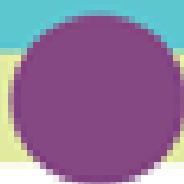


ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СУСПІЛЬСТВО

INFORMATION TECHNOLOGY AND SOCIETY



МІЖРЕГІОНАЛЬНА АКАДЕМІЯ УПРАВЛІННЯ ПЕРСОНАЛОМ
INTERREGIONAL ACADEMY OF PERSONNEL MANAGEMENT

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СУСПІЛЬСТВО

INFORMATION TECHNOLOGY AND SOCIETY

**ВИПУСК 2
ISSUE 2**

2021



**Видавничий дім
«Гельветика»
2021**

*Рекомендовано до друку Вченою радою
Міжрегіональної Академії управління персоналом
(протокол № 4 від 30 серпня 2021 року)*

Інформаційні технології та суспільство / [головний редактор О. Попов]. – Київ: Міжрегіональна Академія управління персоналом, 2021. – Випуск 2. – 94 с.

Журнал «Інформаційні технології та суспільство» є науковим рецензованим виданням, в якому здійснюється публікація матеріалів науковців різних рівнів у вигляді наукових статей з метою їх поширення як серед вітчизняних дослідників, так і за кордоном.

Редакційна колегія не обов'язково поділяє позицію, висловлену авторами у статтях, та не несе відповідальності за достовірність наведених даних і посилань.

Головний редактор: Попов О. О. – член-кор. НАН України, д-р техн. наук, с.н.с., заступник директора з науково-організаційної роботи, Державна установа «Інститут геохімії навколишнього середовища Національної академії наук України».

Редакційна колегія:

Василенко М. Д. – д-р фіз.-мат. наук, проф., професор кафедри кібербезпеки, Національний університет «Одеська юридична академія»; **Горбов І. В.** – канд. техн. наук, с.н.с., старший науковий співробітник, Інститут проблем реєстрації інформації НАН України; **Дуднік А. С.** – д-р техн. наук, доц., доцент кафедри мережевих та інтернет технологій, Київський національний університет імені Тараса Шевченка; **Євсєєв С. П.** – д-р техн. наук, професор кафедри кібербезпеки та інформаційних технологій факультету інформаційних технологій, Харківський національний економічний університет імені Семена Кузнеця; **Зибін С. В.** – д-р техн. наук, доц., завідувач кафедри інженерії програмного забезпечення, Національний авіаційний університет; **Кавун С. В.** – д-р екон. наук, канд. техн. наук, проф., завідувач кафедри комп'ютерних інформаційних систем та технологій Інституту комп'ютерно-інформаційних технологій, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом»; **Комарова Л. О.** – д-р техн. наук, с.н.с., директор Навчально-наукового інституту інформаційної безпеки та стратегічних комунікацій, Національна академія Служби безпеки України; **Мілов О. В.** – д-р техн. наук, професор кафедри кібербезпеки та інформаційних технологій факультету інформаційних технологій, Харківський національний економічний університет імені Семена Кузнеця; **Охріменко Т. О.** – канд. техн. наук, старший науковий співробітник науково-дослідної лабораторії протидії кіберзагрозам в авіаційній галузі, Національний авіаційний університет; **Рудніченко М. Д.** – канд. техн. наук, доц., доцент кафедри інформаційних технологій, Державний університет «Одеська політехніка»; **Фомін О. О.** – д-р техн. наук, доц., професор кафедри комп'ютеризованих систем управління, професор кафедри прикладної математики та інформаційних технологій, Державний університет «Одеська політехніка»; **Скура-товський Р. В.** – канд. фіз.-мат. наук, доц., доцент кафедри обчислювальної математики та комп'ютерного моделювання, Інститут комп'ютерно-інформаційних технологій, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом»; **Супрун О. М.** – канд. фіз.-мат. наук, доц., доцент кафедри програмних систем і технологій факультету інформаційних технологій, Київський національний університет імені Тараса Шевченка; **Табунщик Г. В.** – канд. техн. наук, проф., професор кафедри програмних засобів, Національний університет «Запорізька політехніка»; **Хохлячова Ю. С.** – канд. техн. наук, доц., доцент кафедри безпеки інформаційних технологій, Національний авіаційний університет; **Чолишкіна О. Г.** – канд. техн. наук, доц., директор Інституту комп'ютерно-інформаційних технологій та дизайну, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом»; **Юдін О. К.** – д-р техн. наук, проф., директор центру кібербезпеки Навчально-наукового інституту інформаційної безпеки та стратегічних комунікацій, Національна академія Служби безпеки України; **Гонє-єнко Віктор** – dr. sc. ing., проф., проректор з наукової роботи, директор навчальної програми магістратури «Комп'ютерні системи», Університет прикладних наук ISMA (Латвійська Республіка); **Leszczyna Rafal** – dr hab. inż., професор кафедри комп'ютерних наук у менеджменті, Гданський технологічний університет (Республіка Польща).

*Свідомство про державну реєстрацію друкованого засобу масової інформації
«Інформаційні технології та суспільство» Серія KB № 24815-14755P від 27.04.2021 р.*

Усі електронні версії статей журналу оприлюднюються на офіційній сторінці видання
<http://journals.maup.com.ua/index.php/it>

Статті у виданні перевірені на наявність плагіату за допомогою програмного забезпечення
StrikePlagiarism.com від польської компанії Plagiat.pl.

© Міжрегіональна Академія управління персоналом, 2021

© Авторів статей, 2021

***Recommended for publication
by Interregional Academy of Personnel Management
(Minutes No. 4 dated 30.08.2021)***

Information Technology and Society / [chief editor Oleksandr Popov]. – Kyiv: Interregional Academy of Personnel Management, 2021. – Issue 2. – 94 p.

Journal «Information Technology and Society» is a peer-reviewed scientific edition, which publishes materials of scientists of various levels in the form of scientific articles for the purpose of their dissemination both among domestic researchers and abroad.

Editorial board do not necessarily reflect the position expressed by the authors of articles, and are not responsible for the accuracy of the data and references.

Chief editor: Oleksandr Popov – Corresponding Member of NAS of Ukraine, Doctor of Engineering, Senior Research Scientist, Deputy Director for Scientific-Organizational Affairs, SI “Institute of Environmental Geochemistry of the National Academy of Sciences of Ukraine”.

