 3: Apply the Inverse Lambert W function transformation to the data; Step 5: For the receptive field T of the discriminator, use a sliding window of the corresponding length to preprocess the logarithmic return sequence.

3 Exchange Rate Data Generation

3.1 Selection, Processing of Data Sets

The foreign exchange transaction dataset used to train the Quant-GAN model comes from the Dukascopy trading platform. The descriptive statistics of the unprocessed original dataset are shown in Table 1.

Table 1 – Descriptive statistics of preprocessed exchange rate time series data

currency pair	AUDUSD
date	2019.5.6
type	intraday
amount	4120
maximum value	9.70985
minimum value	9.04364
range	18.75349
average	0
median	0.02582
standard deviation	1
Skewness	0.58207
kurtosis	9.28550

Finally, after steps 1 and 2, the dataset is divided into a training set and a test set according to the ratio of 9:1. used to train the model, and to calculate the fitting error.

3.2 Stylized Facts for Time Series Data in Financial Markets

Aiming at the characteristics of typical financial market time series data represented by stock prices and foreign exchange rates, Cont (2001) summarized some common stylized facts revealed by many studies [10].

Linear Autocorrelation of returns is usually not significant (except for very small intraday time scales), that is, the autocorrelation function converges to near 0 very quickly.

Heavy-tailed distribution refers to the distribution of the return series showing a power-law distribution or Pareto distribution tail characteristics. Leverage effect refers to the price reacting differently to positives and negatives.

Usually, the negative price movement due to bad news is greater. To test the leverage effect, Nelson (1991) used the EGARCH model to estimate the standard deviation of price fluctuations. Asymptotic normality, also known as Aggregational Gaussianity, refers to the phenomenon where the distribution of returns gradually approaches a normal distribution as the time scale increases. Volatility clustering refers to the positive autocorrelation of different volatility measures of a price series within a few time periods. This can be measured with the absolute value correlation coefficient of the return rate.

3.3 Results and Model Evaluation

3.3.1 Data features and model evaluation methods. The exchange rate data generated by the generator after rounds of training is compared with the real dataset. The main features of the recorded data include mean, standard deviation, minimum, quantiles, maximum, skewness and kurtosis. Some typical facts of financial market time series data introduced in the previous section such as autocorrelation, heavy-tailed distribution, leverage effect, aggregational Gaussianity and volatility clustering were tested using both real and generated datasets.

M. Heusel et al. (2017) proposed using Inception Score (IS) and Fréchet Inception Distance (FID) as indicators [11] to evaluate the performance of GAN models. A lower score indicates a more realistic image. According to Xu et al. (2020), this paper trains a KNN classifier to evaluate the performance of the GAN model. When the model is well-trained, the average scores of KNN classifiers calculated using different k values are all around 0.5.

3.3.2 AUDUSD intraday high frequency data. Using the AUDUSD intraday (2019.05.06) high-frequency data, the Quant-GAN model was trained for 500 rounds. The specific parameters are shown in Table 2: Major adjustments include increasing the number of stacked layers in the CNN module to 11 layers in both generator and discriminator and increasing batch size to 2048; increasing stacked layers in CNN module for

daytime data to nine layers and increasing batch size to 512. In addition, due to the small absolute value, the learning rate was increased to 0.0002 and gradient clipping value was decreased.

After the model training is completed, random Gaussian noise is used to generate exchange rate logarithmic return data, and the generated data is used to simulate the exchange rate fluctuation path trend (Fig. 1).

Table 2 – Model hyperparameters

gradient clipping	0.01
learning rate	0.0002
training rounds	500
batch size	32
sliding window	2048
CNN layers	11

With the increase of training batches, the generator loss stabilizes at 0 and the discriminator loss stabilizes at -0.5 (Fig. 2). The training set and the generated data FID score is stably approaching 0, the test set FID score is approaching the training set (Fig. 2), and the mixed data set KNN classifier discriminant result of generated data and real data is very close to 0.5 (Table 3).

Table 3 – KNN classifier results

generator accuracy	0.49891920
--------------------	------------

The data distribution generated by the model is shown in Fig. 3, and the statistics are shown in Table 4. Compared with the training data set, various statistics such as quantile, mean, standard deviation, skewness, and kurtosis can be effectively fitted.

Table 4 – Descriptive statistics for real datasets and model-generated data

	real	fake
count	54487	54487
mean	0.00000006	0.00000013
std dev	0.00002522	0.00002310
min	-0.00057339	-0.00053218
25%	-0.00001430	-0.00001037
50%	0.00000000	0.00000023
75%	0.00001430	0.00001068
max	0.00060206	0.00039266
skew	-0.34678190	-1.43902695
kurtosis	59.04555456	37.37598038

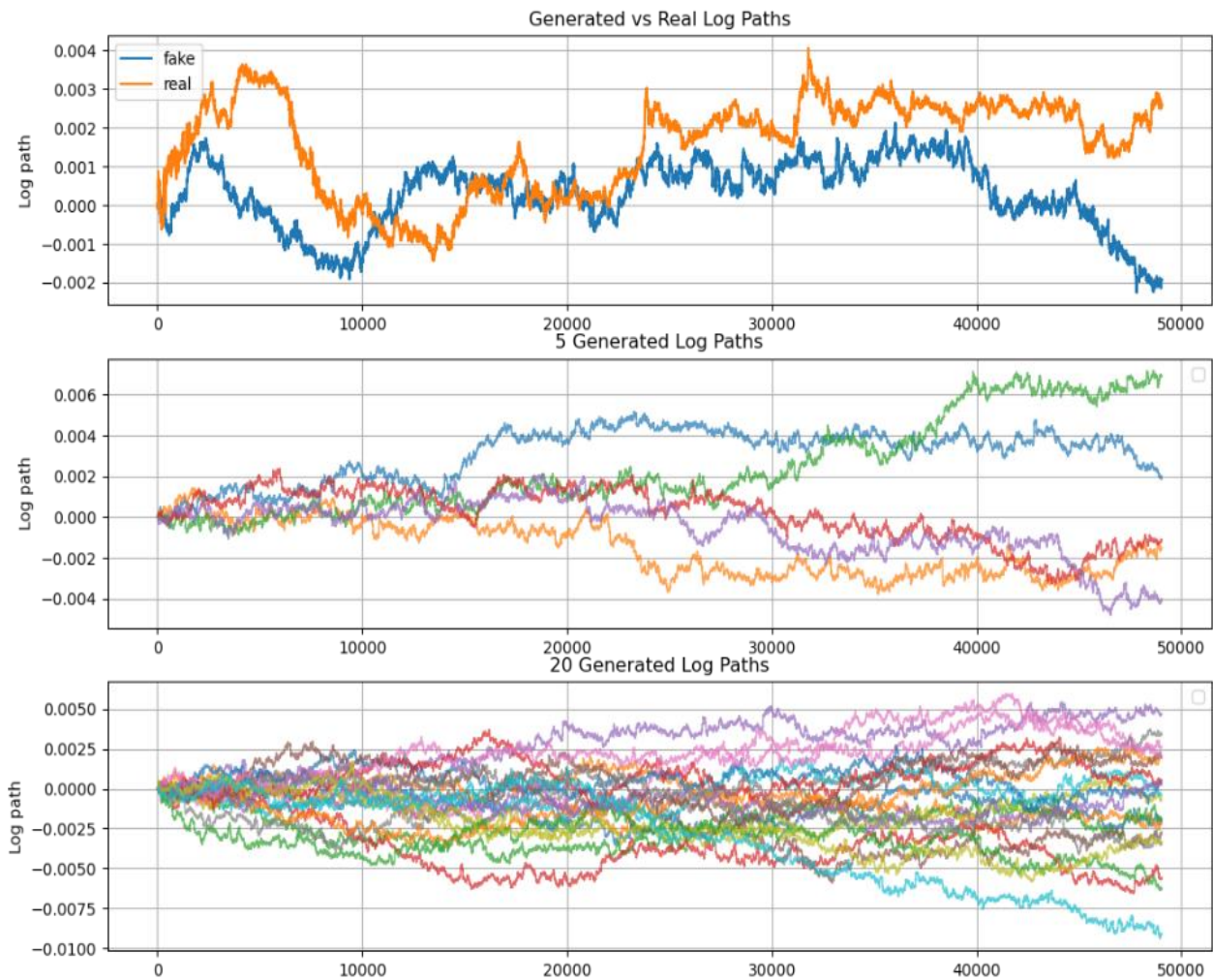


Fig. 1. Generated AUDUSD exchange rate yield trend simulation data

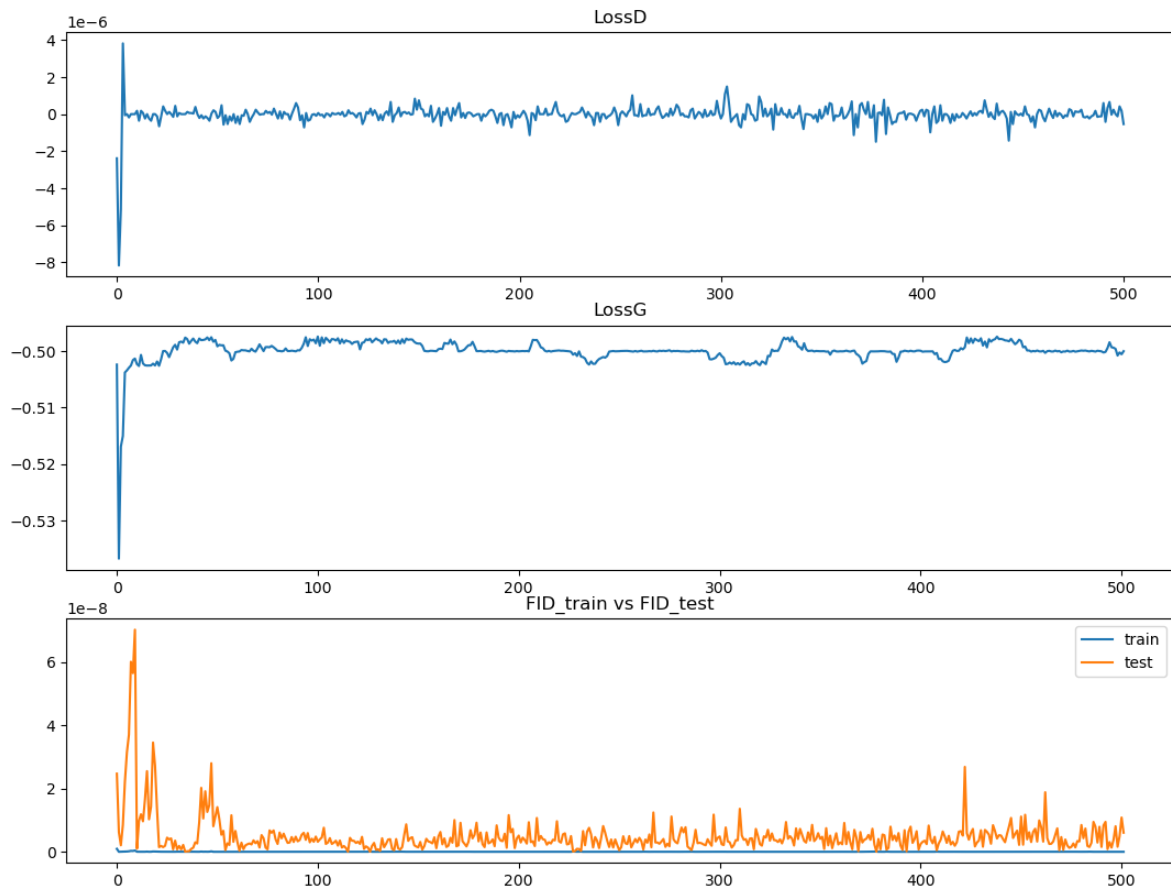


Fig. 2. Quant-GAN model AUDUSD data set learning curve

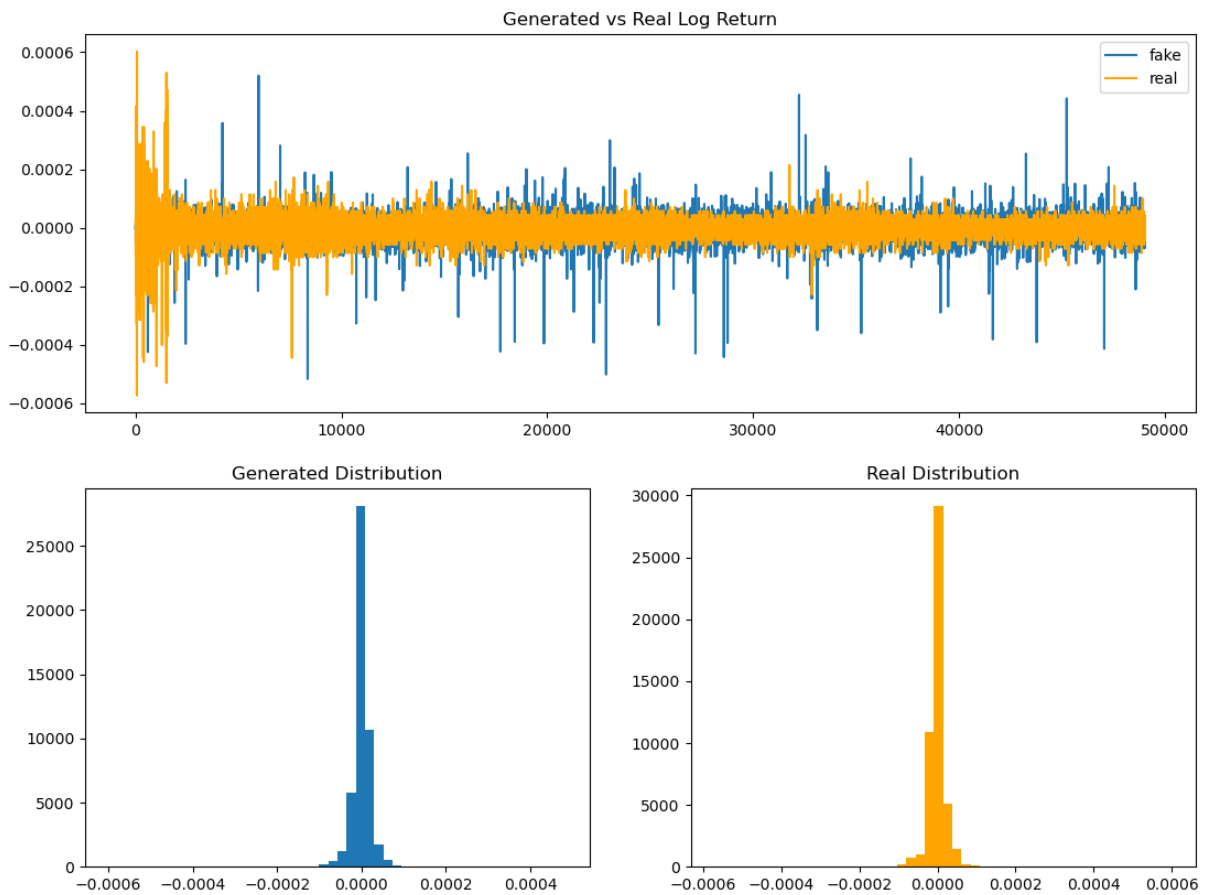


Fig. 3. Generated and real AUDUSD exchange rate logarithmic return distribution

Typical facts of the dataset: as the lag period increases, the autocorrelation between the data generated by the model and the training data is close to 0 (Fig. 4), showing that both lack autocorrelation, the correlation coefficient with the absolute value of the training data set sequence is not 0 and does not decrease as the lag period increases (Fig. 4), indicating an volatility aggregation effect; the tails of the cumulative frequency distribution of the sequence of the training data are far more than those of the normal distribution and the power-law distribution (Fig. 5), indicating that both distributions present heavy tails; the EGARCH parameter estimation

results of the model are significant, and the p-value of the coefficient of the asymmetric ARCH term is less than 0.05 (see Table 5), indicating that both sequences have leverage effects.

Various periods of time-lag distribution diagrams show that the two series are gradually approaching the normal distribution as the time scale t increases (Fig. 6).

Results show that the data generated by the QuantGAN model are in consistent with the training data in terms of main typical facts, such as autocorrelation, distribution heavy tail, leverage effect, aggregational Gaussianity, volatility aggregation, etc.