Editorial Board:

Mykola Vasylenko – Doctor of Physics and Mathematics, Professor, Professor at the Department of Cybersecurity, National University «Odesa Law Academy»; **Ivan Horbov** – PhD in Engineering, Senior Research Associate, Senior Research Fellow, Institute for Information Recording of NAS of Ukraine; **Andrii Dudnik** – Doctor of Engineering, Associate Professor, Senior Lecturer at the Department of Networking and Internet Technologies, Taras Shevchenko National University of Kyiv; **Serhii Yevseiev** – Doctor of Engineering, Professor at the Department of Cybersecurity and Information Technologies of the Faculty of Information Technologies, Simon Kuznets Kharkiv National University of Economics; **Serhii Zybin** – Doctor of Engineering, Associate Professor, Head of the Department of Software Engineering, National Aviation University; **Serhii Kavun** – Doctor of Economics, PhD in Engineering, Professor, Head of the Department of Computer Information Systems and Technologies of the Institute of Computer Information Technologies, PJSC «HEI «Interregional Academy of Personnel Management»; **Larysa Komarova** – Doctor of Engineering, Senior Research Scientist, Laureate of State Prize, Director of Educational-Scientific Institute of Information Security and Strategic Communications, National Academy of the Security Service of Ukraine; **Milov Oleksandr** – Doctor of Engineering, Professor at the Department of Cybersecurity and Information Technologies of the Faculty of Information Technologies, Simon Kuznets Kharkiv National University of Economics; **Tetiana Okhrimenko** – PhD in Engineering, Senior Research Scientist at the Scientific Research Laboratory for Countering Aviation Cyberthreats, National Aviation University; **Mykola Rudnichenko** – PhD in Engineering, Associate Professor, Senior Lecturer at the Department of Information Technologies, Odessa Polytechnic State University; **Oleksandr Fomin** – Doctor of Engineering, Associate Professor, Professor at the Department of Computerized Control Systems, Professor at the Department of Applied Mathematics and Information Technologies, Odessa Polytechnic State University; **Ruslan Skuratovskyi** – PhD in Physics and Mathematics, Associate Professor, Senior Lecturer at the Department of Computational Mathematics and Computer Modeling, Institute of Computer Information Technology, PJSC «HEI «Interregional Academy of Personnel Management»; **Olha Suprun** – PhD in Physics and Mathematics, Associate Professor, Senior Lecturer at the Department of Software Systems and Technologies of the Faculty of Information Technologies, Taras Shevchenko National University of Kyiv; **Halyna Tabunshchuk** – PhD in Engineering, Professor, Professor at the Department of Software Tools, “Zaporizhzhia Polytechnic” National university; **Yuliia Khokhlachova** – PhD in Engineering, Associate Professor, Senior Lecturer at the Department of Information Technology Security, National Aviation University; **Olha Cholyshkina** – PhD in Engineering, Associate Professor, Director of the Institute of Computer Information Technologies and Design, PJSC «HEI «Interregional Academy of Personnel Management»; **Oleksandr Yudin** – Doctor of Engineering, Professor, Director of the Cybersecurity Center of the Educational-Scientific Institute of Information Security and Strategic Communications, National Academy of the Security Service of Ukraine; **Hopeienko Viktor** – dr. sc. ing., Professor, Vice Rector for Research, Director of the study programme “Computer systems”, ISMA University of Applied Sciences (Republic of Latvia); **Leszczyna Rafal** – dr hab. inż., Profesor, Katedra Informatyki w Zarządzaniu, Politechnika Gdańska (Republic of Poland).

*Print media registration certificate «Information Technology and Society»
series KV No. 24815-14755P dated 27.04.2021*

All electronic versions of articles in the collection are available on the official website edition
<http://journals.maup.com.ua/index.php/it>

The articles were checked for plagiarism using the software
StrikePlagiarism.com developed by the Polish company Plagiat.pl.

© Interregional Academy of Personnel Management, 2021
© Copyright by the contributors, 2021

ЗМІСТ

Андрій ДУДНІК

МОДЕЛІ ТА МЕТОДИ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ БЕЗПРОВОДОВИХ МЕРЕЖ
У СКЛАДІ КОМП'ЮТЕРИЗОВАНИХ СИСТЕМ ВИМІРЮВАННЯ МЕХАНІЧНИХ ВЕЛИЧИН.....6

Валерія КОВАЧ, Володимир КУЦЕНКО, Ірина МАРТИНЮК, Олександр КОВАЛЕНКО

ОСОБЛИВОСТІ ВИКОРИСТАННЯ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО
АНАЛІЗУ ДАНИХ ДЛЯ ВИРІШЕННЯ ЗАДАЧ ЕКОЛОГІЧНОЇ БЕЗПЕКИ АТМОСФЕРНОГО ПОВІТРЯ.....16

Пилип ПРИСТАВКА, Ольга ЧОЛИШКІНА

ЧАСТКОВІ ВИПАДКИ ЛОКАЛЬНИХ ПОЛІНОМІАЛЬНИХ СПЛАЙНІВ,
БЛИЗЬКИХ ДО ІНТЕРПОЛЯЦІЙНИХ У СЕРЕДНЬОМУ ДРУГОГО ТА ТРЕТЬОГО ПОРЯДКІВ26

Микола РУДНІЧЕНКО, Сергій ГРИШИН, Денис ШИБАЄВ, Ігор ПЕТРОВ, Максим НОСОВ

ПРОЕКТ ІНФОРМАЦІЙНОЇ СИСТЕМИ ДЛЯ ЗБОРУ, ОБРОБКИ ТА АНАЛІЗУ МЕТЕОДАНИХ.....35

**Микола РУДНІЧЕНКО, Сергій ГРИШИН, Наталя ШИБАЄВА, Володимир ВИЧУЖАНІН,
В'ячеслав РИЧКО**

АНАЛІЗ ОСОБЛИВОСТЕЙ ТА ЗАСОБІВ РОЗРОБКИ ІНТЕРАКТИВНИХ СИСТЕМ
З ПІДТРИМКОЮ ТЕХНОЛОГІЙ ДОПОВНЕНОЇ РЕАЛЬНОСТІ ДЛЯ СФЕРИ ОСВІТИ43

Микола РУДНІЧЕНКО, Тетяна ОТРАДСЬКА, Денис ШИБАЄВ, Ігор ПЕТРОВ, Микита ПОЛІКАРПОВ

РОЗРОБКА СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ
ДЛЯ МЕНЕДЖЕРА З УПРАВЛІННЯ ІТ-ПРОЕКТАМИ51

Ruslan SKURATOVSKIY, Anna BAZARNA

ANALYSIS OF THE MEL SCALE FEATURES USING ELECTROGRAPHY
AND SPEECH SIGNALS BY PARAMETERIZED KNN AND XGBOOST58

Ruslan SKURATOVSKIY

ANALYTICAL OVERVIEW AND SYSTEM ANALYSIS OF ORGANOPHOSPHATES FILTRATION
IN THE AGRICULTURAL SECTOR OF THE EUROPEAN UNION75

Денис ШИБАЄВ, Наталя ШИБАЄВА, Тетяна ОТРАДСЬКА, Микита КУЦ

РОЗРОБКА ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ
З ПРИЙНЯТТЯ РІШЕНЬ В КРЕДИТНОМУ СКОРИНГУ85

CONTENTS

Andrey DUDNIK

MODELS AND METHODS OF GROWTH OF PRODUCTIVITY OF SAFETY NETWORKS IN COMPONENTS
OF COMPUTERIZED SYSTEMS OF MEASUREMENT OF MECHANICAL QUANTITIES6

Valeriia KOVACH, Volodymyr KUTSENKO, Iryna MARTYNIUK, Oleksandr KOVALENKO

PECULIARITIES OF USING METHODS OF INTELLECTUAL DATA ANALYSIS TO SOLVE THE PROBLEMS
OF ENVIRONMENTAL SAFETY OF ATMOSPHERIC AIR16

Pylyp PRYSTAVKA, Olha CHOLYSHKINA

PARTIAL CASES OF LOCAL POLYNOMIAL SPLINES CLOSE TO INTERPOLATIONS
IN THE MIDDLE OF THE SECOND AND THIRD ORDERS26

Mykola RUDNICHENKO, Sergey GRISHIN, Denis SHIBAYEV, Igor PETROV, Maxim NOSOV

PROJECT OF INFORMATION SYSTEM FOR COLLECTION, PROCESSING AND DATA ANALYSIS35

**Mykola RUDNICHENKO, Sergey GRISHIN, Natalia SHIBAYEVA, Vladimir VYCHUZHANIN,
Vyacheslav RYCHKO**

ANALYSIS OF FEATURES AND DEVELOPMENT POTENTIAL OF INTERACTIVE SYSTEMS
SUPPORTED BY AUXILIARY REALITY TECHNOLOGIES FOR EDUCATION FIELD43

Mykola RUDNICHENKO, Tatiana OTRADSKA, Denis SHIBAYEV, Igor PETROV, Nikita POLIKARPOV

DECISION SUPPORT SYSTEM FOR THE IT PROJECT MANAGEMENT MANAGER51

Ruslan SKURATOVSKYI, Anna BAZARNA

ANALYSIS OF THE MEL SCALE FEATURES USING ELECTROGRAPHY
AND SPEECH SIGNALS BY PARAMETERIZED KNN AND XGBOOST58

Ruslan SKURATOVSKYI

ANALYTICAL OVERVIEW AND SYSTEM ANALYSIS OF ORGANOPHOSPHATES FILTRATION
IN THE AGRICULTURAL SECTOR OF THE EUROPEAN UNION75

Denis SHIBAYEV, Natalia SHIBAYEVA, Tatiana OTRADSKA, Mykyta KUTS

DEVELOPMENT OF INFORMATION AND ANALYTICAL SYSTEM
FOR DECISION MAKING IN CREDIT SCORING85

УДК 004.72:531.7.08

DOI <https://doi.org/10.32689/maup.it.2021.2.1>

Андрій ДУДНИК

кандидат технічних наук, доцент, професор кафедри комп'ютерних інформаційних систем і технологій, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом», вул. Фрометівська 2, Київ, Україна, індекс 03039 (a.s.dudnik@gmail.com)

ORCID: <https://orcid.org/0000-0001-5725-5942>

Andrey DUDNIK

Candidate of Technical Sciences, Associate Professor, Professor at the Department of Computer Information Systems and Technologies, Interregional Academy of personnel management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039 (a.s.dudnik@gmail.com)

Бібліографічний опис статті: Дудник А. Моделі та методи підвищення продуктивності безпроводових мереж у складі комп'ютеризованих систем вимірювання механічних величин. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 6–14. DOI: <https://doi.org/10.32689/maup.it.2021.2.1>

Bibliographic description of the article: Dudnik, A. (2021). Modeli ta metody pidvyshchennia produktyvnosti bezprovodovykh merezh u skladi kompiuteryzovanykh system vymiryuvannia mekhanichnykh velychyn [Models and methods of growth of productivity of safety networks in components of computerized systems of measurement of mechanical quantities]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 6–14. DOI: <https://doi.org/10.32689/maup.it.2021.2.1>

МОДЕЛІ ТА МЕТОДИ ПІДВИЩЕННЯ ПРОДУКТИВНОСТІ БЕЗПРОВОДОВИХ МЕРЕЖ У СКЛАДІ КОМП'ЮТЕРИЗОВАНИХ СИСТЕМ ВИМІРЮВАННЯ МЕХАНІЧНИХ ВЕЛИЧИН

Анотація. Стаття присвячена моделям та методам підвищення продуктивності безпроводових сенсорних мереж, що перебувають у складі комп'ютеризованих систем вимірювання механічних величин, побудованих на основі декомпозиції нижніх рівнів еталонної моделі OSI. Запропоновано метод підвищення продуктивності мереж, що функціонально об'єднує фізичний та мережевий рівні, що покращує її ефективність в зонах невпевненого прийому майже в два рази. Розроблено модель структурної схеми пристрою підвищення якості передавання даних в зонах невпевненого прийому або з недостатньою завадостійкістю, що базується на основі так званого методу моніторингу стану якості зв'язку.

Ключові слова: безпроводова мережа, механічні величини, сенсор, продуктивність, сигнал, комп'ютеризована система вимірювання, інформаційно-вимірювальна система.

MODELS AND METHODS OF GROWTH OF PRODUCTIVITY OF SAFETY NETWORKS IN COMPONENTS OF COMPUTERIZED SYSTEMS OF MEASUREMENT OF MECHANICAL QUANTITIES

Abstract. The article is devoted to models and methods of improving the productivity of wireless sensor networks, which are part of computerized systems for measuring mechanical quantities, based on the decomposition of the lower levels of the reference model OSI. The method of increasing the productivity of networks is proposed, which functionally combines physical and network levels, which improves its efficiency in zones of uncertain reception almost twice. The model of the structural scheme of the device for increasing the quality of data transmission in zones of uncertain reception or with insufficient noise immunity, based on the so-called method of monitoring the quality of communication, is developed.

Key words: wireless network, mechanical values, sensor, performance, signal, computerized measurement system, information and measuring system.