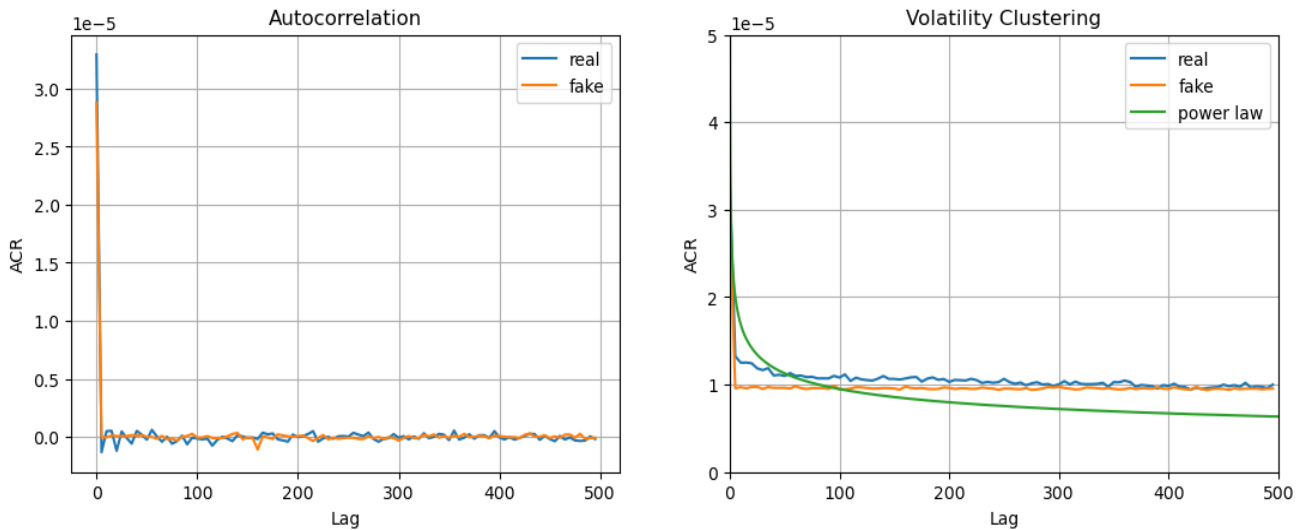


Fig. 4. Generated AUDUSD exchange rate logarithmic rate of return autocorrelation distribution

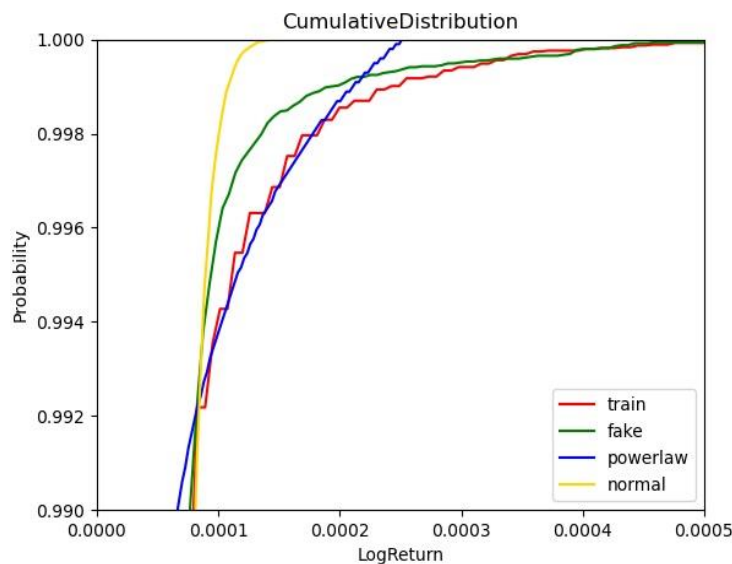


Fig. 5. Generated AUDUSD exchange rate logarithmic return cumulative frequency distribution

Table 5 – Estimated results of GARCH asymmetric term

	coeff	std err	t	P> t	95.0% Conf. Int.
fake: alpha[1]	0.0935	3.065e-02	3.050	2.289e-03	[3.341e-02, 0.154]
real: alpha[1]	0.0250	6.337e-03	3.952	7.751e-05	[1.262e-02, 3.746e-02]

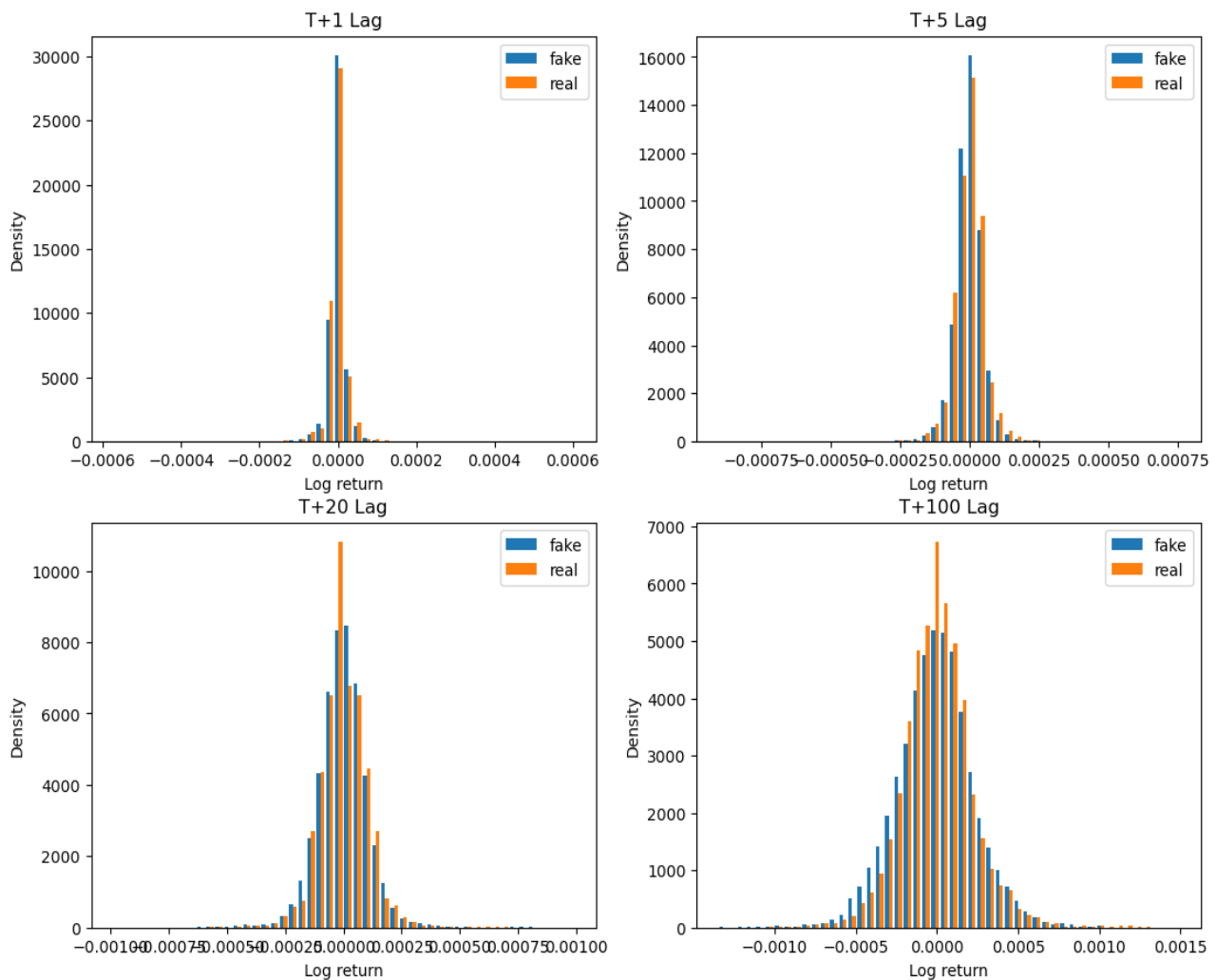


Fig. 6. Distribution of the generated AUDUSD exchange rate logarithmic rate of return

4 Summary and Discussion

This paper introduces the problem of simulating financial market transaction price time series data and related work on machine learning using Generative Adversarial Network (GAN) models. Price data from major currency pairs traded in international markets is selected as the training set to build and train a QuantGAN model.

The trained model is used to generate exchange rate return rate series data for each currency pair in different time dimensions. Experimental results show that data generated by the model effectively fits statistical characteristics and typical facts of real training datasets with good overall fit.

This research provides global foreign exchange market participants with an effective means of generating exchange rate simulation data for stress testing, scenario simulation, trading strategy and portfolio back-testing, derivatives pricing and more. Applying GAN models to

generate and evaluate high-frequency financial time series simulation data expands practical application scenarios for machine learning algorithms such as GAN.

Future work includes accumulating data and increasing computing power, optimizing and improving GAN models, and establishing evaluation standards for generating exchange rate price data. The foreign exchange market is a global over-the-counter (OTC) market where each trader has access to only a small portion of transactions.

Obtaining as much accurate exchange rate transaction data (especially high-frequency data) as possible is essential for model processing. Good data supports good model performance and helps avoid “garbage in, garbage out” situations. Processing massive high-frequency data also requires powerful computing power.

As computing power continues to grow, the GAN model’s ability to process ultra-large-scale datasets is expected to improve.

REFERENCES

- Graves, A. (2012), “Supervised sequence labeling with recurrent neural networks”, *Studies in Computational Intelligence*, vol. 385, Springer, Berlin, Heidelberg, 142 p., doi: <https://link.springer.com/book/10.1007/978-3-642-24797-2>.
- Xu, Q., Gao, H., Yuan, Ya., Guo, Ch., Sun, Yu, Wu, F. and Weinberger, K. (2018), “An empirical study on evaluation metrics of generative adversarial networks”, *arXiv Preprint*, arXiv:1806.07755, doi: <https://doi.org/10.48550/arXiv.1806.07755>.

3. Kingsep, K. (2021), *Quant gan Code In GitHub*, doi: <https://github.com/KseniaKingsep/quantgan>.
4. Hlavcheva, D., Yaloveha, V., Podorozhniak, A., Kuchuk, H. (2021), "Comparison of CNNs for Lung Biopsy Images Classification", 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering, UKRCON 2021 – Proceedings, pp. 1–5, DOI: <https://doi.org/10.1109/UKRCON53503.2021.9575305>.
5. Yaloveha, V., Podorozhniak, A., Kuchuk, H. (2021), "Convolutional neural network hyperparameter optimization applied to land cover classification", *Radioelectronic and Computer Systems*, No. 1(2022), pp. 115–128, DOI: <https://doi.org/10.32620/reks.2022.1.09>.
6. Podorozhniak, A., Liubchenko, N., Kvochka, M., & Suarez, I. (2021). Usage of intelligent methods for multispectral data processing in the field of environmental monitoring. *Advanced Information Systems*, Vol. 5, No. 3, pp. 97–102, doi: <https://doi.org/10.20998/2522-9052.2021.3.13>.
7. Krepych, S., & Spivak, I. (2020). Forecasting system of utilities service costs based on neural network. *Advanced Information Systems*, Vol. 4, No. 4, pp. 102–108, doi: <https://doi.org/10.20998/2522-9052.2020.4.14>.
8. Chalyi, S. & Leshchynskiy, V. (2021), "Hierarchical representation of causal relationships to detail explanations in intelligent systems", *Advanced Information Systems*, Vol. 5, No. 4, pp. 103–108, doi: <https://doi.org/10.20998/2522-9052.2021.4.14>.
9. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. And Bengio, Y. (2014), "Generative Adversarial Nets", *Advances in Neural Information Processing Systems*, 27, NIPS 2014, Montreal, Canada, doi: <https://doi.org/10.48550/arXiv.1406.2661>.
10. 2Wiese, M., Knobloch, R., Korn, R. and Kretschmer, P. (2020), "Quant GANs: Deep Generation of Financial Time Series", arXiv Preprint, arXiv: 1907.06673, doi: <https://doi.org/10.1080/14697688.2020.1730426>.
11. 3Arjovsky, M., Chintala, S. and Bottou, L. (2017), "Wasserstein generative adversarial networks", *International Conference on Machine Learning*, ICML 2017, Sydney, Australia, doi: <https://doi.org/10.48550/arXiv.1701.07875>.
12. 5Cont. R. (2001), "Empirical properties of asset returns: stylized facts and statistical issues", *Quantitative Finance*, vol. 1 (2), pp. 223–236, doi: <http://rama.cont.perso.math.cnrs.fr/pdf/empirical.pdf>.
13. 6Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B. and Hochreiter S. (2017), "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium", *Advances in Neural Information Processing Systems*, 30, NIPS 2017, Long Beach, United States, doi: <https://doi.org/10.48550/arXiv.1706.08500>.

Received (Надійшла) 30.03.2023

Accepted for publication (Прийнята до друку) 24.05.2023

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Дунь Бао – трейдер, Банк Китаю, Пекін, Китай; магістрант кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Bao Dun – Trader, Bank of China, Beijing, China; master's student of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv; e-mail: Bao_Dun@gmail.com; ORCID ID: <https://orcid.org/0000-0002-5349-7618>.

Заковоротний Олександр Юрійович – доктор технічних наук, професор, завідувач кафедри комп'ютерної інженерії та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Oleksandr Zakovorotnyi – Doctor of Technical Sciences, Professor, Head of Computer Engineering and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine; e-mail: Oleksandr.Zakovorotnyi@khp.edu.ua; ORCID ID: <http://orcid.org/0000-0003-4415-838X>.

Кучук Ніна Георгіївна – доктор технічних наук, професор, професор обчислювальної техніки та програмування, Національний технічний університет "Харківський політехнічний інститут", Харків, Україна;

Nina Kuchuk – Doctor of Technical Sciences, Professor, Professor of Computer Science and Programming Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine; e-mail: nina_kuchuk@ukr.net; ORCID ID: <http://orcid.org/0000-0002-0784-1465>.

Генерація даних про обмінний курс валюти на основі моделі Quant-GAN

Дунь Бао, О. Ю. Заковоротний, Н. Г. Кучук

Анотація. Мета дослідження. У цьому документі обговорюється використання алгоритмів машинного навчання для генерації даних, які відповідають вимогам наукових кіл та промисловості в контексті коливань обмінного курсу. **Результати дослідження.** У статті створено модель Quant-GAN з використанням часових згорткових нейронних мереж (CNN) і навчено її на високочастотних курсах валютних пар на світовому ринку наприкінці дня та всередині дня. Згенеровані дані оцінюються за допомогою різних статистичних методів і виявляються такими, що ефективно імітують реальний набір даних. Експериментальні результати показують, що дані, згенеровані моделлю, ефективно відповідають статистичним характеристикам і типовим фактам реальних навчальних наборів даних із загальною хорошою відповідністю. Результати надають учасникам глобального валютного ринку ефективні засоби для виконання різноманітних завдань, таких як стрес-тести та моделювання сценаріїв. **Майбутня робота** включає накопичення даних і збільшення обчислювальної потужності, оптимізацію та вдосконалення моделей GAN, а також встановлення стандартів оцінки для генерації даних про ціни обмінного курсу. Оскільки обчислювальна потужність продовжує зростати, очікується, що здатність моделі GAN обробляти надвеликі масиви даних покращиться.

Ключові слова: згорткова нейронна мережа; генеративні змагальні мережі; валютний курс.

Anastasiya Zakharchenko, Oleksandr Stepanets

National Technical University “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine

DIGITAL TWIN VALUE IN INTELLIGENT BUILDING DEVELOPMENT

Abstract. The aim of the research. This article discusses the use of the Digital Twin in automation and its impact on the resulting solution. The research aims to illuminate the Digital Twin concept explanation and systematise the knowledge base and fulfill information gaps. **Research results.** The paper overviews the history of the concept and determines the main phases of Digital Twin development. The significant attention was paid to the classification issue to show the huge variation depending on Digital Twin’s purpose, lifecycle phase, the scale of the physical twins and data amount in order to explain the twin’s relation and the hierarchy of complex system. The defined capabilities and values of the concept identify the possible use cases and explain the potential benefits of Digital Twin implementation. Also, this paper takes a look at the use of Digital Twin in the area of building automation. This concept potentially may act as the integration platform for building management systems (BMS) and building information modelling (BIM) technologies with IoT solutions. The discussion of Digital Twin implementation for the building automation complex is presented. We conclude that the Digital Twin can integrate human factor to the control system by using the indexes of climate satisfaction, the feedback functionality and human-machine interfaces. As a result, the improvement of system efficiency depends on the coordination and orchestration of equipment operating mode. **Conclusion.** The Digital Twin has a high potential for energy efficiency improvements, as it considers many factors, integrates a huge amount of data and continuously improves themselves with real-world data.

Keywords: Simulation; Data-based control; Real-time control; Internet of Things; Digital Twin; Knowledge-based control; Building Automation.

Introduction

An important topic in control technology today is the collection, management and analysis of data, which has received a significant amount of attention from academics and industry. The trend to digitalisation and intellectualisation of various sectors of production and technology, including building maintenance, is generating enormous interest in the IIoT (Industrial Internet of Things) concept. In this way, data can be directly collected by smart devices and transmitted to the cloud for further processing and storage. However, with the rapid growth of data volume, they will create challenges in data organisation, management and usage.

Improving the energy efficiency of the system operation can be achieved by improving the control algorithms. It is associated with the advanced analytic, additional parameters, predicting the object behavior, equipment diagnostics, etc. All these cases, combined with digitalisation, the need for online monitoring, operational flexibility, lead to the necessity of using a specific integration platform, like the Digital Twin [1]. The technological push for the development of the twin concept includes the availability of low-cost sensors, the development of cloud and edge computing, computing hardware, machine learning and artificial intelligence technologies [2]. From this point of view, Digital Twins become an extension of technological development and an example of data science application that filters, processes and integrates the massive amount of data with different data types [1].

While Digital Twins vary greatly depending on the purposes and data amount, it is a technology for virtualising the physical world. Digital Twins become the proxy for a physical entity, so it can interact with any application to exchange the data [3].

An overview of operating costs indicates the potential for savings by optimising and coordinating the systems’ operation to promote energy efficiency [4]. A

building is a complex object in terms of maintenance, as it has many connected engineering systems, external and internal disturbances that are difficult to predict. So, the system integration of the building is a priority task for today. Therefore, the use of Digital Twin in building automation has a great potential for energy-efficient system design. This research is aimed to provide analysis in three main blocks:

- 1) the current knowledge base on the Digital Twin topic with an enabling technology overview;
- 2) the information systematisation and filling the knowledge gaps;
- 3) applying the described technologies and use cases in building automation.

This paper is organized as follows. Section 2 presents the methodology of the literature review and shows the summaries of considered papers. Section 3 introduces the historical aspects of Digital Twin concept. Section 4 describes the capabilities and values of Digital Twin implementation and provide the general classification and overview of certain enabling technologies. Section 5 discusses possible Digital Twin implementation issue in order of the building automation environment. Finally, Section 6 presents our conclusions.

1 Methodology

The article research approach is divided into several steps. The first step of the study was to identify the purpose of the paper and to frame the research questions. It gives the vector of the following work and helps to define desirable results and its criteria. The search took place in the databases of Science Direct, Google Scholar and among the technical documentation and white papers of the world’s leading companies, scientific and market leaders, including Siemens, Emerson, XMPPro, ABB, Seebo, AnyLogic, ARUP, Oracle, Johnson Controls, Fraunhofer and others. The review of the technical documentation was carried out to study the proposals for practical implementation or

specific examples of application and to consider a more commercial point of view.

The search process was iterative. At first, we considered 60 sources and sifted out those that did not fit the general topic. As their analysis did not answer all the questions, it was decided to expand the list of literature. The next search was more specialized and related to topics in which we did not get information, including Digital Twin values, building Digital Twin application, and papers in the field of building energy consumption. The main inclusion criteria based on suitability for the purposes of the study, its research question and direct relation to the Digital Twin. Also, the authority and availability of full text had significant value of the selection. The final paper allocation, except standards, is shown in Table 1 by the publication type.

Table 1 – Ranking the papers collection by publication type

Publication type	Number of papers
Journal	37
Conference	11
Book (section)	4
White paper and reports	16

2 A brief overview of Digital Twin

The first approach of Digital Twin realization belonged to NASA [5], which was a “pioneer” of this technology during the Apollo program in the 1970s [6]. However, the term Digital twin appeared far more later at the Grieves’s presentation about Product Lifecycle Management (PLM) in 2003 [1, 6–8]. Currently, there is no standardized definition of Digital Twin [5]. It is possible to highlight three different interpretations of the Digital Twin term, which emphasize its continuous evolution (Fig. 1) [9–11].

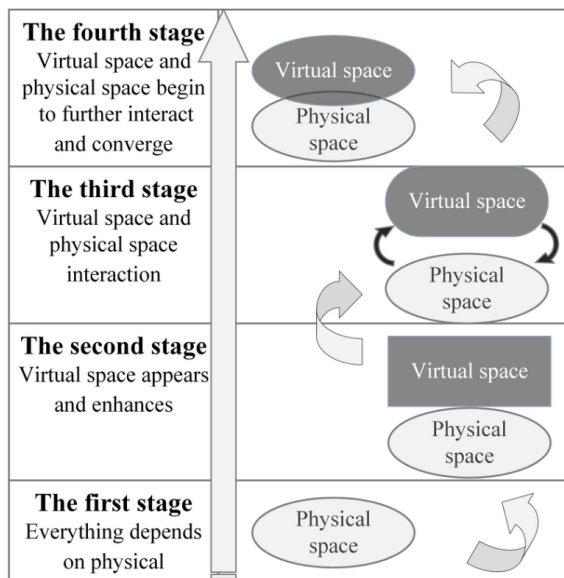


Fig. 1. The four stages of the Digital Twin evolution

The first one was introduced by NASA in 2010 as “an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin. It is ultra-

realistic and may consider one or more important and interdependent vehicle systems” [5, 12–14].

The second interpretation was provided by Siemens in 2016 and focuses on the simulated and visible 3D model and it was the appropriate vision of a broad industrial audience [9]. Thus, the Digital Twin topic turned numerous researchers’ attention and give rise to an exponential increase in the number of relevant publications [7].

The state-of-the-art concept of Digital Twin appeared in 2018 by extending the existing 3-dimension Digital Twin model and adding Data and Services dimensions [7, 10]. This change was intended to facilitate the further application of Digital Twins in various fields. The 5-dimensional model consists of a physical entity (PE), virtual entity (VE), digital twin data (DD), service system (SS), connection (CN) (Fig. 2).

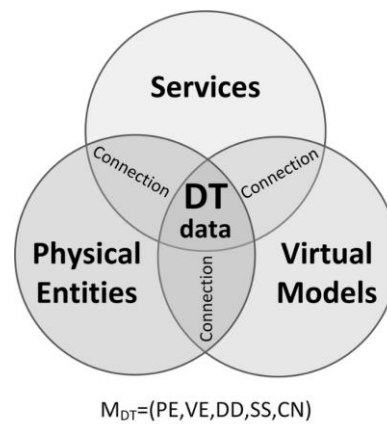


Fig. 2. Five-dimension digital twin model

3 Understanding of Digital Twin

3.1 Classification

3.1.1 General classification

Based on the reviewed literature, it can be concluded that the classification theme is very complex and does not have a clear common position of the authors. If we consider the classification by the level of integration, focused area or product lifecycle phase, and the scale of the physical twins. The division by level of integration separates 3 following groups: digital model, digital shadow and Digital Twin. In this case, the data flow between the physical and digital counterparts are characterized and the differences in understanding of the concept are explained below (Fig. 3) [8, 15–19].

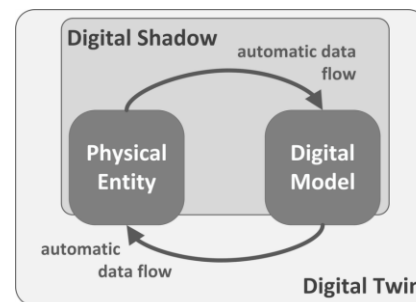


Fig. 3. The data flow in Digital Model, Digital Shadow and Digital Twin

The first group includes digital representations that are manually modelled and have no direct mechanisms for automatic data exchange in any direction. The Digital Shadow is provided by a one-way data stream and changes in the state of the physical object led to updates in its shadow. But there is no feedback from the analytical data. The Digital Twin is characterised by a bidirectional data flow and an integration between the digital and physical objects. In the context of the Digital Twin development path under consideration (section 3), in our opinion, this classification is not indicative, as current trends in technology lead to the view that only the last group meets the true criteria and requirements of this concept.

Classification by industry focus or by product lifecycle phase involves the division into Design, Manufacturing, Service, and Retire phases [17]. However, their usage is more relevant to industrial applications [19]. In the context of the considered industries, the Digital Twin can and should be functional in all phases of the life cycle, which can be realised by having a single Digital Twin with several models of the object that characterise a physical entity from different sides.

It is also possible to classify the Digital Twin by the scale of their physical objects into the following groups: critical component, piece of equipment, single machine, production line, processes and systems [16], [20]. Moreover, the twins of complex objects can

integrate Digital Twins of their component elements, which is in line with the ideology and the trends of this concept. Thus, the relation between the Digital Twins is more complex and is divided into the following groups [21]:

- Hierarchical relation: the twins of complex objects consist of Digital Twins of their component elements. In this manner, the Digital Twin replicates the architecture of a real object.
- Associational relation: the relationship between Digital Twins with technologically associated processes or facilities.
- Peer-to-peer relation: the relationship between Digital Twins of the same or similar type of equipment that perform the same or similar functions.

Despite the spread of such a general classification, in our opinion, its relevance will decrease with the development of Digital Twins, as predicted in section 3. Thus, there is much more interest in classifying the technologies that can be used to create the Digital Twin.

3.1.2 Enabling technologies classification

In accordance with the definitions of the Digital Twin, described in the previous sections, the following four categories of enabling technologies can be identified: modelling, data management, services and connection technologies [7, 19]. Fig. 4 summarises this information and presents the technological architecture of the Digital Twin.

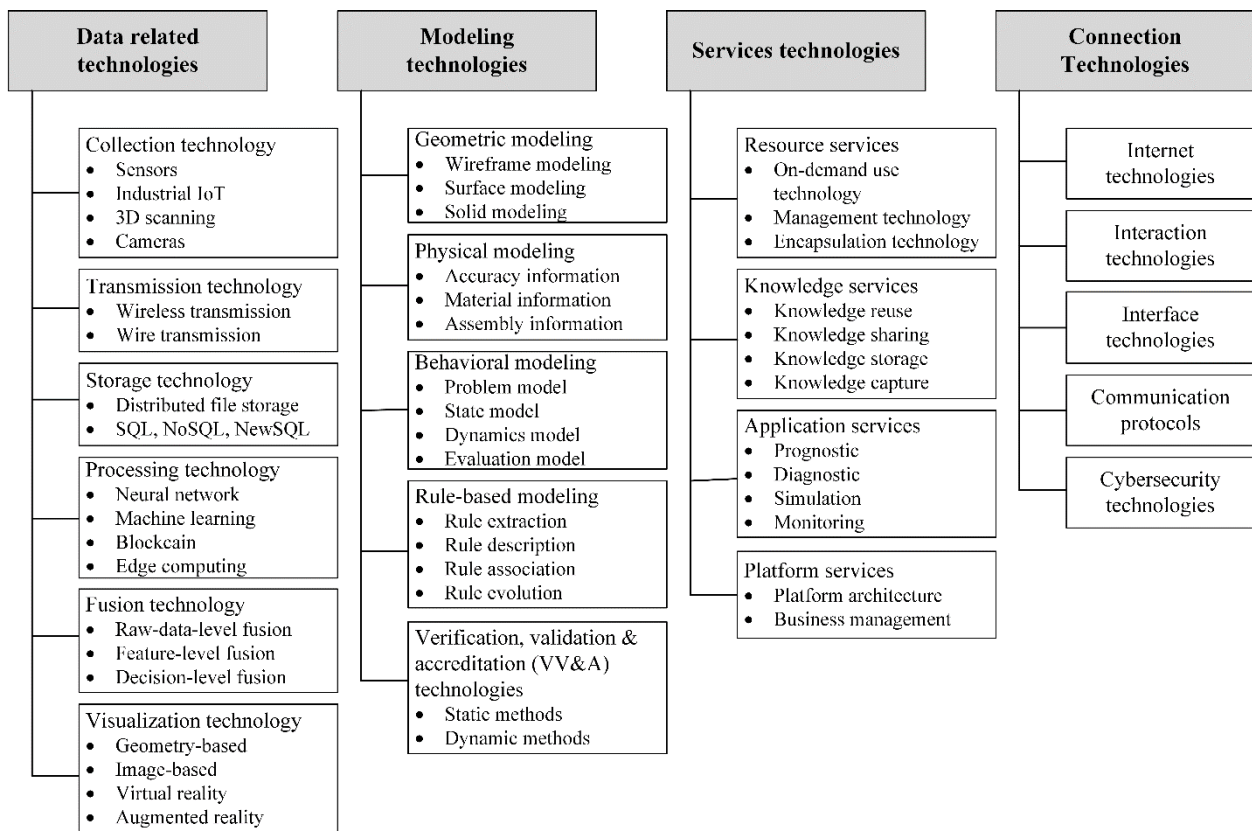


Fig. 4. The technological architecture of the Digital Twin

Data related technologies summarise data lifecycle and their transfer into information and further into knowledge. The Digital Twin should contain data

about the physical environment that are required to represent and analyze the behavior and states of the real-world object [22]. It is characterised by the

following phases: collection, transmission, storage, processing, fusion and visualization [7, 19]. The beginning of the data path starts with the operation of sensors, cameras, RFID tags, IT gadgets, 3D scanning, etc. to collect complete information about a particular physical environment.

The next step is *data transmission* using wired and wireless technologies that need to be performed in real or near real-time to provide data processing and analytics. Due to the volume of information collected and the high demands on processing time, it may be necessary to pre-process the data. For this purpose, the concept of edge computing is ideally suited and reduces the complexity and cost of transmission [19].

Obtaining multisource data requires *data fusion* through their synthesis, filtering, correlation, and integration. Data fusion includes raw-data-level fusion, feature-level fusion, and decision-level fusion [7].

An important stage of the data life cycle is *visualization*, which includes various graphical methods of presentation: histogram, pie chart, line chart, map, bubble chart, tree chart, dashboards, etc. Virtual and augmented reality technologies can also bring development in this direction [23].

Modelling technologies could be classified in a couple of ways. First of all, the modelling approaches can be divided according to the amount of physical knowledge used [5]:

- black-box is a mathematical data-driven model;
- white-box is a physical-based model;
- gray-box is a combination of previous types.

On the other hand, Digital Twin involves geometric, physical, behavioral, and rule-based modelling. As mentioned earlier, single Digital Twin can obtain several models, that bring different information about the physical object [7]:

- the geometric model describes geometric and tolerance information;
- the physical model includes accuracy material and assembly information;
- the behavioral model provides information about the way in which the physical entity performs functions, reacts to changes, interacts with others and maintains health;
- the rule-based model defines rules derived from historical data, expert knowledge and predefined logic.

To ensure the model accuracy and simulation confidence, *Verification Validation and Accreditation (VV&A)* analyses the relevance of the running result to the requirements, evaluates the models' sensitivity and tests the transformation from models to program codes [6, 7, 10, 24]. *Service technologies* could be divided into four groups: resource services, knowledge services, application services, platform services [7] (Fig. 4). A number of services operate at the upper most level and use integrated knowledge to generate added value through advanced analysis and validation [25]. The main benefit of services using is an on-demand operating organization. This means that the services will be provided at the right time and to the right extent. Otherwise, it was necessary to provide all resources,

software mechanisms and organise their maintenance, updates and management locally, which makes the owner of the system responsible for its operation. Also, the computing capacity is selected with a reserve to cover peak loads and ensure the appropriate level of reliability. In case of service, we delegate this responsibility, which could save time and money.

Connection technologies achieve bidirectional data interaction based on intelligent data interaction devices and communication interfaces [25–27]. The definition and understanding of the required technology are closely linked to the 5-D Digital Twin model that was described earlier. Each of its elements is interconnected with the others and forms six different connection options [7]:

- Physics-Services;
- Physics-Model;
- Physics-Data;
- Service-Model;
- Service-Data;
- Model-Data.

It also includes the links between the components of each element, the environment, among the various Digital Twins and even human-machine relations. As an example, Digital Twins can store their data in different databases [14]. There are vast differences in the technology's usage depending on the architecture of the Digital Twins, the environment and the availability of edge computing. The connections are enabled using several technologies, such as Internet technologies, human-computer interaction technologies, security technologies, interface technologies, communication protocols, etc [7].

3.2 The Digital Twin capabilities

The Digital Twin concept includes a wide range of capabilities. However, the list of its functions were changed with Digital Twin evolution [9], so the range of possible capabilities is generalized in Table 2 [28]. It is describing the Digital Twin value and use cases throughout the lifecycle of a real object with the classification of selected papers by their focus area and described technology.

The brief review of the mentioned capabilities is given below.

The PLM theory became the source of Digital Twin creation and it originally was an idea of the digitalization process and mostly means the approach for the holistic control and administration of product-related information [5]. PLM integrates all data and information collected along the item lifecycle in order to increase productivity and efficiency level by the interconnection of different isolated and fragmented processes and characteristics [29]. It supports five lifecycle phases: planning phase, create-phase, build phase, sustain phase and disposal phase [5].

The mount of the Digital Twin documentation includes operational instructions and service records that are used in further analysis. For example, maintenance records may be used to inform the manufacturer about detecting issues that would be noted and used to increase the quality of its product [22].

Table 2 – The Digital Twin capabilities

FUNCTIONALITY	FEATURE	PLAN	BUILD	OPERATE	MAINTAIN
All documents associated to equipment throughout its lifecycle	Document management	PLM	PLM	Operation instructions	Service record
		[5, 11, 14, 17, 29, 54, 62–64]		[25, 54]	[15, 54]
Digital representation of the equipment that can mimic properties and behaviours of a physical device	Model	Physical properties predict		Optimization	Diagnostics
		[2, 5, 7, 44, 55, 65]		[3, 5, 30, 31, 44, 65, 66]	
Representation of a physical device in a simulation environment to study its behaviour	Simulation	Design simulation	Virtual commissioning		
		[5, 35, 55]	[5, 22, 25, 45, 67]		
Properties of a physical device mapped to a 3D digital representation	3D representation	Design drawings	Manufacturing instructions		Service instructions
		[7, 37]			[27]
Standardized data model for connectivity, analytics, and/or visualization	Data model	Engineering data	Production data	Operational data	Service data
		[1, 14, 37, 49, 68]		[1, 49, 68–70]	[36]
Graphical representation of the object either on a supervisory screen or personal device	Visualization			Operational state display	Health status display
Alignment of a model with real-world parameters	Model synchronization			[2, 7, 11, 21, 44, 65]	
				Real-time movement	Model inversion
				[23, 30, 55, 66, 68]	
Algorithms and computational results based on measured properties of a physical device	Connected analytics			Operational KPIs	Asset health KPIs
				[37]	

The modelling function usually is mentioned by a different author as a central component of Digital Twins. Generally, Digital Twin is a simulation of the real world and the capable modelling environment is crucial [30], [31]. It obtains the wide research area directed at the development of more accurate digital models. The virtual model can adapt in real-time to the physical object changes to generate up-to-date data, predict properties, optimization and even diagnostic for maintaining equipment [27, 32]. It could represent a system, part or a family of parts while the Digital Twin represents an instance. Simulation has been underlined as the most appropriate method to model dynamic material and energy flows in the conditions of the complexity of process interactions and a large volume of variables [33, 34]. The purpose of a simulation is to study behaviour and performance of a physical device [14] and to determine the characteristics or make predictions about the real system using the data that cannot be measured during the operation process. In this context, the Digital Twin gets a wide analytical possibility and has a lot of data that is supported by 3D representation of the system or object and contains the geometrical information about the entity starting with design drawings. The advantage of this feature is not only the provision of 3D content and real-time analysis to workers [20], but its use for training sessions with remote, hazardous or sensitive equipment of users, technicians or operators [7, 35]. Virtual representation might be so accurate that a human could operate the system or object exactly the same as in the physical world [20]. So it reduces the waste of time and money to gain users or workers experience.

3D simulation not only provides insight into the physical and geometric organisation of objects, but is

also important for fault diagnostics, helping to identify the cause of a malfunction or to find a solution [36]. The Digital Twin is data-driven technology. The data model is designed to achieve different tasks: control, processing, diagnostics and condition monitoring, and analytics [37]. As a result, it integrates a vast amount of data that are generated during the object life cycle, including the planning, operating, maintaining and end-of-life phase. The data cycle consists of collection, transmission, storage, processing, fusion and visualization stages [7] and it mainly means that all Digital Twin function is dependent on the data flow and its transformation to information. The accumulated knowledge needs to be properly presented to the user or operator in an intuitive and interactive way for the purpose of his prompt and correct reaction. The generation of data begins even before the physical object appears, by preserving the planning phase documentation, developing digital models, etc. and inheriting data and knowledge from previous generations of similar physical objects.

The digitisation process leads to a perception complication of information through its complexity and volume, which caused a growth in tools, technologies and services for processing and visualization of information for its analysis, support for operator, user and manager decision-making [27]. The Digital Twin's ability to visualise data should be a key parameter to use, evaluate and understand the amount of information it generates.

The range of mathematical models, calculation methods, software and service solutions allow the model to be synchronised with a real object in real-time [37]–[39]. It enables obtaining additional data during operation and use it to optimise control, diagnostics and

monitoring [17]. Comparison and processing of the Digital Twin data with real measurements keep the modelling accurate and relevant by ensuring a bidirectional connection between the physical process and the virtual system.

The *operation state display* comes from a real-time data stream and is presented as plots, characteristics and time-series data generated by the simulations. The health status is displayed on the local application, gateway or cloud as an alarm and notification [37]. Information dashboards based on Digital Twins could be improved with included information (historical data, real-time data, and predicted data) and models that enable the use of 3D visual monitoring technologies, Augmented Reality (AR) or Virtual Reality (VR) as well as displaying near real-time analytics [19], [20].

A set of *Key Performance Indicators (KPIs)* is generated to assess the performance of the system, which can be linked to various targets to help evaluate the relevance of performance to specific business objectives by measuring assets' health, wear and performance [2], [40]. The set of KPIs includes [37]:

- operational KPIs - are based on process, electrical, mechanical and control data;

- asset health KPIs - use specific services and diagnostic KPIs - process, electrical, mechanical and control data.

The measurable performance indicators for a smart building context may include energy efficiency, equipment service and operating costs, number of failures, specific emissions of pollutants, room efficiency etc. The KPIs are based on the process data. The evaluation of the asset health often uses algorithms that determine the risk of failure and its remaining useful life. Although, as a rule, only these health KPIs do not reflect a performance degradation, which can impair performance long before the actual failure or end of life. An analyst based on digital twins can detect performance degradation using access to large amounts of data [41].

3.3 Value of Digital Twin

In addition to the technical features presented above, it is equally significant to describe the strategic advantages of using digital twin solutions. These values were partly mentioned in the capabilities of the Digital Twin itself (Table 2). All advantages were generalised in Fig. 5 and are described below. Generally, we divided the advantages into 5 groups.