Вступ. Серед різноманітних класів комп'ютерних інформаційних систем і мереж особливе місце займають системи і мережі, транспортна служба яких побудована на використанні радіоефіру в якості середовища передачі даних комп'ютеризованих систем вимірювання механічних величин (безпроводові сенсорні мережі). Тому при створення наукових основ побудови комп'ютеризованих систем вимірювання механічних величин, важливе місце займає продуктивність безпроводових сенсорних мереж на основі модифікації існуючої класичної еталонної моделі взаємодії відкритих систем (ЕМ OSI/ISO), згідно з якою проектується, створюються і експлуатуються більшість засобів передачі даних інформаційно-вимірювальних систем, а також теоретичний аналіз та пошук оптимальних методів моделювання, взаємодії та управління комп'ютеризованих систем вимірювання механічних величин. Питанням дослідження інформаційно-вимірювальних систем, в тому числі і дослідженням технологій моделювання, управління і взаємодії комп'ютеризованих систем вимірювання механічних величин, присвячено

роботи сучасних вчених Кваснікова В.П., Орнадського Д.П., Осмоловського А.І., а також роботи Геєра Д., Ірвіна Дж., Лієрі Дж., Рошана П., Столлінгса В., Харля Д. та ін.

Постановка задачі. Метою даного дослідження є аналіз та обґрунтування теоретичних основ, а також розробка моделей та методів, пов'язаних з підвищенням продуктивності безпроводових мереж, що входять до складу комп'ютеризованих систем вимірювання механічних величин.

Виклад основного матеріалу. Проведено порівняльний аналіз основних існуючих технологій передавання показників вимірювання. На основі експериментальних досліджень виявлено недоліки та ступені свободи для покращення систем управління, взаємодії та підвищення продуктивності в існуючих технологіях безпроводового передавання даних інформаційно-вимірювальних систем. Розглянуто моделі безпроводових комп'ютеризованих систем вимірювання механічних величин та особливості моделювання передачі показників вимірювання у безпроводовому середовищі [1–5].

Багатопроменевий характер поширення радіохвиль, коли кожна точка простору характеризується тим, що до неї надходять сигнали з різних напрямків та з різними часовими затримками, спливає багатократній інтерференції сигналів та їх суттєвому спотворенню. Такі спотворення мають суттєвий вплив на характеристики системи, якщо тривалість затримки перевищує тривалість передавання пакету даних вимірювання. Тобто, у комп'ютерних системах безпроводового зв'язку, як і для проводових систем, велике значення має головне правило, згідно з яким пакет повинний досягти кінцевої точки раніше, ніж буде переданий останній його символ. Але у разі різнонаправленого поширення безпроводових сигналів виконати цю умову значно важче [1–5].

Концептуальна модель простого варіанту мережі скатернет (розподіленої мережі з багатьма розгалуженнями на основі технології *Bluetooth*), що входить до складу комп'ютеризованої системи вимірювальної системи, представлена відкритою (незамкнутою) багатофазною системою масового обслуговування. З погляду класифікації еталонної моделі описувалися два нижні рівні. Транзакт, що є неподільним об'єктом в системі імітаційного моделювання загального призначення, породжувався бітом, що переміщується в скатернет від джерела інформації до споживача. Кожна фаза моделювалася СМО $G/M/n$ з відмовами і дисципліною *FIFO*. Універсальна система імітаційного моделювання забезпечує збір і статистичну обробку даних про транзакти, затримані в кожній точці моделі, а також інтенсивності потоків відмов. Час затримки бітів в каналах передавання (*Wчерги*) даної мережі визначається у відповідності до формули для обчислення часу затримки в черзі багатоканального пристрою з очікуванням:

$$W_{\text{черги}} = \frac{1}{\sum_{i=1}^N \lambda_i} L_{\text{черги}}, \quad (1)$$

де λ – інтенсивність надходження бітів, що передаються для i -го стану;

$L_{\text{черги}}$ – середня кількість бітів, що передаються і визначається за наступною формулою:

$$L_{\text{черги}} = \frac{\sum_{i=1}^N \rho_i P_0}{n \cdot n! (1 - \sum_{i=1}^N \rho_i / n)^2},$$

де n – кількість розподілених підканалів безпроводового передавання даних мережі *Bluetooth*, що для даного випадку складає 23 підканали;

P_0 – імовірність того, що підканал на даний момент зайнятий, визначається за формулою:

$$P_0 = \left(1 + \frac{\sum_{i=1}^N \rho_i}{1!} + \frac{\sum_{i=1}^N \rho_i}{2!} + \dots + \frac{\sum_{i=1}^N \rho_i^n}{n!} + \frac{\sum_{i=1}^N \rho_i^{n+1}}{n!(n - \sum_{i=1}^N \rho_i)} \right)^{-1},$$

де ρ – навантаження на дану мережу безпроводового передавання даних, що визначається за формулою:

$$\rho = \frac{\sum_{i=1}^N \lambda_i}{\mu}, \quad (2)$$

де λ – інтенсивність надходження бітів до мережі передавання даних при i -тому стані;

μ – інтенсивність обслуговування бітів у мережі передавання даних. Підставивши усі ці значення у формулу (1), отримаємо наступну результуючу формулу:

$$W_{\text{черги}} = \frac{1}{\lambda} \frac{\left(\frac{\sum_{i=1}^N \lambda_i}{\mu} \right)^{n+1} \left(1 + \frac{\sum_{i=1}^N \lambda_i}{\mu} + \frac{\sum_{i=1}^N \lambda_i}{2!} + \dots + \left(\frac{\sum_{i=1}^N \lambda_i}{\mu} \right)^n + \left(\frac{\sum_{i=1}^N \lambda_i}{\mu} \right)^{n+1} \right)^{-1}}{n! \left(n - \frac{\sum_{i=1}^N \lambda_i}{\mu} \right)}$$

$$n \cdot n! \left(1 - \frac{\sum_{i=1}^N \lambda_i}{\mu} \right)^2 \frac{\mu}{n}$$

Головний пристрій “Master” передає біти в непарні моменти часу, а підлеглий пристрій “Slave” – у парні. Дана модель має ієрархічну будову адже пристрій, який для окремої ділянки мережі є головним, для іншої ділянки може бути підлеглим. Експериментальне дослідження даної моделі було проведене за допомогою інструментального засобу GPSS, при 10^6 прогонів. З отриманих результатів виявлено, що співвідношення часу затримки відповідає реально наявним співвідношенням параметрів мережі [6].