Design	System interaction and integration	Monitoring and Diagnostic	Prediction and Optimization	Advanced services
Continuous improvement	Disparate systems connection	Device operating analysis	What-if analysis	Operator training
Provide data integrity	Learning from all devices	Real-time remote monitoring and control	Decision support system	Digitize expert experience
Virtual testing			Forecast future operating condition	Human-machine integration
Virtual evaluation & verification	Filling gaps		Predictive maintenance	Documentation automation
Advanced personalization of product and services	Business analytic application	Fault detection and diagnostic	Energy consumption planning	Intra- and inter-team synergy & collaboration
Improve trustworthiness	Gain control over system-of-systems	Anomaly detection		Troubleshoot equipment in remote service center
		Intellectual Alarm Management	Scenario and risk assesment	
PLUGandWORK enviroment				

Fig. 5. Value of Digital Twin

3.3.1 Design

The digital twin has extensive facilities at the system design stage. It includes virtual testing of the digital twin or system before its actual operation by applying 3D models, behavioural and other models and their subsequent virtual validation and verification with data integrity analysis [19, 25, 27, 42]. It ensures, above all, an increase in system reliability through automatic verification of system requirements, element compatibility analysis, etc. [3, 25]. The Digital Twin makes it possible to personalise the system and services to meet the requirements and preferences of users, market trends and legislative requirements at the design stage [2], while the coverage of a large number of digital technologies and innovations becomes the basis for continuous improvement of the system [43].

3.3.2 System interaction and integration

The Digital Twin provides mechanisms for interconnecting a system of heterogeneous objects [11], their communication, organisation and analytics. It enables data from all devices to be received, compared and analysed [44]. The Digital Twin allows to provide complete, detailed information and increase the level of control over an object or a system of objects, on the one hand, and facilitate the integration of the object into the system, on the other. The creation of complex systems can lead to the fragmentation of knowledge about its elements and processes, which complicates access to useful information. The use of the Digital Twins concept helps to solve this problem by making it easier to obtain data for the next generations of objects [19] and transfer information when there is a change of

ownership, audit, etc. [3]. The continuous accumulation of data throughout the life cycle and its analysis supports decision-making during the design and operation phase without the need for additional requests and information collection [19]. The intelligence and self-description capability of the Digital Twin makes it possible to create a Plug-and-Work environment where new components can be linked at runtime and automatically integrated at the functional level, business analytics and other IIoT solutions [45].

3.3.3 Monitoring and diagnostic

A Digital Twin can be described as a near-real-time digital image of a physical system providing monitoring, diagnostics, optimisation functions, etc. [7], [43]. An important advantage of Digital Twins is the visibility through the transparency of the processes and operations that take place in the system, the mapping of the links between structural elements, which are necessary to improve productivity and a deeper understanding of the object by the user or operator [20]. It enables real-time visualisation of equipment that is used by real people and operating in a particular environment [11]. The integration of heterogeneous systems and real-time work simulations makes it possible to fill in data that cannot be obtained at a physical object or is not rational, for example [44]:

- It is too expensive to install sensors on all equipment in all locations;
- The sensor can make the equipment run worse or interfere with it;
- It is not possible to provide reliable, economically justified communication with the sensor;
- Data is collected at different time intervals.

The Digital Twin makes it possible to monitor the status of equipment in real-time and to detect when devices are not working efficiently or are not interacting properly and provide extended information about them. The ability to detect anomalies is one of the pressing themes of IIoT [44]. It detects when the system is not working properly and helps to determine what may be the cause or how long it will last before it breaks down [39]. The anomaly does not always indicate an error, but may only show a state that has not been seen before [44].

3.3.4 Intellectual Alarm Management

Intellectual Alarm Management is another important application of the Digital Twin. Its purpose is to alert and support operator or user solutions when they arise. Alarm flooding means that one emergency causes many other errors and prevents the quick identification of the root cause of the fault [46]. Digital Twin helps to avoid such flooding.

3.3.5 Prediction and optimization

The use of physics models makes it possible to predict the behaviour and state of equipment, which includes forecasting future operating conditions [2, 43]. It is important for the implementation of the decision support system, making it possible to obtain more informed or effective solutions and to deepen knowledge about the physical object. Scenarios and risk assessments are performed to detect unexpected

behaviour situations, study them and find ways to mitigate or solve them [2]. Data from the different states of the object that are generated by the what-if analysis can be used [27]:

- to improve future system designs;
- to optimise service cycles;
- to outline ideas for new system applications;
- to approve preliminary design solutions;
- to predict the response of the system to various types of field disruptions.

The analysis of the system's operation enables predictive maintenance, based on the anomaly detection, what-if analysis, fault detection and planning capabilities described earlier [5, 27, 42, 47, 48]. It helps to reduce maintenance costs and prevent serious equipment breakdowns. The set of prediction functions described above makes it possible to carry out energy consumption, maintenance and cost planning, which is positive for commercial institutions and private use.

It provides an opportunity to improve operating conditions, to understand the relationship between indoor climate and the amount of energy used or its cost. The values of the Digital Twin involve a positive impact on the environmental and energy consciousness for users and teach them to understand their responsibility.

4 Digital Twin in building automation

4.1 General overview

The application of Digital Twin technology in the area of intelligent buildings has its specifics. In particular, some technologies, principles and standards already exist to support the building's life cycle. It is necessary to determine the place of the Digital Twin concerning them and answer the questions:

- What benefits does it offer for building automation?
- What practical sense do the mentioned features have?

First of all, it is necessary to get acquainted with the terminology used in the field of building automation.

The terms building automation and control system (BACS), building management system (BMS) are synonymous according to EN 15232. They incorporate all systems, equipment and services designed to manage, monitor, and optimise the performance of building engineering systems to improve their energy efficiency, cost savings and safety [4]. The term building automation system (BAS), which is widely used in literature, is also equivalent in meaning [49]. This system integrates the control functions of multiple electrical devices, heating, ventilating, cooling, air conditioning, hot water and lighting appliances and has the task of providing user comfort with energy performance improving [50].

A similar meaning is expressed in the energy management system (BEMS), which is part of the BMS and focuses on optimising energy consumption. BEMS is a computerised system for monitoring and controlling, which includes data collection, logging, reporting, and analysis of energy consumption. There are many

realisations of BEMS from numerous manufacturers, generally performing seven functions [50]:

- installation management and control (sequencing control and process control);
- energy-management functions;
- risk-management functions (fire alarm systems and security systems);
- information-processing functions;
- performance monitoring and diagnosis (fault detection and diagnosis technology, automatic commissioning technology, and smart maintenance technology);
- facility management.

BMS needs to use models of building and its engineering systems behaviour to achieve its potential in the field of energy efficiency.

Often the building’s structural elements are planned separately by different consultants, contractors, and other stakeholders. The lack of coordination can lead to serious problems and collisions during the construction phase [51, 52]. The term building information modelling (BIM) refers to the process of developing an intellectual model that brings together the work of specialists in the architecture, engineering, and construction industry [53,54]. Initially, the BIM concept was considered to improve collaboration between the various participants in design and construction processes, but it quickly became incorporated into related areas, gaining applications throughout the building’s life cycle and providing strong interdisciplinary cooperation [8, 52, 53]. BIM is applicable at various stages of the life cycle for both new and existing buildings [24, 53]. BIM provides for the collection and comprehensive processing of all architectural, technological, economic and other information about the object [49] from conception to dismantling of the building and enables the integration of knowledge from different disciplines into one model [53]. On the other hand, in combination with the digitalization

trend in the real estate sector, the volume of information obtained by BIM, their complexity and abundant have increased exponentially, causing a “drowning in data” [27, 52], which must be processed and stored to obtain valuable knowledge.

The difference between BIM and the Digital Twin is that BIM represents a process, while the Digital Twin is a virtual representation of an object in real-time [55]. Due to outdated formats and standards, the current state of BIM does not allow for integration with IIoT solutions [52] or BMS platforms [53], which limits the application of data. The unique aspect of the building as a control object is that it is a complex socio-technological system. This feature cannot be involved with BIM. Moreover, it requires a large amount of memory and computation power and includes advanced end-user interaction by BMS [50].

The result of BIM is a static informational model, which does not support dynamic modelling, but contains most of the information required by the Digital Twin. Siemens defines the Digital Twin of a building as “a digital copy of a physical building, which includes a 3D model of a facility combined with dynamic data to allow easy-to-understand visualization and analysis” and consists of [51]:

- Digital Twin of the devices in a building;
- Digital Twin of the structural components of a building;
- Digital Twin of the dynamic data (performance data, time-series data).

Thus, the Digital Twin capabilities mentioned in the previous chapters have been partly explored in BIM, BMS and BEMS. The Digital Twin should provide them with enhanced analytics and interact to form a complex for efficient building management or take over their functions (Fig. 6).

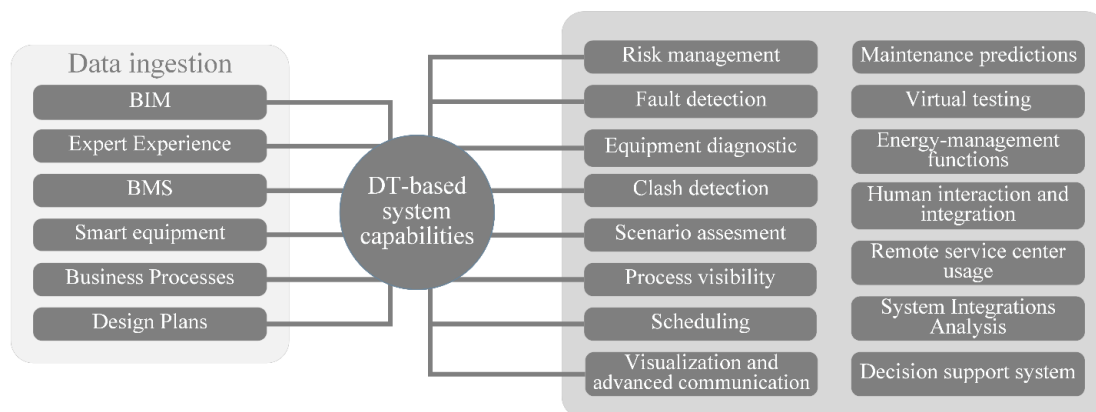


Fig. 6. Building Digital Twin implementation

The application of the Digital Twin solution comes through the potential of using dynamic realtime modelling and the ability to connect services to implement or improve certain BMS functions. Building energy systems are characterised by non-linear behaviour, stochastic environmental interactions [5] and the uniqueness of each object and operating mode. The accuracy of energy analysis and building performance forecasts is closely linked to the accuracy of modelling, which is influenced by several factors. Firstly, the use of real object data makes it

possible to refine the model and maintain its relevance in the process of equipment operation [49, 51]; secondly, the use of real-time data for calculations makes it possible to improve the accuracy of forecasts [53].

The Digital Twin makes it possible to create a digital version of the building infrastructure, including its engineering systems. Such a system can investigate building usage modes and user preferences using AI technology. The obtained information is used to optimise the performance of the physical twin using

analysis of different system scenarios, which gives different modelling results parallel with the actual operation [5]. Based on the results, it is possible to select the workflow scenario automatically according to predetermined criteria, or with the user's participation when the Digital Twin acts as an advisor.

4.2 Control strategy

Considering the heating and cooling sector as an example, which is one of the most energy-intensive building systems, the realisation of control algorithms will be highly dependent on the type of building and its operating mode. The control of non-residential and residential buildings is a little different in principle. The first category includes industrial and commercial (showrooms, offices, shopping malls, etc.) buildings, theatres, and museums. They are characterised by the presence of defined working hours or production plan, the ability to predict the occupancy of a building and the type of human activities. Also, they could have some technological constraints associated with indoor climate parameters. Moreover, non-residential buildings sometimes have a facility manager, who is responsible for maintaining a comfortable and safe environment and making operational decisions. These parameters allow us to improve the building model, which increases the accuracy of the technological parameter prediction.

The occupants' presence and behaviour can significantly affect the correctness of predictions and represents one of the essential factors that determine energy consumption in buildings [56]. Therefore, this factor needs to be considered in both building types. The easiest option for incorporating occupancy into a building behaviour model is the use of schedules. However, it may not be flexible enough for residential building management, as the daily routines may vary due to certain events in the lives of the residents.

In this case, users manage and adapt the building performance to their own needs during operation, without having as much competence as, for example, a facility manager. Thus, the management of a residential building requires an interaction between the occupants and the building. The occupant actively participates in supervision using flexible, user-friendly decision-supported interfaces. The collected operational and historical data are used to examine the correlations in human behaviour and to integrate them into the Digital Twin.

There are different objectives for building management, which will determine the used analytics and datasets. It is usual to consider criteria such as comfort, cost of maintenance and energy efficiency. The last two are often wrongly equated with one another.

The calculation of an optimal operating plan for heating, cooling and other energy-consuming equipment aims to minimise the total energy usage, its cost or the prioritisation of energy sources and is based on weather forecasts, operating conditions, etc [57]. Energy consumption minimisation does not necessarily lead to maintenance cost minimisation.

This result requires the introduction of algorithms that take into account the availability of different energy sources, their efficiency and tariff variability. Thus, the

accumulation of thermal energy in "lowcost" periods may reduce the overall energy cost, but may not be optimal in terms of the energy consumed. Other options for control objectives are shifting energy use away from peak periods of the power grid, prioritising the use of renewable energy sources, minimising greenhouse gas emissions, etc. [58].

One of the main functions of building automation systems is to provide quality of the indoor environment, including acoustic comfort, visual comfort, indoor air quality and thermal comfort for occupants or users. Some standards and methods represent the requirements and allow the assessment of the human satisfaction level in the numerical form [59, 60].

Optimal climate control can improve productivity and reduce the risk of human illness. The simplest and most common control option is to keep the temperature in the building within a defined range, although this option does not always give the most comfortable result. Consequently, the Predicted Mean Vote (PMV), and Predicted Percentage Dissatisfied (PPD) indexes have been developed for this purpose. PMV indicates the heat sense for the human body and PPD assesses the percentage of people who feel heat discomfort. PPD can be derived from PMV [61].

The PVM index is a complex non-linear equation that considers a large number of parameters, such as metabolic rate, clothes heat transfer resistance, airspeed and temperature, air velocity and others. The index value is dimensionless and represents a scale from -3 to +3, where zero value will correspond to the most comfortable conditions [61].

The use of non-linear indicators in the optimisation process makes the assessment of the environment more realistic, but it complicates the calculation and requires more computing power. Thus, the inclusion of the human factor in the building's Digital Twin provides additional subjective information about the object, which enables a more comprehensive data analysis for making management decisions. Moreover, the Digital Twin can include far more virtual sensors than are actually installed, thereby filling information gaps that could not be obtained from the sensor data.

Conclusion

Although the idea of a Digital Twin is not a new one, the development of modern data collection, transmission, processing, modelling and service technologies together with the development of hardware finally gives the possibility to consider its wide usage. This complex of technologies offers a wide range of capabilities during all lifecycle of physical twins. No matter the focus area, product lifecycle phase, the scale of the physical twins or used enabling technologies, Digital Twin save its concept of being the virtual representation of the physical world and act as a proxy to organise the complex analytical, controlling, diagnostic and interaction processes.

The main Digital Twin ideology is visibility, which is open to interpretation in several ways. Basically, this means the possible presence of 3D models that helps to inspect and explore the equipment even if it is hidden from human eyes. Also, it includes visibility throughout

the timeline. It means the representation of physical twins' past historical, real-time and predicted future operating states. Digital Twin promotes the accuracy of prediction by updating the object model with new data during its work. The prediction function directly impacts on the anomaly detection, what-if analysis, fault detection and planning capabilities. Finally, visibility means filling the gap by using virtual sensors if it is not able to get real-world measurements and provide the user interface that represents the information in an intuitive form.

All these features are applicable in predictive control concepts, predictive maintenance, system improvement testing, staff training and business analytics, etc. And these conclusions are fully suitable for the building automation area.

Concerning the building automation environment, the Digital Twin solutions should provide BIM and BMS with advanced functionality, solve the problem of system integration with IIoT solutions, or take over their function. The Digital Twin for the building has to consider the human component as a significant source of disturbance and as the main purpose of facility controlling. The realisation of human satisfaction assessment may include the complex indexes, feedback functionality and intuitive interfaces for human-machine interaction. Whether it is a residential or commercial building, the Digital Twin has a high potential for energy efficiency improvements, as it considers many factors, integrates a huge amount of data and continuously improves themselves with real-world data.

REFERENCES

1. F. Tao, F. Sui, A. Liu, *et al.*, "Digital twin-driven product design framework," *International Journal of Production Research*, vol. 57, pp. 3935–3953, 12 2019. doi: [10.1080/00207543.2018.1443229](https://doi.org/10.1080/00207543.2018.1443229).
2. A. Rasheed, O. San, and T. Kvamsdal, "Digital twin: Values, challenges and enablers," *IEEE Access*, vol. 8, pp. 21980–22012, 2019. doi: [10.1109/ACCESS.2020.2970143](https://doi.org/10.1109/ACCESS.2020.2970143).
3. P. van Schalkwyk, "XMPPro. The Ultimate Guide to Digital Twins," Tech. Rep., 2019. [Online]. Available: <https://xmpro.com/digital-twins-the-ultimate-guide/>.
4. "En 15232:2012 - Energy performance of buildings—impact of building automation, controls and building management," European Committee for Standardization, Standart, 2012. Available at: <https://standards.iteh.ai/catalog/standards>.
5. C. Vering, M. Nu'renberg, P. Mehrfeld, D. Coakley, M. Lauster, and D. Mueller, "Unlocking potentials of building energy systems' operational efficiency: Application of digital twin design for HVAC systems," in *Building Simulation 2019*, 2019, pp. 1304–1310. doi: [10.26868/25222708.2019.210257](https://doi.org/10.26868/25222708.2019.210257).
6. D. J. Wagg, K. Worden, R. J. Barthorpe, and P. Gardner, "Digital Twins: State-of-The-Art and Future Directions for Modeling and Simulation in Engineering Dynamics Applications," *ASCEASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, vol. 6, p. 030901, 3 2020. doi: [10.1115/1.4046739](https://doi.org/10.1115/1.4046739).
7. Q. Qi, F. Tao, T. Hu, *et al.*, "Enabling technologies and tools for digital twin," *Journal of Manufacturing Systems*, pp. 3–21, 2019. doi: [10.1016/j.jmsy.2019.10.001](https://doi.org/10.1016/j.jmsy.2019.10.001).
8. D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the Digital Twin: A systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, 2020. doi: [10.1016/j.cirpj.2020.02.002](https://doi.org/10.1016/j.cirpj.2020.02.002).
9. R. Drath, "The Digital Twin: The evolution of a key concept of Industry 4.0," *visIT Industrial IoT - Digital Twin*, pp. 6–7, 2018. [Online]. Available: www.iosb.fraunhofer.de.
10. F. Tao and M. Zhang, "Digital twin shop-floor: A new shop-floor paradigm towards smart manufacturing," *IEEE Access*, vol. 5, pp. 20418–20427, 2017. doi: [10.1109/ACCESS.2017.2756069](https://doi.org/10.1109/ACCESS.2017.2756069).
11. L. Kitain, "The new age of manufacturing: Digital Twin technology & IIoT," Tech. Rep., 2018. [Online]. Available: <https://www.seebo.com/digital-twin-technology/>.
12. M. Macchi, I. Roda, E. Negri, and L. Fumagalli, "Exploring the role of digital twin for Asset Lifecycle Management," *IFAC-PapersOnLine*, vol. 51, pp. 790–795, 11 2018. doi: [10.1016/j.ifacol.2018.08.415](https://doi.org/10.1016/j.ifacol.2018.08.415).
13. J. Ovtcharova and M. Grethler, "Beyond the Digital Twin – making analytics come alive," *visIT Industrial IoT - Digital Twin*, pp. 4–5, 2018. [Online]. Available: www.iosb.fraunhofer.de.
14. G. N. Schroeder, C. Steinmetz, C. E. Pereira, and D. B. Espindola, "Digital twin data modeling with AutomationML and a communication methodology for data exchange," *IFAC-PapersOnLine*, vol. 49, pp. 12–17, 2016. doi: [10.1016/j.ifacol.2016.11.115](https://doi.org/10.1016/j.ifacol.2016.11.115).
15. H. Zhang, L. Ma, J. Sun, H. Lin, and M. Thü'rer, "Digital twin in services and industrial product service systems: Review and analysis," in *Procedia CIRP*, vol. 83, Elsevier B.V., 2019, pp. 57–60. doi: [10.1016/j.procir.2019.02.131](https://doi.org/10.1016/j.procir.2019.02.131).
16. C. Cimino, E. Negri, and L. Fumagalli, "Review of digital twin applications in manufacturing," *Computers in Industry*, vol. 113, p. 103130, 2019. doi: [10.1016/j.compind.2019.103130](https://doi.org/10.1016/j.compind.2019.103130).
17. W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, pp. 1016–1022, 11 2018. doi: [10.1016/j.ifacol.2018.08.474](https://doi.org/10.1016/j.ifacol.2018.08.474).
18. A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108952–108971, 2020. doi: [10.1109/ACCESS.2020.2998358](https://doi.org/10.1109/ACCESS.2020.2998358).
19. M. Liu, S. Fang, H. Dong, and C. Xu, "Review of digital twin about concepts, technologies, and industrial applications," *Journal of Manufacturing Systems*, vol. 58, pp. 346–361, 2020. doi: [10.1016/j.jmsy.2020.06.017](https://doi.org/10.1016/j.jmsy.2020.06.017).
20. P. Raj and C. Surianarayanan, "Chapter Twelve - Digital twin: The industry use cases," in *The Digital Twin Paradigm for Smarter Systems and Environments: The Industry Use Cases*, ser. Advances in Computers, vol. 117, Academic Press Inc., 2020, pp. 285–320. doi: [10.1016/bs.adcom.2019.09.006](https://doi.org/10.1016/bs.adcom.2019.09.006).
21. P. van Schalkwyk, S. Malakuti, and S.-W. Lin, "A Short Introduction to Digital Twins," The Industrial Internet Consortium, Tech. Rep., 2019. Available: <https://www.iiconsortium.org/news/joi-articles/2019-November-JoI-A-Short-Introduction-to-Digital-Twins.pdf>.
22. S. Malakuti, P. Schalkwyk, B. Boss, *et al.*, "Digital Twins for Industrial Applications. Definition, Business Values, Design Aspects, Standards and Use Cases," Tech. Rep., 2020. Available: <https://www.iiconsortium.org/digital-twins-for-industrial-applications>.
23. Y. Lu and X. Xu, "Resource virtualization: A core technology for developing cyber-physical production systems," *Journal of Manufacturing Systems*, vol. 47, pp. 128–140, 2018. doi: [10.1016/j.jmsy.2018.05.003](https://doi.org/10.1016/j.jmsy.2018.05.003).

24. G. Angjeliu, D. Coronelli, and G. Cardani, "Development of the simulation model for digital twin applications in historical masonry buildings: The integration between numerical and experimental reality," *Computers and Structures*, vol. 238, p. 106282, 2020. doi: [10.1016/j.compstruc.2020.106282](https://doi.org/10.1016/j.compstruc.2020.106282).
25. G. F. Schneider, H. Wicaksono, and J. Ovtcharova, "Virtual engineering of cyber-physical automation systems: The case of control logic," *Advanced Engineering Informatics*, vol. 39, pp. 127–143, 2019. doi: [10.1016/j.aei.2018.11.009](https://doi.org/10.1016/j.aei.2018.11.009).
26. F. Xiang, Y. Huang, Z. Zhang, and Y. Zuo, "Digital twin driven energy-aware green design," in *Digital Twin Driven Smart Design*. 2020, pp. 165–184, isbn: 9780128189184. doi: [10.1016/B9780-12-818918-4.00006-3](https://doi.org/10.1016/B9780-12-818918-4.00006-3).
27. A. Madni, C. Madni, and S. Lucero, "Leveraging digital twin technology in model-based systems engineering," *Systems*, vol. 7, p. 7, 1 2019. doi: [10.3390/systems7010007](https://doi.org/10.3390/systems7010007).
28. K. E. Harper, C. Ganz, S. Malakuti, and S. Scientist, "Digital twin architecture and standards," Tech. Rep., 2019.
29. F. Tao, M. Zhang, and A. Nee, "Chapter 2 applications of digital twin," in *Digital Twin Driven Smart Manufacturing*. Elsevier, 2019, pp. 29–62. doi: [10.1016/b978-0-12-817630-6.00002-3](https://doi.org/10.1016/b978-0-12-817630-6.00002-3).
30. Anylogic, "An introduction to digital twin development," Tech. Rep., 2018. [Online]. Available: <https://www.anylogic.com/resources/white-papers/an-introduction-to-digital-twindevelopment/>.
31. A. Papacharalampopoulos and P. Stavropoulos, "Towards a digital twin for thermal processes: Control-centric approach," in *Procedia CIRP*, vol. 86, Elsevier B.V., 2020, pp. 110–115. doi: [10.1016/j.procir.2020.01.015](https://doi.org/10.1016/j.procir.2020.01.015).
32. P. Verboven, T. Defraeye, A. K. Datta, and B. Nicolai, "Digital twins of food process operations: The next step for food process models?" *Current Opinion in Food Science*, vol. 35, pp. 79–87, 2020. doi: [10.1016/j.cofs.2020.03.002](https://doi.org/10.1016/j.cofs.2020.03.002).
33. V. J. Mawson and B. R. Hughes, "Thermal modelling of manufacturing processes and HVAC systems," *Energy*, vol. 204, p. 117984, 2020. doi: [10.1016/j.energy.2020.117984](https://doi.org/10.1016/j.energy.2020.117984).
34. G. P. Lydon, S. Caranovic, I. Hischier, and A. Schlueter, "Coupled simulation of thermally active building systems to support a digital twin," *Energy and Buildings*, vol. 202, p. 109298, 2019. doi: [10.1016/j.enbuild.2019.07.015](https://doi.org/10.1016/j.enbuild.2019.07.015).
35. Emerson, "Emerson Digital Twin: A key technology for digital transformation," Tech. Rep., 2019. [Online]. Available: <https://www.emerson.com/documents/automation/emerson-digitaltwin-a-key-technology-for-digital-transformation-en-5262472.pdf>.
36. V. Gopinath, A. Srijia, and C. N. Sravanthi, "Re-design of smart homes with digital twins," *Journal of Physics: Conference Series*, vol. 1228, p. 012031, 1 2019. doi: [10.1088/1742-6596/1228/1/012031](https://doi.org/10.1088/1742-6596/1228/1/012031).
37. P. Lipnicki, D. Lewandowski, D. Pareschi, E. Ragaini, and G. Perrone, "Digital avatars for powertrains expand digital twin concept," *ABB Review. Digital twins and simulations*, pp. 52–57, 2019. Available: <https://global.abb/group/en/technology/abb-review>.
38. M. P. Acharya and M. J. Mousavi, "Real-time AI powered by edge-deployed digital twins," *ABB Review. Digital twins and simulations*, pp. 14–19, 2019. [Online]. Available: <https://global.abb/group/en/technology/abb-review>.
39. M. Bevilacqua, E. Bottani, F. E. Ciarapica, et al., "Digital twin reference model development to prevent operators' risk in process plants," *Sustainability (Switzerland)*, vol. 12, p. 1088, 3 2020. doi: [10.3390/su12031088](https://doi.org/10.3390/su12031088).
40. C. Ganz, "Cloud, edge and fog computing," *ABB Review. Digital twins and simulations*, pp. 78–79, 2019. [Online]. Available: <https://global.abb/group/en/technology/abb-review>.
41. K. Smiley, X. Qu, T. Galoppo, et al., "Managing solar asset performance with connected analytics," *ABB Review. Digital twins and simulations*, pp. 34–41, 2019. [Online]. Available: <https://global.abb/group/en/technology/abb-review>.
42. T. Y. Melesse, V. D. Pasquale, and S. Riemma, "Digital twin models in industrial operations: A systematic literature review," in *Procedia Manufacturing*, vol. 42, Elsevier B.V., 2020, pp. 267–272. doi: [10.1016/j.promfg.2020.02.084](https://doi.org/10.1016/j.promfg.2020.02.084).
43. L. Manca, R. Grugini, and R. Mirzazadeh, "Digital twin. a digital copy of reality that enables you to simulate and find answers in a risk free and secure environment," Tech. Rep., 2019. Available: <https://www.eng.it/en/white-papers/digital-twin>.
44. B. Klentz, "How to use streaming analytics to create a real-time Digital Twin," in *Proceedings of the SAS Global Forum 2018 Conference*. Cary, NC: SAS Institute Inc., 2018, pp. 2004–2018. [Online]. Available: <https://www.sas.com/content/dam/SAS/support/en/sas-global-forumproceedings/2018/2004-2018.pdf>.
45. O. Sauer, "The digital twin – a key technology for Industry 4.0," *visIT Industrial IoT - Digital Twin*, pp. 8–9, 2018. [Online]. Available: www.iosb.fraunhofer.de.
46. M. Fullen and F. Pethig, "Digital twins for industrial alarm management," *visIT Industrial IoT Digital Twin*, pp. 10–11, 2018. [Online]. Available: www.iosb.fraunhofer.de.
47. A. Werner, N. Zimmermann, and J. Lentz, "Approach for a holistic predictive maintenance strategy by incorporating a digital twin," *Procedia Manufacturing*, vol. 39, pp. 1743–1751, 2019. doi: [10.1016/j.promfg.2020.01.265](https://doi.org/10.1016/j.promfg.2020.01.265).
48. Oracle, "Digital twins for IoT applications: A comprehensive approach to implementing IoT digital twins," Tech. Rep., 2017. [Online]. Available: <https://www.oracle.com/assets/digital-twinsfor-iot-apps-wp-3491953.pdf>.
49. O. Maryasin, "Design of ontologies for the digital twin of buildings," *Ontology of designing*, vol. 9, pp. 81–96, 4 2019. doi: [10.18287/2223-9537-2019-9-4-480-495](https://doi.org/10.18287/2223-9537-2019-9-4-480-495).
50. "Chapter 8 - building automation for energy efficiency," in *Handbook of Energy Efficiency in Buildings*, F. Asdrubali and U. Desideri, Eds., Butterworth-Heinemann, 2019, pp. 597–673. doi: [10.1016/B978-0-12-812817-6.00008-5](https://doi.org/10.1016/B978-0-12-812817-6.00008-5).
51. Siemens, "Digital twin-Driving business value throughout the building life cycle," Tech. Rep., 2018. [Online]. Available: <https://new.siemens.com/uk/en/products/buildingtechnologies/whitepapers/digital-twin.html>.
52. C. Boje, A. Guerriero, S. Kubicki, and Y. Rezgui, "Towards a semantic construction digital twin: Directions for future research," *Automation in Construction*, vol. 114, p. 103179, 2020. doi: [10.1016/j.autcon.2020.103179](https://doi.org/10.1016/j.autcon.2020.103179).
53. C. Panteli, A. Kylili, and P. A. Fokaides, "Building information modelling applications in smart buildings: From design to commissioning and beyond a critical review," *Journal of Cleaner Production*, vol. 265, p. 121766, 2020. doi: [10.1016/j.jclepro.2020.121766](https://doi.org/10.1016/j.jclepro.2020.121766).
54. A. Khalil, S. Stravoravdis, and D. Backes, "Categorisation of building data in the digital documentation of heritage buildings," *Applied Geomatics*, 2020. doi: [10.1007/s12518-020-00322-7](https://doi.org/10.1007/s12518-020-00322-7).
55. J. Controls, "Applying digital twins to the built environment. a digital twin is a connected, virtual replica of a physical product," Tech. Rep., 2019. Available: <https://www.johnsoncontrols.com/insights/2019/bts/applying-digital-twins-to-the-built-environment>.
56. F. C. Sangogboye, K. Arendt, A. Singh, C. T. Veje, M. B. Kjærgaard, and B. N. Jørgensen, "Performance comparison of occupancy count estimation and prediction with common versus dedicated sensors for building model predictive control," *Building Simulation*, vol. 10, pp. 829–843, 6 2017. doi: [10.1007/s12273-017-0397-5](https://doi.org/10.1007/s12273-017-0397-5).

57. A. Jain, D. Nong, T. X. Nghiem, and R. Mangharam, "Digital Twins for efficient modeling and control of buildings: An integrated solution with SCADA systems.," in *Proc. of the 2018 Building Performance Analysis Conf. and SimBuild*, 2018.
58. J. Drgon'a, J. Arroyo, I. C. Figueroa, *et al.*, "All you need to know about model predictive control for buildings," *Annual Reviews in Control*, vol. 50, pp. 190–232, 2020. doi: [10.1016/j.arcontrol.2020.09.001](https://doi.org/10.1016/j.arcontrol.2020.09.001).
59. A. Zaballos, A. Briones, A. Massa, P. Centelles, and V. Caballero, "A smart campus' digital twin for sustainable comfort monitoring," *Sustainability*, vol. 12, p. 9196, 2020. doi: [10.3390/su12219196](https://doi.org/10.3390/su12219196).
60. J. Ridley, "Our mission is therefore not merely to liberate data, but to filter it. Our job is to extract meaning from the noise. To turn data into insight and insight into action.," *Digital twin: towards a meaningful framework subtitle. Report. ARUP*, pp. 88–91, 2019. [Online]. Available: www.arup.com/digitaltwinreport.
61. "Ergonomics of the thermal environment analytical determination and interpretation of thermal comfort using calculation of the pmv and ppd indices and local thermal comfort criteria," International Organization for Standardization, Standart, 2005.
62. K. Dohrmann, B. Gesing, and J. Ward, "A DHL perspective on the impact of digital twins on the logistics industry DHL trend research digital twins in logistics," DHL Customer Solutions & Innovation, Tech. Rep., 2018. Available at: <https://www.dhl.com/global-en/home/insightsand-innovation/thought-leadership/trend-reports/virtual-reality-digital-twins.html>.
63. F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, "Digital twin-driven product design, manufacturing and service with big data," *Int.J. of Adv. Manufacturing Techn.*, vol. 94, pp. 3563–3576, 9-12 2018. doi: [10.1007/s00170-017-0233-1](https://doi.org/10.1007/s00170-017-0233-1).
64. R. Vrabı̄c, J. A. Erkoyuncu, P. Butala, and R. Roy, "Digital twins: Understanding the added value of integrated models for through-life engineering services," in *Procedia Manufacturing*, vol. 16, Elsevier B.V., 2018, pp. 139–146. doi: [10.1016/j.promfg.2018.10.167](https://doi.org/10.1016/j.promfg.2018.10.167).
65. H. Aiki, K. Saito, K. Domoto, and H. Hirahara, "Boiler digital twin applying machine learning," *Mitsubishi Heavy Industries Technical Review*, vol. 55, 4 2018. Available at: <https://www.mhi.co.jp/technology/review/pdf/e554/e554130.pdf>.
66. T. S. Cognizant, "Nurturing digital twins: How to build virtual instances of physical assets to boost performance," Tech. Rep., 2019.
67. C. Wagner, J. Grothoff, U. Epple, *et al.*, "The role of the Industry 4.0 asset administration shell and the digital twin during the life cycle of a plant," in *2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2017, pp. 1–8. doi: [10.1109/ETFA.2017.8247583](https://doi.org/10.1109/ETFA.2017.8247583).
68. A. Roxin, W. Abdou, D. Ginhac, W. Derigent, D. Dragomirescu, and L. Montegut, "Digital building twins-contributions of the ANR McBIM project," in *15th International Conference on Signal Image Technology & Internet Based Systems, SITIS 2019*, 2019, p. 2420387. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-02420387>.
69. Altran, "Digital Twins: The Altran point of view," Tech. Rep., 2019. Available at: <https://www.altran.com/no/en/insight/the-digital-twin-how-to-create-value-now-andprepare-for-digital-operations-for-tomorrow/>.
70. D. A. Howard, J. Mazanti, A. Agrotech, Z. Ma, and B. N. Jørgensen, "Data architecture for digital twin of commercial greenhouse production," in *The 2020 RIVF International Conference on Computing & Communication Technologies*, 2020, pp. 1–7. doi: [10.1109/RIVF48685.2020.9140726](https://doi.org/10.1109/RIVF48685.2020.9140726).

Received (Надійшла) 20.03.2023

Accepted for publication (Прийнята до друку) 17.05.2023

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Захарченко Анастасія Сергіївна – аспірантка, кафедра автоматизації енергетичних процесів, Національний технічний університет України «КПІ ім. Ігоря Сікорського», Київ, Україна;

Anastasiya Zakharchenko – Post-Graduate student, Department of Automation of Energy Processes, National Technical University "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine;

e-mail: zakharchenko.anastasia@iit.kpi.ua; ORCID: <https://orcid.org/0000-0001-5767-3028>;

Степанець Олександр Васильович – кандидат технічних наук, доцент кафедри автоматизації енергетичних процесів, Національний технічний університет України «КПІ ім. Ігоря Сікорського», Київ, Україна;

Oleksandr Stepanets – Candidate of Technical Sciences, Associate Professor, Department of Automation of Energy Processes, National Technical University "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine;

e-mail: Stepanets.Oleksandr@iit.kpi.ua; ORCID: <https://orcid.org/0000-0003-4444-0705>.