Розглянута модель інфрачервоного трафіку. Мета даного дослідження – побудова одноканальної моделі безпроводового передавання параметрів вимірювання інфрачервоної сенсорної мережі, результати роботи якої б відображали часові та кількісні параметри передавання даних вимірювання механічних величин. Імітаційну модель про інфрачервоного зв’язку подано відкритою одноканальною системою масового обслуговування. Транзакт, що є неподільним об’єктом в системі імітаційного моделювання загального призначення, породжується бітом, що переміщується в інфрачервоній мережі від джерела інформації до споживача. Модель, за класифікацією є моделлю, $G/M/1$ з дисципліною *FIFO*. Час затримки бітів в каналах передавання ($W_{\text{черги}}$) даної системи буде визначатись у відповідності до формули до формули для обчислення часу затримки в черзі одноканального пристрою з очікуванням [7–10]:

$$W_{\text{черги}} = \frac{L_{\text{черги}}}{\sum_{i=1}^N \lambda_i (1 - P_N)}, \quad (3)$$

де $L_{\text{черги}}$ – середня кількість бітів, що передаються, визначається за наступною формулою:

$$L_{\text{черги}} = \sum_{i=1}^N n P_n,$$

де P_n – імовірність того, що в каналі інфрачервоного передавання даних знаходиться n бітів і визначається за формулою:

$$P_n = P_0 \sum_{i=1}^N \rho_i^n,$$

P_0 – імовірність того, що підканал на даний момент зайнятий, визначається за формулою:

$$P_0 = \frac{1 - \sum_{i=1}^N \rho_i^n}{1 - \sum_{i=1}^N \rho_i^{n+1}},$$

де ρ – навантаження на мережу бездротового передавання даних, що визначається за формулою (2). Підставивши дані значення у формулу (3), отримаємо наступну результуючу формулу:

$$W_{\text{черги}} = \frac{\sum_{n=0}^N n \frac{1 - \frac{\sum_{i=1}^N \lambda_i}{\mu}}{\sum_{i=1}^N \lambda_i} \left(\frac{\sum_{i=1}^N \lambda_i}{\mu} \right)^n}{\sum_{i=1}^N \lambda_i (1 - P_N)}.$$

У результаті моделювання були отримані параметри роботи мережі, зокрема затримка в чергах каналів передавання обох пристроїв, що відповідає пропускній спроможності каналу передавання у випадку, коли навантаження на даний канал є максимальним [6].

Досліджено стандарти мережі типу *IEEE 802.11*, застосовуючи ту чи іншу дисципліну черг, що задається за допомогою застосування одного з алгоритмів управління чергами. Основною задачею даних досліджень являється моделювання багатоканальної безпроводової сенсорної мережі передавання даних вимірювання, із застосуванням різних алгоритмів управління чергами. Визначення на практиці оптимального алгоритму управління передавання даних, за допомогою проведення порівняльних характеристик. Алгоритми, що застосовуються для моделювання беруться довільно [7–10].

Алгоритми управління чергами потрібні для роботи в періоди тимчасових перевантажень при максимальному завантаженні каналів передавання безпроводової мережі, що знаходиться у складі комп'ютеризованої системи вимірювання механічних величин, якщо мережевий пристрій не може впоратися з передаванням бітів на вихідний інтерфейс в тому темпі, в якому вони поступають. Якщо причиною перевантаження є недостатня продуктивність процесорного блоку мережевого пристрою безпроводової, то необроблені біти тимчасово накопичуються у вхідній черзі відповідного вхідного інтерфейсу. Черг до вхідного інтерфейсу може бути декілька, якщо диференціюються запити на обслуговування по декількох класах. У тому ж випадку, коли причина перевантаження полягає в обмеженій пропускній спроможності вихідного інтерфейсу, біти тимчасово зберігаються у вихідній черзі (або чергах) цього інтерфейсу. Концептуальні моделі мереж з застосуванням різних алгоритмів управління чергами в мережах WI-FI представлені відкритою багатозадачною системою масового обслуговування. Кожна фаза моделювалася СМО $G/M/n$ побудована згідно умов того чи іншого алгоритму управління чергами. Результуючою формулою для визначення часу затримки бітів в каналах передавання (*Wчерги*) даної мережі буде формула 1 тому, що модель даної мережі також представлена у вигляді багатоканальної СМО з очікуванням. Модель складається з вимірювальних сенсорів та точки безпроводового доступу. Опишемо роботу даної системи: біти, що надходять з сенсору, який передає дані вимірювання, стають у чергу на обслуговування до точки доступу. Далі у точці доступу, вони стають в чергу на розподілення до робочих станцій, які є їхніми адресатами. Потім вони стають у чергу на обробку до потрібного адресату [7–10].

Розглянуто 3 моделі відповідної структури у яких застосовувались 3 різні алгоритми управління трафіком, що застосовуються у різних стандартах *IEEE 802.11*, а саме «FIFO», «Пріоритетне обслуговування», «Зважені черги». Проведено дослідження з метою порівняльної характеристики між цими алгоритмами за критеріями пропускної спроможності каналів передавання, максимальної швидкості передавання показників, дальності передавання показників, кількості підканалів. Після ряду досліджень з метою порівняння, проведених за правилами процедури підтримки прийняття рішень «*Analitic hierarchy process*» (АНП) по даним алгоритмам було отримано наступні підсумкові коефіцієнти, а саме алгоритм «FIFO» – 0,228, алгоритм «пріоритетне обслуговування» – 0,222, алгоритм «Зважені черги» – 0,55. У результаті даних порівнянь досліджено, що алгоритм «зважені черги» в двічі кращий ніж 2 інших. Проведено дослідження, які показали, що результати, отримані при використанні моделі типу $G/M/n$ для моделювання передачі показників вимірювання сенсорними мережами, є найбільш точними серед інших моделей.

Досліджено модель сенсорної мережі на концептуальному рівні. На основі результатів моделювання даної мережі при стандартній смузі частот мережі «*Bluetooth*» (2400 – 2483,5 МГц *ISM*-діапазону), за формулою (1) була одержана характеристика на рис. 1 [7–10].

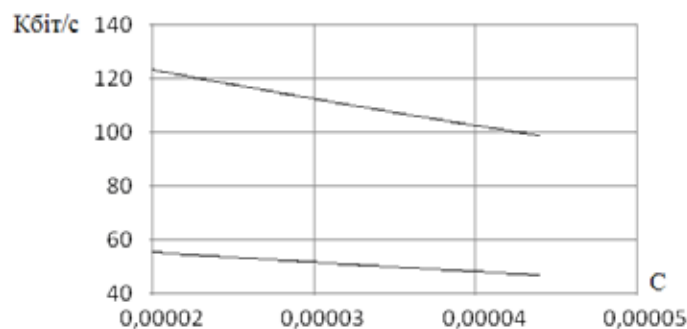


Рис. 1. Графік залежності пропускної спроможності каналу передавання від часу затримки на обробку

Стандартну смугу частот було розширено на 80 МГц, та за формулою (1), знято характеристику, що на рис. 2.