Роль цифрового двійника в проектуванні інтелектуальних будівель

А. С. Захарченко, О. В. Степанець

Анотація. Мета дослідження. В роботі розглядається можливість Цифрових двійників і їх потенціал для систем автоматизації. Дослідження спрямоване на розгляд та систематизацію знань в сфері двійників як через призму історії виникнення та розвитку так і сучасного погляду на концепцію. **Результати дослідження.** Значну увагу було приділено питанню класифікації Цифрових двійників, що відображає значну варіацію функціональності в залежності від їх призначення, фази життєвого циклу, масштабу фізичних близнюків, та включає наявність зв'язків з іншими Цифровими двійниками, їх ієрархія з метою побудови двійників для складних систем. Чітко сформульовані можливості та цінності даної концепції допомагають визначити можливі варіанти практичного застосування технології та демонструють потенційні вигоди від впровадження Цифрових двійників. Водночас, у цій роботі розглядається використання Цифрових двійників у сфері автоматизації будівель, які потенційно можуть стати платформою для інтеграції систем управління будівлями та технологій інформаційного моделювання будівель з рішеннями в області технології Інтернету речей. Висвітлено питання впровадження Цифрового двійника для комплексу автоматизації інженерних систем будівель та можливість інтегрувати людський фактор в систему управління за допомогою індексів задоволеності кліматом та зворотного зв'язку з користувачами через людино-машинні інтерфейси. **Висновок** Цифровий двійник має високий потенціал для підвищення енергоефективності, оскільки враховує багато факторів, інтегрує величезну кількість даних і постійно вдосконалюється за допомогою реальних даних.

Ключові слова: моделювання; контроль на основі даних; контроль в реальному часі; інтернет речей; Цифровий двійник; контроль на основі знань; автоматизація будівель.

Methods of information systems protection

UDC 004.934:141.1

doi: <https://doi.org/10.20998/2522-9052.2023.2.12>

Serhii Datsenko, Heorhii Kuchuk

National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine

BIOMETRIC AUTHENTICATION UTILIZING CONVOLUTIONAL NEURAL NETWORKS

Abstract. Relevance. Cryptographic algorithms and protocols are important tools in modern cybersecurity. They are used in various applications, from simple software for encrypting computer information to complex information and telecommunications systems that implement various electronic trust services. Developing complete biometric cryptographic systems will allow using personal biometric data as a unique secret parameter instead of needing to remember cryptographic keys or using additional authentication devices. **The object of research** the process of generating cryptographic keys from biometric images of a person's face with the implementation of fuzzy extractors. **The subject of the research** is the means and methods of building a neural network using modern technologies. **The purpose of this paper** to study new methods for generating cryptographic keys from biometric images using convolutional neural networks and histogram of oriented gradients. **Research results.** The proposed technology allows for the implementation of a new cryptographic mechanism - a technology for generating reliable cryptographic passwords from biometric images for further use as attributes for access to secure systems, as well as a source of keys for existing cryptographic algorithms.

Keywords: biometric cryptographic systems; cryptographic keys; fuzzy extractors; convolutional neural network.

Introduction

Relevance. In contemporary cybersecurity, cryptographic algorithms and protocols are crucial tools. They are applied in a variety of applications, from straightforward computer information encryption software to sophisticated communications and information systems that incorporate multiple electronic trust services. Implementing comprehensive biometric cryptography systems will enable the use of personal biometric data as a unique secret parameter rather than having to remember cryptographic keys or rely on extra authentication devices.

Interest in biometric methods has grown drastically in recent years. Modern technologies replace traditional biometric systems by forming cryptographic keys on the spot, as discovered by comparing acquired biometric photos with preserved reference copies [1 - 4]. The development of full-fledged biometric cryptography systems, in which biometric data of personality should be applied as a source of unique secret parameters, might be the next step in the advancement of such technology [5, 6]. The end user no longer has to remember cryptographic keys (passwords) or utilize extra devices to transmit, store, and etc. The biometric cryptosystem may be initialized at anytime and anywhere by removing the required parameters on the spot (with practicable erasures, mistakes, etc.) without causing harm to the given pictures [5, 7].

An overview of scientific works. The process of authentication involves employing several identifying measures to verify the user's validity [8 - 10]. In the security system, during authentication procedure the information provided by the user will be compared with the database and upon the match, user will be granted access to this system [11, 12]. For user identification, biometric authentication systems rely on their distinctive traits [4, 5, [13]. Process whereby the person is

automatically identified based on a vector of characteristics selected from their physiological or behavioral features [5, 7]. This leads to a classification of biometric approaches into two categories: physiological and behavioral [4, 13, 14]. Physical characteristics that a person already holds, such as their hand, fingerprint, or face, are used in physiological biometrics. This usually originates from the fact that a person's features remain constant over time. Behavioral biometrics, on the other hand, are based on the user's actions, such as how they take notes or write articles [15].

Setting objectives. Analysis of biometric features involves a variety of research methods. Convolutional neural networks (CNN) are the most widespread. Histogram of oriented graphs (HOG) is another mathematical tool for pattern identification in computer vision systems. Based on this, **the goal of this paper** is to study these methods, their software implementation and experimental researches of their performance to solve problems of biometric authentication. In particular, the authentication precision and biometric image processing speed of CNN and HOG are evaluated.

Software implementation and convolutional neural network model description

The Python programming language utilizing `face_recognition` and `dlib` modules was used to develop software implementation of authentication algorithms for biometric images of facial features. These modules provide functions for HOG and CNN technologies, as well as the choice between 2 models (small standard and larger one) to read additional biometric features. The program allows us to detect facial features and compare them to those that remained in the collection.

A. Applied libraries and functions

The following libraries were used during software development.

1) Dlib - a modern C++ toolkit that incorporates machine learning techniques and tools for constructing multiplex software to solve real-world problems.

2) Mathematical library NumPy, which offers the following functionalities:

- powerful N-dimensional arrays;
- tools for integrating C/C++ and Fortran code;
- useful features of linear algebra, Fourier transformations and random numbers;
- complex (notification) functions;
- and much more.

3) The face_recognition library has been generated with the most recent dlib face recognition library. Its functionality enables us to recognize and manage faces. The Face Recognition library is made up of multiple code files and models that have been trained on a large number of examples and are freely available. It's also important to mention that the model scored 99.38% on the Labeled Faces in the Wild test

4) Face_recognition_models models from the auxiliary library.

5) Pillow Library (PIL) - a set of functions that adds the ability to process images to the Python interpreter. This library supports a huge range of file types, has efficient internal representation, and has a rather broad image processing capacity. The basic set of modules is intended to provide easy access to data saved in many basic pixel formats. It should serve as a solid foundation for the image processing tool as a whole.

During the development of the test software the following functions were utilized [16]:

1) The function load_image_file, which loads pictures for analysis.

2) The face_encodings function, which gives us access to the transmitted image array's face encoding.

3) The face_locations function provides an array of bounding squares for faces on the picture after taking the image file, the type of model (algorithm) to be used and the number of sampling times.

4) The face_distance function, which calculates the Euclidean distance by comparing the original image's encoding to the encoding of another face. How similar the faces are can be determined by this distance.

5) A list of base face encodings is compared to a list of candidate face encodings using the compare_faces function.

6) The percentdif function determines the likelihood that two bit strings will match.

7) A binar function that, using the received face metrics, produces a bit string.

B. Convolutional neural network model description

CNN-based object detector is trained by implementing the loss layer, known as loss_mmod, that is provided in the dlib library. To create this layer of loss, a unique method known as Max-Margin Object Detection (MMOD) is used. HOG functions are replaced by CNN in order to produce the same loss layer as the popular SVM+HOG object detector, and after that, the entire detector will be trained from beginning to end [17]. This opens up an opportunity for the development of stronger detectors.

First step is to define our CNN. CNN will be assessed in a complicated manner across the picture pyramid, namely the standard sliding window classifier. Herewith the CNN must be identified with the intention to view a part of the image and determine whether we are looking for the right object. In order to clearly determine whether a face is present in a 50x50 image, for example, we could define a CNN that uses a receptive field and has a roughly 50x50 pixel size, which is suitable for face detection. CNN with unique architectures may be advantageous for other applications.

For instance, CNN is configured with three layers of sample reduction [17]. The image will be scaled eight times with help of these layers, presenting a useful map with 32 sizes. The outcome of the preceding stage is then passed through 4 more convoluted layers to create the final network output. When the network considers locating an object in a specific place, the values will be enormous since the remaining layer consists of only 1 channel.

Network definition starts from creating some network blocks.

To reduce the sample in two times a 5x5 conversion layer is presented:

```
template <long num_filters, typename SUBNET> using
con5d = con<num_filters, 5, 5, 2, 2, SUBNET>;
```

A 3 × 3 conversion layer without reducing sampling is presented:

```
template <long num_filters, typename SUBNET> using
con3 = con<num_filters, 3, 3, 1, 1, SUBNET>;
```

With reference to convolutional 5d blocks we can now define an 8x sampling reduction block. Use of ReLU and batch normalization in the conventional manner is also applicable:

```
template <typename SUBNET> using downsampler =
relu<bn_con<con5d<32, relu<bn_con<con5d<32,
relu<bn_con<con5d<32, SUBNET>>>>>>>>>>>>;
```

The remaining network transforms into 3x3 convolutional layers with batch normalization and reuse. Therefore, the 3x3 block, that is used here, is defined:

```
template <typename SUBNET> using rcon3 =
relu<bn_con<con3<32, SUBNET>>>;
```

The entire network is completely defined. A special layer known as input_rgb_image_pyramid forces the network to operate on the spatial pyramid, which keeps the detector's scale constant:

```
using net_type =
loss_mmod<con<1, 6, 6, 1, 1, rcon3<rcon3<rcon3<downsa
mpler<input_rgb_image_pyramid<pyramid_down<6>>>>
>>>>;
```

In this scenario, the face detector will be trained using the catalog's small face data set. So, first and foremost, we must acquire this dataset:

```
const std::string faces_directory = argv[1];
```

A training data set and a separate test data set are comprised in the face catalog. The training data consists of four pictures, each of which is highlighted by rectangles that delimit each distinct human face. The goal is to use this training data to figure out how to detect people's faces in new photographs.

Furthermore, once an object detector has completed its training, it must be tested on data that has not been educated. As a result, a separate test set of five pictures is also loaded. The efficiency of the face detector derived from training data will be assessed by running it on other test photos.

As a result, variables containing a set of data are generated here. The position of training image faces will be stored in `face_boxes_train`, whereas `images_train` will contain four training images. For example, the image `images_train [0]` includes faces described by rectangles in the array `face_boxes_train [0]`:

```
std::vector<matrix<rgb_pixel>> images_train,
images_test;
std::vector<std::vector<mmod_rect>>
face_boxes_train, face_boxes_test;
```

XML files containing images from each data set as well as the positions of face borders from that point can be downloaded. Any input format can be used without a doubt if the data is stored in `images_train` and `face_boxes_train`:

```
load_image_dataset(images_train,
face_boxes_train,
faces_directory+"/training.xml");
load_image_dataset(images_test, face_boxes_test,
faces_directory+"/testing.xml");
```

The Max-Margin Object Detection method contains numerous parameters that can be adjusted to control how it functions. In any case, we can provide the constructor with training notes and the size of the targeted object, and it will naturally adjust itself to solve our problems. Faces, on the other hand, are still recognisable at 40x40 pixels. In most cases, we should go with the smallest size possible. In accordance with the preceding rule, the constructor, defined as `mmod_options`, will invariably determine the required width and height of the sliding window. It will also automatically select a fair maximum for the suppression parameters:

```
mmod_options options(face_boxes_train, 40,40);
```

If necessary, multiple sliding windows can be applied to the detector automatically. However, for these faces, only one is required.

A network and a simulator can now be built:

```
net_type net(options);
```

The loss of the MMOD necessitates a number of options. detector filters equal to `windows.size()`. As a result, it is established here:

```
net.subnet().layer_details().set_num_filters(opt
ions.detector_windows.size());
dnn_trainer<net_type> trainer(net);
trainer.set_learning_rate(0.1);
trainer.be_verbos();
trainer.set_synchronization_file("mmod_sync",
std::chrono::minutes(5));
trainer.set_iterations_without_progress_threshol
d(300);
```

The network must be educated at this point. 150-image miniature bundles will be used. The images can be acquired by selecting random samples from the training set:

```
std::vector<matrix<rgb_pixel>>
mini_batch_samples;
std::vector<std::vector<mmod_rect>>
mini_batch_labels;
random_cropper cropper;
cropper.set_chip_dims(200, 200);
```

Shredder requires any minimum dimensions that have been transferred to the constructor `mmod_options`, which is accomplished here:

```
cropper.set_min_object_size(40,40);
dlib::rand rnd;
```

The simulator will continue to operate until the rate of training becomes insignificant. It takes a long time. There is an option of randomly mixing colors, which typically helps the detector better infer new images:

```
while(trainer.get_learning_rate() >= 1e-4) {
cropper(150, images_train, face_boxes_train,
mini_batch_samples, mini_batch_labels);
for (auto&& img : mini_batch_samples)
disturb_colors(img, rnd);
trainer.train_one_step(mini_batch_samples,
mini_batch_labels);}
```

Training flows are scheduled to end soon:

```
trainer.get_net();
```

The network has been saved to disk:

```
net.clean();
serialize("mmod_network.dat") << net;
```

When a face detector is obtained, it can now be examined. The initial operation checks it on training input, whereas the secondary operation examines it on test input. Recall, accuracy, and then average accuracy will be outputted. This should indicate that the network operates properly when learning new information:

```
cout << "training results: " <<
test_object_detection_function(net, images_train,
face_boxes_train) << endl;
cout << "testing results: " <<
test_object_detection_function(net, images_test,
face_boxes_test) << endl;
```

C. Testing methods

The algorithm of software implementation and research methodology lies in performing the following steps:

1) A basic biometric image is introduced, and the face recognition method (CNN or HOG) is chosen for the pictures.

2) Applying the `face_locations` and `load_image_file` functions to load and locate the face in the base image. Using the `face_encodings` function, the found face image is processed, resulting in the creation of an array showing the distances between the face's primary points (biometric image encoding).

3) The test image is loaded (from a particular sample), processed, and encoded (as for the base image from step 2), and then stored.

4) Examining the array of distances between the base image and the test (using the `compare_faces` function).

5) Steps 3 and 4 are repeated in a loop for every image in the sample.

6) List of solutions (the outcome of arrays of distances being compared on all test photos) is created.

7) Experimental results (match probabilities, execution time, etc.) are outputted.

Therefore, using HOG and CNN technologies, experimental research will produce digital data from biometric images. The face_encodings function, in particular, enables us to encode a face from the resulting biometric image and generates a list of 128 real numbers that characterize various facial features. Then, a single binary number with 128-digit number is created using all of the obtained values. We used the following rule: a real number is assigned a "1" if it is more than or equal to zero, and a "0" if it is less than zero. The acquired 128-digit number can be also compared to one another and utilized as a model for a key (access password) in the future. By employing passwords created in this manner, the efficiency of biometric authentication may be evaluated. The results indicate a pretty high level of match probability. We have a requirement for authentication if, specifically, the generated passwords match the matching biometric photos by more than 80%, indicating that only one user is responsible for them.

Results of Experiment

The program was evaluated using both ways on a sample of 480 test photos of various faces in accordance with the aforementioned algorithm. It was assessed how quickly biometric features were processed and how likely it was that certain features from the original image would match those in all other images. The results were compared to one another and summarized in Table 1 and Fig. 1, respectively. Each value in the table and figure, in particular, corresponds to a different biometric image

from the test. The zero (0) image is the baseline (beginning) image used to compare with all other images. The procedures used before extracting the metrics from the null file refer to the model that was used to obtain the face detector, hence they take a lot longer than the following images. The processing time is roughly the same for the remaining photos. The total processing time for all 480 test photos is displayed in the table's final column. We can see that HOG's technology is a little behind CNN's.

The probabilities of matching between biometric pictures of human faces acquired using the HOG and CNN technologies are shown in Fig. 1, respectively. Only 10 pairs of typical values from 480 test results are represented in the figure. As we can see, the outcomes produced by both technologies—HOG and CNN—are essentially the same. However, CNN typically provides a more accurate answer.

The following general conclusions can be taken from the results:

- Although the HOG method addresses the face detector a little bit faster, it receives facial metrics a lot slower. On the basis of this, CNN requires less time overall to run the authentication process;
- The use of CNN technology offers somewhat more precise face metrics, which increases the likelihood that specific biometric traits will match.

Thus, it is best to use convolutional neural networks to resolve the task of recognizing bio-metric images. With the help of this technology, we can quickly and with a high degree of probability compare face metrics extracted from any image. Additionally, this technology will perform better on a wider range of samples. The requirement to use an efficient graphics processor with CUDA cores in its architecture is the main drawback of the practical application of the created software solution.

Table 1 – Time to Obtain Face Metrics

	0	1	2	3	4	5	6	7	8	...	479	Σ
HOG	1,89	0,04	0,04	0,34	0,75	0,22	0,08	0,08	0,08	...	0,03	21,77
CNN	1,21	0,03	0,03	0,03	0,69	0,85	0,09	0,08	0,08	...	0,03	20,88

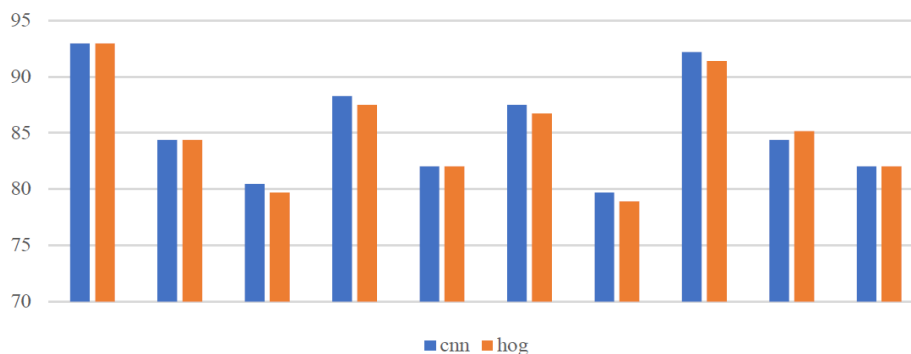


Fig. 1. Comparative diagram of the probability of match of faces

Conclusions

The following results were obtained in this paper:

1. A methods for generating cryptographic keys from biometric images using convolutional neural

networks and histogram of oriented gradients was proposed.

2. Conclusions of methods comparison by probability and performance were presented.

REFERENCE

1. Dodis, Y., Reyzin, L. and Smith, A. (2004), "Fuzzy Extractors: How to Generate Strong Keys from Biometrics and Other Noisy Data", *Advances in Cryptology - EUROCRYPT 2004*, Berlin, pp. 523–540. doi: https://doi.org/10.1007/978-3-540-24676-3_31.
2. Boyen, X. (2004), "Reusable cryptographic fuzzy extractors", *Proceedings of the 11th ACM conference on Computer and communications security*, New York, NY, USA, Oct. 2004, pp. 82–91, doi: <https://doi.org/10.1145/1030083.1030096>.
3. Álvarez, F. H. and Encinas, L. H. (2009), "Security Efficiency Analysis of a Biometric Fuzzy Extractor for Iris Templates", *Computational Intelligence in Security for IS*, Berlin, pp. 163–170. doi: https://doi.org/10.1007/978-3-642-04091-7_20.
4. Uludag, U., Pankanti, S., Prabhakar, S. and Jain, A. K. (2004), "Biometric cryptosystems: issues and challenges," *Proceedings of the IEEE*, vol. 92, no. 6, pp. 948–960, Jun. 2004, doi: <https://doi.org/10.1109/JPROC.2004.827372.Liu>
5. Jin, Z., A. Teoh, B. J., Goi, B.-M. and Tay, Y.-H. (2016), "Biometric cryptosystems: A new biometric key binding and its implementation for fingerprint minutiae-based representation", *Pattern Recognition*, vol. 56, pp. 50–62, Aug. 2016, doi: <https://doi.org/10.1016/j.patcog.2016.02.024>.
6. Fuller, B., Reyzin, L. and Smith, A. (2014), *When are Fuzzy Extractors Possible?* Available at: <http://eprint.iacr.org/2014/961>.
7. Lutsenko, M., Kuznetsov, A., Kiian, A., Smirnov, O. and Kuznetsova, T. (2021), "Biometric Cryptosystems: Overview, State-of-the-Art and Perspective Directions," *Advances in Information and Communication Technology and Systems*, Cham, pp. 66–84. doi: https://doi.org/10.1007/978-3-030-58359-0_5.
8. Schneier, B. (1996), *Applied cryptography : protocols, algorithms, and source code in C*, Wiley, New York, available at: http://archive.org/details/appliedcryptogra00schn_328.
9. Menezes, A. J., van Oorschot, P. C., Vanstone, S. A., van Oorschot, P. C. and Vanstone, S. A. (2018), *Handbook of Applied Cryptography*. CRC Press, doi: <https://doi.org/10.1201/9780429466335>.
10. Klima, R. E. and Sigmon, N. P. (2018), *Cryptology: Classical and Modern*, Chapman and Hall/CRC, doi: <https://doi.org/10.1201/9781315170664>.
11. Rubinstein-Salzedo, S. (2018), *Cryptography*, Cham: Springer International Publishing, 2018. doi: <https://doi.org/10.1007/978-3-319-94818-8>.
12. Delfs, H. and Knebl, H. (2015), *Introduction to Cryptography*, Springer, Berlin, doi: <https://doi.org/10.1007/978-3-662-47974-2>.
13. Amin, R., Gaber, T., ElTaweel, G. and Hassanien, A. E. (2014), "Biometric and Traditional Mobile Authentication Techniques: Overviews and Open Issues", *Bio-inspiring Cyber Security and Cloud Services: Trends and Innovations*, A. E. Hassanien, T.-H. Kim, J. Kacprzyk, and A. I. Awad, Eds. Springer, Berlin, pp. 423–446. doi: https://doi.org/10.1007/978-3-662-43616-5_16.
14. Jain, A. K., Ross, A. and Prabhakar, S. (2004), "An introduction to biometric recognition", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, Jan. 2004, doi: <https://doi.org/10.1109/TCSVT.2003.818349>.
15. Clarke, N. L. and Furnell, S. M. (2007), "Advanced user authentication for mobile devices", *Computers & Security*, vol. 26, no. 2, pp. 109–119, Mar. 2007, doi: <https://doi.org/10.1016/j.cose.2006.08.008>.
16. Geitgey, A. (2021), *ageitgey/face_recognition*, available at: https://github.com/ageitgey/face_recognition.
17. King, D. E. (2015), "Max-Margin Object Detection", *arXiv:1502.00046 [cs]*, available at: <http://arxiv.org/abs/1502.00046>.

Received (Надійшла) 14.03.2023

Accepted for publication (Прийнята до друку) 24.05.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Даченко Сергій Сергійович – магістрант кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Serhii Datsenko – master's degree student at Department of Computer Engineering and Programming, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine.
e-mail: sergdacenko@gmail.com; ORCID ID: <https://orcid.org/0000-0001-9514-0433>.

Кучук Георгій Анатолійович – доктор технічних наук, професор, професор кафедри комп'ютерної інженерії та програмування, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;

Heorhii Kuchuk – Doctor of Technical Sciences, Professor, Professor of the Department of Computer Engineering and Programming, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine.
e-mail: kuchuk56@ukr.net; ORCID ID: <https://orcid.org/0000-0002-2862-438X>.

Біометрична автентифікація, що використовує згорткові нейронні мережі

С. С. Даченко, Г. А. Кучук

Анотація. Актуальність. Криптографічні алгоритми та протоколи є важливими інструментами сучасної кібербезпеки. Вони використовуються в різних додатках, від простого програмного забезпечення для шифрування комп'ютерної інформації до складних інформаційних і телекомунікаційних систем, які реалізують різні електронні довірчі служби. Розробка повних біометричних криптографічних систем дозволить використовувати персональні біометричні дані як унікальний секретний параметр замість необхідності запам'ятовувати криптографічні ключі або використовувати додаткові пристрої автентифікації. **Об'єкт дослідження** – процес генерації криптографічних ключів з біометричних зображень обличчя людини з реалізацією нечітких екстракторів. **Предмет дослідження** – засоби та методи побудови нейронної мережі з використанням сучасних технологій. **Метою даної статті** є дослідження нових методів генерації криптографічних ключів із біометричних зображень за допомогою згорткових нейронних мереж та гістограм орієнтованих градієнтів. **Результати дослідження.** Запропонована технологія дозволяє реалізувати новий криптографічний механізм – технологію генерації надійних криптографічних паролів з біометричних зображень для подальшого використання їх як атрибутів доступу до захищених систем, а також джерела ключів для існуючих криптографічних алгоритмів.

Ключові слова: біометричні криптографічні системи; криптографічні ключі; нечіткі екстрактори; згорткова нейронна мережа.

Dmytro Salnikov, Dmytro Karaman, Viktoriia Krylova

National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine

HIGHLY RECONFIGURABLE SOFT-CPU BASED PERIPHERAL MODULES DESIGN

Abstract. Research motivation. When developing microcontrollers, manufacturers try to include as many different types of peripherals as possible in order to increase the marketing attractiveness of their products. On the one hand, with a large assortment of various peripheral modules, it is very difficult to implement several devices of the same type in the microcontroller: manufacturers are mainly limited to 1-2 instances, in rare cases 4 modules of the same type are included. On the other hand, most software projects do not use all the peripherals of modern microcontrollers and many devices are left unused, while there may be a shortage of other types of modules. Another problem that has become especially noticeable for microcontrollers used in the field of IoT is the cryptographic protection of data that is transmitted through built-in information exchange interfaces. The main efforts of researchers and developers of cryptographic data protection methods were aimed at reducing energy-intensive operations, memory access iterations and speeding up encryption processes while maintaining a high level of cryptographic protection and enabling efficient data distribution within IoT devices networks. **Research results.** This paper presents an alternative approach to the manufacture of peripheral modules as part of microcontrollers. The authors propose to use a configurable software processor module based on the MIPS architecture with a reduced instruction set and limited capabilities. **Conclusions.** This approach would make it possible to dynamically change the functionality of peripheral modules in accordance with the requirements of the developed software solution, which in turn will increase the efficiency of the microcontroller chips capabilities utilization. In addition, the transfer of data stream encryption functions to the reconfigurable core of the peripheral module will provide fast and transparent cryptographic protection, as well as allow offloading the microcontroller core and increasing the energy efficiency of chips while reducing their production cost.

Keywords: peripheral module; soft-CPU; RISC; MIPS architecture; FPGA; internet of things; encryption, lightweight cryptography; AEAD; Ascon.

Introduction

Modern microcontrollers must have the widest possible capabilities so that devices based on them can meet the requirements of users. To do this, microcontroller manufacturers are constantly improving the computing cores in microcontrollers, as well as increasing the number and expanding the functionality of peripheral modules. However, this approach leads to a narrowing of promising areas of application of microcontrollers, especially from electronic devices with intelligent capabilities for everyday life, as part of smart home systems, as well as distributed systems with stringent requirements for energy efficiency: wireless sensor networks, devices as part of the Internet of Things (IoT) etc.

At the moment, the problem of the efficiency of the use of energy resources is one of the most important on the European continent and even in the world. At the same time, humanity uses millions of electrical devices that use microcontrollers to deliver specific features to their users.

Internet of things devices are widely used in most modern homes in the form of climate control systems, safety or security systems, voice-controlled assistants etc. The number of such devices continues to grow. Thus, each device's efficiency impacts energy usage.

Despite the relatively low level of energy consumption of each such device, the total costs for millions of sold devices are significant, and optimizing the energy consumption of each of them is an important task of modern science and technology.

Another feature that has become relevant recently for distributed embedded systems, especially for IoT, is ensuring the security, integrity and confidentiality of transmitted information, as well as providing protection from unauthorized access. Implementations of traditional

encryption, integrity and confidentiality methods are very resource intensive. During their development, first of all, the issues of reliability and resistance to cryptanalytic attacks were considered, rather than the possibility of use on platforms with limited computing capabilities and limited power supplies. That is why recently there have been many attempts to adapt existing solutions to the operating conditions of IoT devices, as well as to search for new algorithms that initially take into account the limitations of the IoT platform elements and provide a level of protection no less than traditional cryptographic methods.

Relevance of the topic and related works

The design of electronic devices is directly related to the choice of optimal microprocessor used to implement all planned functional capabilities of the device. Modern microcontrollers use different architectures of the processor core and have a different set of peripheral units that can be configured to perform various tasks. Selection of an appropriate architecture and a set of such peripheral blocks affects a possibility to expand and/or improve the functional capabilities of the device. With this in mind, electronics and software engineers often use overpowered devices to avoid problems in the future and let the project grow without significant printed circuit boards and software changes.

Most of the changes in the software can be related to the lack of peripheral modules of the required type or the inability to configure them to work in the required mode.

Problems of this kind can be solved by including CPLD or FPGA configurable blocks in the circuit [1, 2, 3]. In this case, the programmer can include the implementation of the necessary peripheral modules in the Verilog or VHDL hardware description languages as part

of the project. Which requires additional programming skills and, more importantly, significantly increases the cost of such solutions and the process of working with them [4]. Solutions with embedded CPU and FPGA in a single chip are available from most semiconductor manufacturers like Intel FPGA, AMD Xilinx, Analog Devices [5, 6], etc. Another option is highly configurable hybrid ICs like Cypress microcontrollers [7].

Flexible combination of the embedded microprocessor functions core (CPU) and the programmable part (FPGA) allows developers, if necessary, to expand the functionality of the designed solution by transferring complex and intensive operations for the CPU to be executed in a separate specialized module in the FPGA part. In the same way, you can expand the pool of available peripherals, since the FPGA part in most cases has access to the lines and I/O ports of the chip.

The issue of creating flexible auxiliary peripheral modules has been considered for the past decade. For example, [8] studies the possibility of expanding the capabilities of the PIC microcontroller by connecting external programmable peripherals to general purpose I/O ports. This solution allows you to add the missing peripheral module, but occupies one of the few input/output ports of the microcontroller. In addition, this approach requires the involvement of external additional hardware resources.

The use of hardware emulation through software is proposed in [9] that allows you to recreate the behavior of the I2C serial bus controller through the GPIO ports in a microcontroller based on the RISC-V core. Thus, the expansion of the functionality of the microcontroller without the involvement of additional equipment is provided. However, in this case, the kernel receives an additional load, since it must spend additional processor time emulating the I2C transceiver. In addition, questions arise regarding the performance parameters of such a solution. Definitely, such a module will not be able to work with devices on the bus at high data exchange rates.

A specific way to use custom peripheral modules is discussed in [10]. The authors create digital hardware twins of peripheral modules using FPGA, which make it easier for students to learn embedded systems in distance education. Access to the modules is provided via the Internet using a client-server architecture.

Another example of a remote peripheral module is discussed in [11]. A method for remote control of microcontroller input/output interfaces using WebUSB technology is proposed.

Thus, the problem of the efficient, easy to use and low-cost architecture for microcontroller peripheral unit implementation is a relevant topic to research and development nowadays.

The issues of security and data protection for elements of the Internet of Things (IoT) system have become especially acute since the concept was adopted by leading electronics manufacturers and began to be massively introduced into various fields of human activity. The methods and technologies that initially relied on that concept of devices networking did not assume that devices would be available to a wide range

of users, among which there could be many intruders.

Numerous attempts to adapt existing methods of cryptographic protection and authentication have shown that in order to implement these methods in systems with limited resources, it is necessary either to take resources from other functions in the IoT system, or to curtail the capabilities or even modify the main algorithms [12], which can lead to a decrease in cryptographic resilience or the emergence of vulnerabilities.

This state of affairs led to the need to develop a separate class of cryptographic protection methods – lightweight cryptography [13]. These methods had to take into account the features of distributed embedded systems: reduced power supply and low level of performance, and at the same time provide a level of protection that is not inferior to traditional methods.

Authenticated Encryption with Attached Data (AEAD) [14] occupies a special place among the methods of lightweight cryptography. Their peculiarity lies in the fact that not the entire message is encrypted, but only a part of it with the most sensitive data, while the entire message is authenticated. This allows you to ensure data protection and message integrity, while the data necessary for successful and efficient routing remains open.

In February 2023, the US National Institute of Standards and Technology (NIST) announced the results of a competition to select a lightweight cryptography algorithm [15] that will form the basis of the corresponding standard. As a result of many years of thorough selection among 57 applicants, the family of cryptographic algorithms Ascon [16] was chosen. The implementation of the block cipher algorithm, which is part of this family, is discussed in this article.

Soft-CPU-based peripheral module

From the authors' point of view, the complexity of implementing peripheral modules in hardware significantly reduces the possibilities of their usage, and most companies prefer usage-ready solutions that do not require additional development process.

At the same time, most programmers involved in the creation of microcontroller-based systems have an understanding of the processes and operations which take place in peripheral units and can create their grammatical description using C/C++ or assembler language.

Worthless to say that availability of mature programming development tools, in particular, a C language compiler and/or an assembly parser, plays an even more important role in the life cycle of any software project.

Thus, it is promising to use general purpose CPU architectures without additional license conditions, mature development tools and well-known instruction sets, such as MIPS, RISC-V etc.

At the current stage, a software processor that uses MIPS architecture has been chosen to construct the experimental architecture. Issues of the efficiency of MIPS/RISC-V software processors and development for such systems on chip are widely considered in the literature, particularly in [17].

Basic architecture of MIPS soft processor was modified for this experiment. Modules for working with data memory were removed. Registers \$a0-\$a3 were used to load data from the queue. Registers \$s0-\$s7 – for setting the output GPIO lines to the required state. Other registers are available to the programmer for intermediate calculations.

For such a system, basic UART module can be executed in the form of the following program assembly language (Listing 1).

```

_start:
li      $t1, 10      # set bit counter st+8d+sp
move    $t0, $a0     # load data from queue to tx buffer
andi    $t0, $t0, 0xFF # cut data to 1 byte with mask
sll     $t0, $t0, 1  # prepare start bit
or      $t0, $t0, 0x01
loops:
move    $s0, $t0     # copy to s0.0 bit data to tx from buf
jal     baud_delay
srl     $t0, $t0, 1  # prepare next bit to tx
subu    $t1, $t1, 1  # count bits transmitted
bgtz   $t1, loops   # test if tx complete
# rest of code
nop
j_start
baud_delay:
# delay routine according to baudrate
jr     $ra
    
```

Listing 1

The process of designing and manufacturing ICs is overcomplicated and involves significant workforce requirements with licensing expenses. It is not possible for small IT companies to plan a set of required peripherals and produce a chip or ASIC which matches needs of a project exactly. It increases time-to-market of the device, expenses and complexity of any device. Moreover, there is no way to deal with changes in requirements. It is not possible to reconfigure such chips in future. A possible way of solving this problem presented in [8], while our proposal is more generic and configurable.

Implementation of some peripheral modules consume significant hardware resources. This affects chip cost and energy consumption. Thus, widely adopted practice is to use cheaper devices and implement required peripheral in software using general purpose input/output pins. An example of such a study is presented in [9]. It provides software implementation of I2C peripheral.

Proposed in this work, MIPS-based block, is 3 times smaller in terms of consumed FPGA resources.

Another field where highly reconfigurable peripherals are needed is a study laboratory. Any university in the world now has remote laboratory classes and requires students to perform experiments and laboratory assignments from home. A solution that uses FPGA logic to mimic some type of hardware is proposed in [10, 11, 19]. In addition, hardware emulated using FPGA can be used to improve testing methods [18], and to measure and characterize network architectures [19]. Our approach of realization of peripheral blocks may simplify implementation of generic execution units suitable for remote/concurrent usage.

We can't analyze hardware expenses to create specific peripheral blocks in ASIC. It depends on the technical process of the manufacturer and a lot of design choices of the engineer. We assume that ASIC resource consumption is compatible with FPGA resources consumed to implement such blocks. Mentioned blocks and their configuration can be seen in [20].

In Table 1 one can see resource consumption for widely used Intel FPGA peripheral IP blocks. Note that some of them use additional memory which is not used to implement register memory maps, queues or FIFO buffers. Thus, additional memory is required to implement an interface to control them.

To verify the solutions proposed in the article, we used a 5CSEBA6U23I7 device of Intel Cyclone V SoC with 2-core ARM Cortex A9 FPGA.