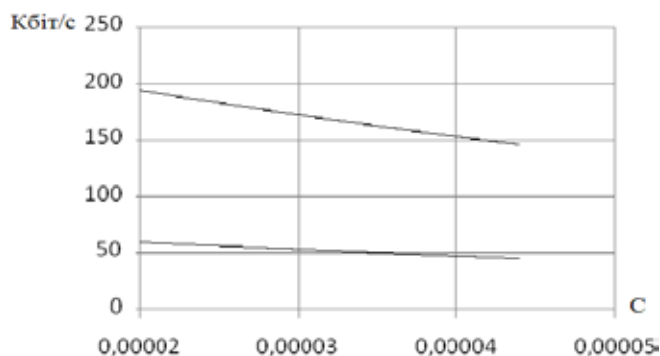


Рис. 2. Графік залежності пропускної спроможності каналу передавання від часу затримки на обробку при розширенні смуги частот

У підсумку зроблений висновок, що розширення смуги частот хоч і погіршує на початку швидкісні показники передавання показників вимірювання сенсорною мережею *Bluetooth*, але значно поліпшує інші параметри, зокрема пропускну спроможність.

Досліджено інфрачервону сенсорну мережу. Передавання показників вимірювання, в даному випадку, буде відбуватися від одного вимірювального пристрою до іншого. Досліджено залежність пропускної спроможності каналу передавання (C) від часу затримки на обробку в обслуговуючому пристрої каналу, яку зображено на рис. 3, було одержано за формулою (3) [7–10].

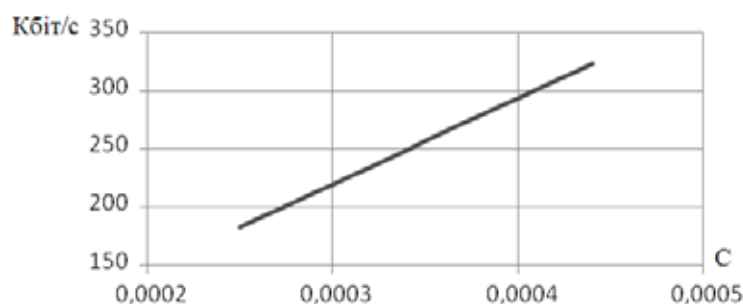


Рис. 3. Графік Залежності пропускної спроможності каналу передавання від часу затримки на обробку в обслуговуючому пристрої каналу

На даному графіку спостерігається лінійне зростання навантаження. Що свідчить про стабільну поведінку обслуговуючих пристроїв при моделюванні стаціонарних процесів за допомогою моделей типу $G/M/1$.

Запропоновано алгоритм функціонального об'єднання нижніх рівнів еталонної моделі *OSI*.

Дану технологію було застосовано для досліджень мереж стандарту *IEEE 802.11* із застосуванням трьох вищезазначених алгоритмів. Це дало змогу одержати наступні результати, за формулою 5, що зображені на рис. 6, 7.

Алгоритм базується на контролюванні бітів фрейму фізичного рівня. Він побудований з тою метою, щоб з полів фрейму, які містять інформацію про середовище передавання даних, отримувати потрібну інформацію. У даному методі мережевий рівень через певні проміжки часу відсилає запити на фізичний про стан сигналу на даний момент. Інформація про стан сигналу міститься у полі, в якому перші біти тримають у собі інформацію про швидкість передавання, а інші про стан сигналу [7–10].

Зазвичай кожному з вище перелічених станів сигналу відповідає число від 0 до 6. На основі одержаних даних мережевий рівень формує уявлення про зміну станів зв'язку, постійно порівнюючи теперішній та попередній стани ($n < n_0$ чи $n \geq n_0$). У випадку, коли зміна станів відповідає $n < n_0$, мережевий рівень відсилає повторний запит. Даний цикл буде відбуватися до тих пір, доки ситуація не стане протилежною ($n \geq n_0$) (гілка «так»). Лише тоді мережевий рівень направить до каналного пакет і віддасть

наказ на його передавання. Потім каналний рівень, при наявності вільного каналу, надасть наказ на генерацію бітів фізичному рівню, у відповідності до конкретного пакету [7–10].



Рис. 4. Блок-схема алгоритму аналізу стану якості сигналу

Особливості даного алгоритму було враховано в структурній схемі пристрою безпроводового передавання даних лінії, що виконувала б функцію зв'язку між додатками фізичного та мережевого рівнів еталонної моделі *OSI*. Тобто працювала б на основі даного алгоритму. Саме з цієї причини в існуючу структурну схему пристрою безпроводового зв'язку вводиться аналізатор якості сигналу 5 рис. 5.

При побудові пристрою, блоки розбито на модулі, згідно їхньої належності до того чи іншого рівня еталонної моделі. Пристрій містить керуючий блок 1, що є частиною *NMS*, модуль мережевого рівня 2, модуль підрівня *LLC* каналного рівня моделі 3, блоку хост інтерфейсу 3.1, вбудованого мікроконтролера 3.2, блоку додатку прийомо/передавача 3.3, блоку шинного інтерфейсу 3.4, пам'яті *S*, модуль підрівня *MAC* каналного рівня 4, контролера смуги частот 4.1, радіочастотного прийомо/передавача 4.2, аналізатора сигналу 5, модуля фізичного рівня 6, інтерфейсу фізичного рівня 6.1, антени 6.2., блоку автоматичного налаштування частоти 7.

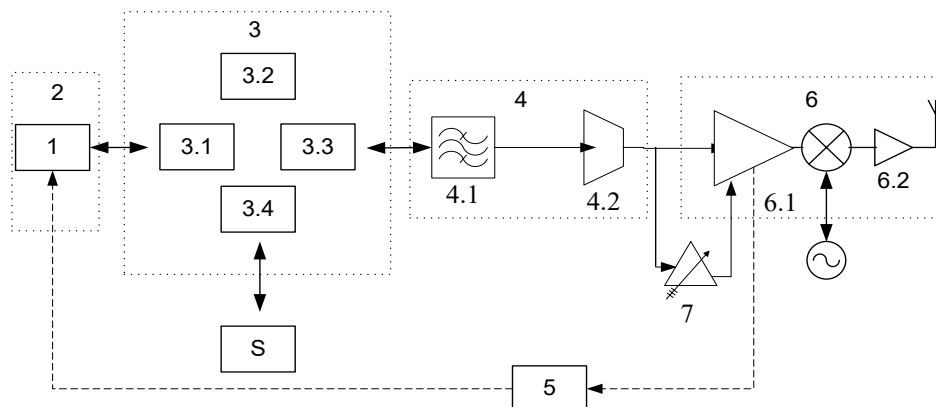


Рис. 5. Безпроводовий мережевий пристрій з системою підвищення якості передачі показників вимірювання в зонах невпевненого прийому або з недостатньою завадостійкістю

Точка насичення X_0 визначається за формулою:

$$X_0 = \frac{1}{V_d R_d},$$

де X_0 – найбільше навантаження для даної мережі;

V_d – коефіцієнт відвідування для вузла d ;

R_d – час перебування бітів у мережі для вузла d .