Table 1 – FPGA Resources consumption for different types of IP cores

	Registers	ALM blocks	Block memory bits
Modified MIPS core	74	72	0
Intel SPI IP Core	56	23	0
Intel 16550 UART IP Core	156	114	0
Intel UART IP Core	0	1	0
Intel I2C Host IP Core	215	143	40

In addition to peripheral blocks, we provide resources consumed to synthesize a modified MIPS processor suitable for software implementation of the listed peripherals. The structure of the synthesized CPU is show on Fig. 1.

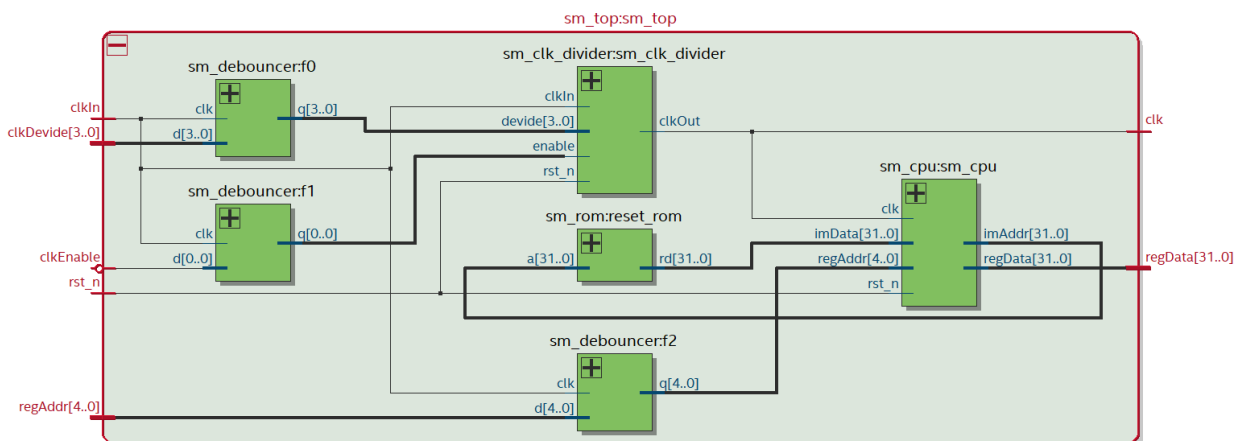


Fig. 1. Structure of the synthesized MIPS core

Modern CPU implementations use pipelining to allow computing cores to work faster by using higher frequencies. The problem is that on high frequencies timing delay on logic used to implement CPU becomes significant. Such delay limits frequency and, as a result, CPU performance. Thus, adding a pipeline (e.g., save intermediate states of signals to registers) to the design allows to improve capabilities of the CPU.

Moreover, it is a common practice to include hazard units to CPU implementations to allow pipelining of the CPU and achieve higher performance. In contradiction to

common CPUs, reconfigurable peripheral modules do not have a requirement to run with frequencies higher than 30-40MHz. Such relaxation of the requirements allows to drop these modules and significantly reduce resources required to implement peripheral module.

As most peripheral blocks of modern microcontrollers the suggested implementation requires additional FIFO buffers and input/output multiplexers to achieve efficient signal transmitting/receiving. Such logic can be connected to the registers of the peripheral module as shown on Fig. 2.

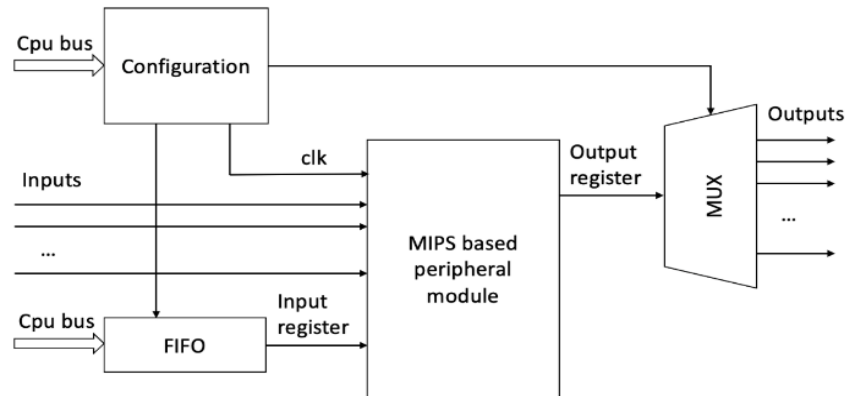


Fig. 2. Additional logic needed to used proposed modules

In this article we don't compute resources that required to implement such buffers and interconnect. In most cases resources spent to synthesize such logic is equivalent to the resources needed for functioning of regular peripheral block.

Implementing Encryption Functions in the Peripheral Module

One of the modern requirements for devices that are oriented to work as part of the IoT is to support a certain stack of cryptographic data protection methods: block or stream encryption, various ciphertext block chaining modes, hash sum calculation, message authentication mechanism (MAC) and even digital signature generation. In most cases, the necessary protection mechanisms are implemented at the software level and are executed by the main computing core of the microcontroller. Peripheral modules are not involved in the process in any way and perform only the functions of receiving, transmitting and transport integrity control of data packets.

Much less common are devices in which cryptographic functions or their composite operations are implemented at the level of the computing core (hardware support for cryptographic operations). However, despite the excellent speed of the encryption and authentication processes, the use of such tools faces significant limitations: high price and export control by government agencies.

Since the core of the MIPS soft processor was used in the development of the reconfigurable peripheral module, the possibility of transferring the execution of encryption functions from the main core of the microcontroller to the soft processor core of the peripheral module was considered.

Such a solution will allow to unload the main core of the microcontroller and making the encryption process transparent: in the main program, it will only be enough to send data for transmission and read the received data, the peripheral module will automatically perform encryption during transmission and decryption during reception. In the main program code, it is only necessary to provide for the process of initial initialization of encryption functions in the peripheral module (loading keys, initialization vectors, mode selection), which is performed when the system is initialized after switching on or by special request.

As an example, the family of lightweight algorithms Ascon [16] was considered. It includes an authenticated encryption algorithm and a hash function calculation algorithm. Both algorithms use a transformation called the sponge function. The authors of the algorithm set offer a wide range of software implementations for various microprocessor architectures with different degrees of optimization in public repository on GitHub.

The implementation of the entire set of algorithms by means of a peripheral module does not seem appropriate, since after compilation, the entire code, taking into account optimizations, occupies at least 6456 bytes in memory. Therefore, it is proposed to perform all operations for initialization and preparation of the encryption process in the main core of the microcontroller, and to transfer the initial state for the encryption or decryption functions to the peripheral module.

An estimate of the size of the executable code is given in Table 2. For estimation, a general implementation version of the algorithm was compiled from the authors' repository on GitHub, ascon128v12 version for 32-bit processors with instructions on integer

operands (bi32). The compilation was done on mips gcc ver. 12.2.0, emulation and debugging were performed on the MIPS32r5 generic kernel simulator.

Table 2 – Estimation of the code size of the encryption system

Implemented function	Code size, bytes
Full implementation of the algorithm stack	6456
Implementation of encryption and decryption (auxiliary code)	5344 (3928)
Encryption function only (including auxiliary code)	4544
Decryption function only (including auxiliary code)	4720

An example of the previously discussed program for the operation of the UART module with the implementation of preliminary encryption of the transmitted data is shown below (Listing 2).

Conclusions

The use of processor architectures with reduced command and pipeline functionality is seen as a profitable replacement for traditional configurable peripheral modules that are widely used at the moment.

The FPGA implementation of software-based MIPS core, modified to be lightweight from the recourse usage perspective, shows that such design consumes comparable, with most peripheral blocks, among FPGA resources.

Thus, microcontrollers with 10 to 20 such modules may be a competitive replacement to widely used solutions.

```

_start:
move    $t0, $a0    # load plaintext from queue to buffer
xor     $t0, $t0, $t4    # plain text being xored with
                        # state block to get ciphertext
jal     ascon_encrypt    # call encryption function -
                        # permutations of state block

li      $t1, 10    # set bit counter st+8d+sp
andi   $t0, $t0, 0xFF    # cut data to 1 byte with mask
sll    $t0, $t0, 1    # prepare start bit
or     $t0, $t0, 0x01

loops:
move    $s0, $t0    # copy to s0.0 bit data to tx from buf
jal     baud_delay
srl    $t0, $t0, 1    # prepare next bit to tx
subu   $t1, $t1, 1    # count bits transmitted
bgtz   $t1, loops    # test if tx complete
# rest of code
nop
j      _start
baud_delay:
# delay routine according to baudrate
jr     $ra

```

Listing 2

The transfer of encryption functions for transmitted/received data from the main core program of the microcontroller to the core of the soft-processor allows to significantly reduce load of the main core, increase performance, and also make the processes of ensuring the protection and authentication of data transmitted by IoT devices transparent to user software. As the results of the implementation showed, the transfer of all encryption functions to the microcode that is executed in the peripheral module can lead to a high consumption of a very limited amount of soft processor memory and reduce the performance of the peripheral module. In this regard, the issue of implementing hardware support for encryption at the level of assembler commands of a soft-processor is considered.

REFERENCES

- Liu, C., Liu Q. and Cheng L. (2011), "CPLD based MCU coprocessor design and experiment platform", *2011 Int. Conf. on Electronics, Communications and Control*, Ningbo, China, 2011, pp. 1365-1368, doi: <https://doi.org/10.1109/ICECC.2011.6066408>.
- Schiavone, P.D., Rossi, D., Mauro, A. Di, Gürkaynak, F.K., Saxe, T., Wang, M., Yap, K.C. and Benini, L. (2021), "Arnold: An eFPGA-Augmented RISC-V SoC for Flexible and Low-Power IoT End Nodes", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 29, no. 4, pp. 677-690, April 2021, doi: <https://doi.org/10.1109/TVLSI.2021.3058162>.
- Matsumura, T., Okada, N., Kawamura, Y., Nii, K., Arimoto, K., H. Makino and Y. Matsuda (2014), "The LSI implementation of a memory based field programmable device for MCU peripherals", *17th Int. Symposium on Design and Diagnostics of Electronic Circuits & Systems*, Warsaw, Poland, pp. 183-188, doi: <https://doi.org/10.1109/DDECS.2014.6868787>.
- Amano, H., Abe, S., Hasegawa, Y., Deguchi, K. and Suzuki, M. (2005), "Performance and cost analysis of time-multiplexed execution on the dynamically reconfigurable processor", *Proc. IEEE Symposium on Field-Programmable Custom Computing Machines*, pp. 315-316, doi: <https://doi.org/10.1109/FCCM.2005.52>.
- (2023), *Intel® Products / Intel® FPGAs and SoC FPGAs*, available at: <https://www.intel.com/content/www/us/en/products/details/fpga.html>.
- (2023), *AMD Xilinx Products / Adaptive SoCs: ZYNQ and Versal*, available at: <https://www.xilinx.com/products/silicon-devices/soc.html>.
- (2023), *Cypress PSoC® 6 Microcontrollers Purpose-Built for the Internet of Things*, available at: https://www.infineon.com/dgdl/Infineon-PSoC_6_MCU_The_New_Standard_for_the_Internet_of_Things-ProductBrochure-v05_00-EN.pdf?fileId=8ac78c8c7d0d8da4017d0f64f95450c7.
- Penteado, C.G. and Moreno, E.D. (2009), "A Specialized Processor for Emulating Peripherals of the PIC Microcontroller", *IEEE Latin America Transactions*, vol. 7, no. 2, pp. 133-140, June 2009, doi: <https://doi.org/10.1109/TLA.2009.5256820>.
- Molina-Robles, R., García-Ramírez, R., Chacón-Rodríguez, A., Rimolo-Donadio, R. and Arnaud, A. (2021), "Low-level algorithm for a software-emulated I²C I/O module in general purpose RISC-V based microcontrollers", *2021 IEEE URUCON*, Montevideo, Uruguay, pp. 90-94, doi: <https://doi.org/10.1109/URUCON53396.2021.9647309>.
- Buysse, L., Van den Broucke, Q., Verslype, S., Peuteman, J., Boydens, J. and Pissoort, D. (2021), "FPGA-based digital twins of microcontroller peripherals for verification of embedded software in a distance learning environment", *2021 XXX Int. Scientific Conference Electronics (ET)*, Sozopol, Bulgaria, pp. 1-4, doi: <https://doi.org/10.1109/ET52713.2021.9579770>.
- Huang, L. and Shu, Y. (2022), "Design and Research of Microcontroller I/O Control Technology", *2022 IEEE 4th International Conference on Power, Intelligent Computing and Systems (ICPICS)*, Shenyang, China, 2022, pp. 263-266, doi: <https://doi.org/10.1109/ICPICS55264.2022.9873583>.

12. Kane, L. E., Chen, J. J., Thomas, R., Liu, V. and Mckague, M. (2020), "Security and Performance in IoT: A Balancing Act", *IEEE Access*, vol. 8, pp. 1219.69-1219.86, doi: <https://doi.org/10.1109/ACCESS.2020.3007536>.
13. Biryukov, A. and Perrin, L. (2017), "State of the art in lightweight symmetric cryptography", *Cryptology ePrint Archive*, Nov. 2017, available at: <https://eprint.iacr.org/2017/511.pdf>.
14. Jutla, C. S. (2008), "Encryption Modes with Almost Free Message Integrity", *Journal of Cryptology*, vol. 21, pp. 547–578, doi: <https://doi.org/10.1007/s00145-008-9024-z>.
15. (2023), *NIST Selects 'Lightweight Cryptography' Algorithms to Protect Small Devices*, National Institute of Standards and Technology (NIST) Website, February 07, 2023, available at: <https://www.nist.gov/news-events/news/2023/02/nist-selects-lightweight-cryptography-algorithms-protect-small-devices>.
16. Dobraunig, C., Eichlseder, M., Mendel, F. and Schl affer, M. (2023), *Ascon: Lightweight Authenticated Encryption & Hashing*, available at: <https://ascon.iaik.tugraz.at/index.html>.
17. Dewangan, G. K., Prasad, G. and Mandi, B.C. (2021), "Design and Implementation of 32 bit MIPS based RISC Processor", *2021 8th International Conference on Signal Processing and Integrated Networks (SPIN)*, Noida, India, 2021, pp. 998-1002, doi: <https://doi.org/10.1109/SPIN52536.2021.9566007>.
18. Li, T. and Liu, Q. (2016), "Cost effective partial scan for hardware emulation", *2016 IEEE 24th Annual Int. Symposium on Field-Programmable Custom Computing Machines (FCCM)*, pp. 131-134, doi: <https://doi.org/10.1109/FCCM.2016.39>.
19. Khamis, M., El-Ashry, S., Shalaby, A., AbdElsalam M. and El-Kharashi M. W. (2018), "A configurable risc-v for noc-based mpocs: A framework for hardware emulation", *2018 11th International Workshop on Network on Chip Architectures (NoCArc)*, pp. 1-6, doi: <https://doi.org/10.1109/NOCARC.2018.8541158>.
20. (2021), *Embedded Peripherals IP User Guide for Quartus Prime 21.4*, Intel, UG-01085, available at: <https://www.intel.com/content/www/us/en/docs/programmable/683130/21-4/introduction.html>.

Received (Надійшла) 24.02.2023

Accepted for publication (Прийнята до друку) 10.05.2023

ВІДОМОСТІ ПРО АВТОРІВ / ABOUT THE AUTHORS

Сальніков Дмитро Валентинович – кандидат технічних наук, старший викладач кафедри автоматизації та управління в технічних системах, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Dmytro Salnikov – Candidate of Technical Sciences, Senior Lecturer of the Department of automation and control in technical systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine.
e-mail: dmytro.salnikov@khp.edu.ua; ORCID ID: <http://orcid.org/0000-0007-6201-5370>.

Караман Дмитро Григорович – старший викладач кафедри автоматизації та управління в технічних системах, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Dmytro Karaman – Senior Lecturer of the Department of automation and control in technical systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine.
e-mail: dmytro.karaman@khp.edu.ua; ORCID ID: <http://orcid.org/0000-0002-7252-3172>.

Крилова Вікторія Анатоліївна – кандидат технічних наук, доцент кафедри автоматизації та управління в технічних системах, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Viktoriia Krylova – Candidate of Technical Sciences, Associate Professor of the Department of automation and control in technical systems, National Technical University «Kharkiv Polytechnic Institute», Kharkiv, Ukraine.
e-mail: vika.hpi@outlook.com; ORCID ID: <http://orcid.org/0000-0002-4540-8670>.

Архітектура гнучко конфігурованих периферійних модулів на базі програмованих процесорних ядер

Д. В. Сальніков, Д. Г. Караман, В. А. Крилова

Анотація. Мотивація дослідження. При розробці мікроконтролерів виробники намагаються включити якнайбільше різних видів периферійних пристроїв, щоб підвищити маркетингову привабливість своєї продукції. З одного боку, при великому асортименті різних периферійних модулів дуже складно включити до складу мікроконтролера велику кількість пристроїв одного типу: виробники переважно обмежуються 1-2 екземплярами, дуже рідко зустрічаються 4 модулі однакового типу. З іншого боку, більшість програмних проектів не використовують всю периферію сучасних мікроконтролерів і багато пристроїв залишаються незадіяними, тоді як може відчуватися нестача модулів іншого типу. Ще однією проблемою, яка стала особливо відчутною для мікроконтролерів, що застосовуються у сфері ІТ, це криптографічний захист даних, які передаються через вбудовані інтерфейси обміну інформацією. Основні зусилля дослідників та розробників криптографічних методів захисту даних були спрямовані на зниження енерговитратних операцій, звернень до пам'яті та прискорення процесів шифрування за одночасного збереження високого рівня криптографічного захисту та забезпечення можливості ефективного поширення даних у мережах пристроїв IoT.
Результати дослідження. У роботі подано альтернативний підхід до виготовлення периферійних модулів у складі мікроконтролерів. Пропонується використовувати конфігурований модуль програмного процесора на базі архітектури MIPS з укороченим набором команд та обмеженими можливостями. **Висновки.** Такий підхід дозволяє динамічно змінювати функціонал периферійних модулів відповідно до вимог програмного рішення, що розробляється, що в свою чергу дозволить підвищити ефективність використання можливостей мікросхем мікроконтролерів. Крім того, перенесення функцій шифрування потоку даних в ядро периферійного модуля, що реконфігурується, дозволить забезпечити швидкий і прозорий криптографічний захист, а також дозволить розвантажити ядро мікроконтролера і підвищити енергоефективність мікросхем при одночасному зниженні собівартості їх виробництва.

Ключові слова: периферійний модуль; софт-процесор; RISC; архітектура MIPS; ПЛІС; інтернет речей; шифрування; легка криптографія; AEAD-режим, Ascon.