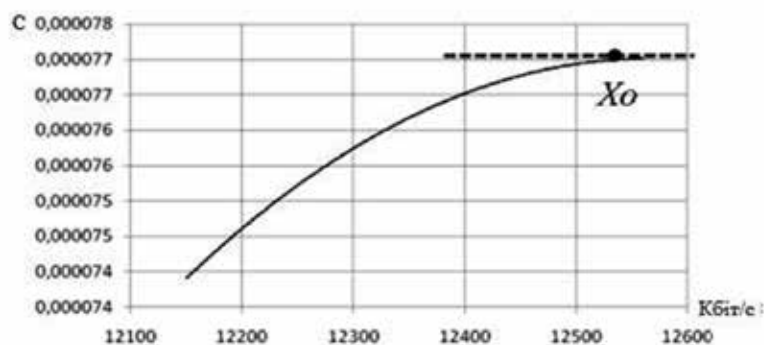


Рис. 6. Графік залежності часу затримки від навантаження каналу передавання

Швидкість передавання показників вимірювання, яку було використано при моделюванні відповідала *Peak Cell Rate (PCR)* – максимальній швидкості передавання даних складає 54 Мбіт/с.

В безпроводових каналах передавання даної мережі перерозподіл потоків буде відбуватися за допомогою алгоритму курування чергами «Зважені черги». Тобто розподіл бітів інформації в каналах передавання буде відбуватись згідно відсотку пропускної спроможності, яка надана даному класу трафіку. Кількість відсотків пропускної спроможності надається тому чи іншому класу трафіку згідно його пріоритету.

Проведено експеримент, згідно згаданого вище алгоритму.

На основі даних досліджень, за формулою (1), було одержано характеристику наведену на рис. 6.

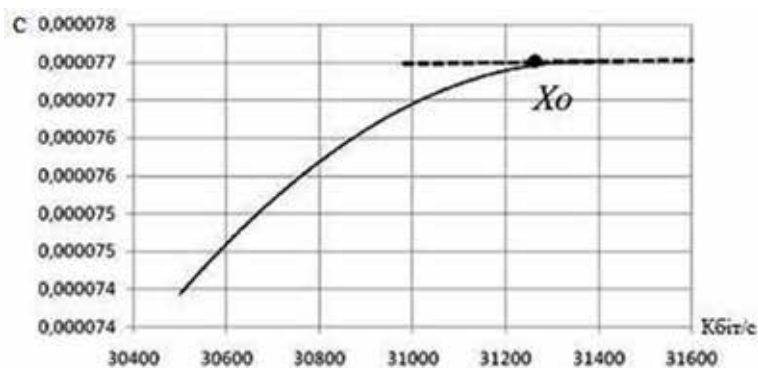


Рис. 7. Графік залежності часу затримки від навантаження каналу передавання

Як показано на графіку, даний алгоритм вносить в роботу безпроводової сенсорної мережі досить позитивні зміни. Це видно з того, що при значному зростанні навантаження в каналах передавання час затримки бітів в черзі на обробку майже не змінився. Тобто дані алгоритми більше майже в три рази покращують швидкісні показники роботи мережі. Цілком імовірно, що таким покращенням сприяє зменшення відсотку помилкових пакетів та запитів на їх повторну відправку, що також потребують певного часу для обробки, а також перехід на стандарт з більшою потужністю при падінні рівня сигналу, що безумовно покращує стійкість роботи сенсорної мережі, значно зменшуючи імовірність відсутності сигналу в зонах невпевненого прийому.

Висновки. Головною особливістю способів управління передаванням даних бездротових інформаційно-вимірювальних систем являється обмеженість пропускної спроможності системи та необмежений вплив перешкод на середовище передавання даних від сторонніх джерел, що випромінюють надвисокі частоти. Можливе зниження пропускної спроможності мережі може сягнути 90% у порівнянні з аналогічним випадком у мережах на основі кабелю.

Моделювання алгоритмів управління чергами безпроводових IBC дозволило виявити те що при використанні алгоритмів «Зважені черги» сумарна пропускна спроможність на 9 Мбіт/с краща ніж при використанні інших алгоритмів.

Знайдено наступні рішення проблеми покращення кількісних та якісних показників безпроводового передавання показників відмірювання механічних величин: розширення частотного діапазону

мережі короткого радіусу дії дає змогу покращити пропускну спроможність мережі майже на 50 %; запропоновано метод підвищення продуктивності для безпроводових сенсорних мереж, що функціонально об'єднує фізичний та мережевий рівні, що може покращити пропускну спроможність мережі в зонах невпевненого прийому майже в два рази; алгоритм зваженого обслуговування заявок трафіку безпроводових комп'ютеризованих систем вимірювання механічних величин, є найбільш ефективною технологією передавання даних у безпроводових сенсорних мережах, за умови рівня неточності вихідних даних 0,9 та на його основі запропоновано метод «Перерозподіл пропускну спроможності каналу передавання сенсорної мережі», що покращує дисципліну обслуговування в черзі, який буде описано в наступних роботах.

Список використаних джерел:

1. Квасніков В.П., Хаєйн Т.М. Концепція повірки координатно-вимірювальних машин через Інтернет. *Методологія та прилади*. 2013. № 6. С. 48–53. URL: <http://ua.amu.in.ua/journal1>.
2. Квасніков В.П., Ганева Т.І. Нова методика розрахунку датчика деформації. *Вимірювальна та обчислювальна техніка в технологічних процесах*. 2015. № 3(52). С. 169–173. URL: <http://inmad.vntu.edu.ua/portal/static/41E185F3-5121-491D-84A7-F1E92DE5FB6E.pdf>.
3. Kvasnikov V.P., Ganeva T.I. Method of measurement of stress in loaded structures. *Metallurgical and Mining Industry*. 2015. № 7. Р. 340–343.
4. Квасніков В.П., Лещенко Ю.П. Розробка та дослідження інтелектуальної системи для аналізу ймовірнісних характеристик випадкових процесів механічних величин. *Технологический аудит и резервы производства*. 2015. № 5/3(25). С. 100–102. URL: <http://journals.uran.ua/tarp/article/viewFile/52011/48196>.
5. Квасніков В.П., Орнатский Д.П., Осмоловский А.И. Оптимизация аналогового интерфейса информационно-измерительных систем для дистанционных измерений механических величин. *Восточно-Европейский журнал передовых технологий*. 2014. № 4(70). С. 31–37. URL: <http://cyberleninka.ru/article/n/optimizatsiya-analogovogo-interfeysa-informatsionno-izmeritelnyh-sistem-dlya-distantsionnyh-izmereniy-mekhanicheskikh-velichin>.
6. Дуднік А.С. Метод застосування сучасних безпроводових технологій в інформаційних економічних системах. *Проблеми та перспективи розвитку економічної кібернетики* : монографія / під ред. Т.І. Олешко. Київ : ВД ТОВ «Agrar Media Group», 2013. С. 214–225.
7. Дуднік А.С., Бондаренко Ю.В. Система покращеного передавання даних безпроводових комп'ютерних мереж стандарту IEEE 802.11. *Наука і молодь* : зб. наук. праць. Київ : НАУ, 2012. № 11(12). С. 113–116.
8. Дуднік А.С., Шевцова Є.В., Зубарева О.О. Спосіб підвищення продуктивності безпроводових комп'ютерних мереж на основі міжрівневої взаємодії та пристрій для його реалізації. *Проблеми інформатизації та управління* : зб. наук. Праць, 2011. № 4(36). С. 45–50.
9. Дуднік А.С., Шевцова Є.В., Яценко М.М., Зубарева О.О. Пристрій підвищення якості передавання даних в бездротових мережах в зонах невпевненого прийому або з недостатньою завадостійкістю. *Пат. № 60400 України, МПК H04B 7/005; № u201007469; Заявл. 15.06.10; Опубл. 25.06.11*. Бюл. № 12. 4 с.
10. Дуднік А.С. Динамічний алгоритм управління перерозподілом мережевого трафіку в безпроводових інформаційних системах. *Проблеми модернізації України* : тези допов. всеукраїнської наук.-практ. конф. (м. Київ, 12 листопада 2015 р.) Київ, 2015. С. 126.

References:

1. Kvasnikov, V.P., Khaeyn, T.M. (2013). Kontseptsiya povirky koordynatno-vymiryuval'nykh mashyn cherez Internet [The concept of calibration of coordinate measuring machines via the Internet]. *Metrolohiya ta prylady*, no. 6, pp. 48–53. URL: <http://ua.amu.in.ua/journal1> [in Ukrainian].
2. Kvasnikov, V.P., Hanyeva, T.I. (2015). Nova metodyka rozrakhunku datchyka deformatsiy [A new method of calculating the deformation sensor]. *Vymiryuval'na ta obchyslyuval'na tekhnika v tekhnologichnykh protsesakh*, no. 3(52), pp. 169–173. URL: <http://inmad.vntu.edu.ua/portal/static/41E185F3-5121-491D-84A7-F1E92DE5FB6E.pdf> [in Ukrainian].
3. Kvasnikov, V.P., Ganeva, T.I. (2015). Method of measurement of stress in loaded structures. *Metallurgical and Mining Industry*, no. 7, pp. 340–343 [in English].
4. Kvasnikov, V.P., Leshchenko, Yu.P. (2015). Rozrobka ta doslidzhennya intelektual'noyi systemy dlya analizu ymovirnisnykh kharakterystyk vypadkovykh protsesiv mekhanichnykh velychyn [Development and research of an intelligent system for the analysis of probabilistic characteristics of random processes of mechanical quantities]. *Tehnologicheskij audit i rezervy proizvodstva*, no. 5/3(25), pp. 100–102. URL: <http://journals.uran.ua/tarp/article/viewFile/52011/48196> [in Ukrainian].
5. Ornatskij, D.P., Kvasnikov, V.P., Osmolovskij, A.I. (2014). Optimizatsiya analogovogo interfeysa informacionno-izmeritelnyh sistem dlya distantsionnyh izmerenij mekhanicheskikh velychin [Optimization of the analog interface of information-measuring systems for remote measurements of mechanical quantities]. *Vostochno-Evropejskij zhurnal peredovykh tekhnologij*, no. 4(70), pp. 31–37. URL: <http://cyberleninka.ru/article/n/optimizatsiya-analogovogo-interfeysa-informatsionno-izmeritelnyh-sistem-dlya-distantsionnyh-izmereniy-mekhanicheskikh-velichin> [in Russian].
6. Dudnik, A.S. (2013). Metod zastosuvannya suchasnykh bezprovodovykh tekhnologij v informatsiynykh ekonomichnykh systemakh [Method of application of modern wireless technologies in information economic systems]. *Problemy ta perspektyvy rozvytku ekonomichnoi kibernetiky* : monohrafiia / pid red. T.I. Oleshko. Kyiv: VD TOV «Agrar Media Group», pp. 214–225 [in Ukrainian].

7. Dudnik, A.S., Bondarenko, Yu.V. (2012). Systema pokrashchenoho peredavannia danykh bezprvodovykh kompiuternykh merezh standartu IEEE 802.11 [IEEE 802.11 Wireless Computer Networking System]. *Nauka i molod: zb. nauk. prats.* Kyiv: NAU, no. 11(12), pp. 113–116 [in Ukrainian].
8. Dudnik, A.S., Shevtsova, Ye.V., Zubarieva, O.O. (2011). Sposib pidvyshchennia produktyvnosti bezprvodovykh kompiuternykh merezh na osnovi mizhrivnevoi vzaiemodii ta prystrii dlia yoho realizatsii [Method of increasing productivity of wireless computer networks on the basis of interlevel interaction and device for its implementation]. *Problemy informatyzatsii ta upravlinnia: zb. nauk. Prats*, no. 4(36), pp. 45–50 [in Ukrainian].
9. Dudnik, A.S., Shevtsova, Ye.V., Yatsenko, M.M., Zubareva, O.O. (2011). Prystrii pidvyshchennia yakosti peredavannia danykh v bezdrotovykh merezhakh v zonakh nevpevnenoho pryimu abo z nedostatnoiu zavadostiikistiu [Device for improving data transmission quality in wireless networks in areas of uncertain reception or with insufficient noise immunity]. *Pat. № 60400 Ukrainy, MPK N04V 7/005; № u201007469; Zaiavl. 15.06.10; Opubl. 25.06.11.* Biul, no. 12, 4 p. [in Ukrainian].
10. Dudnik, A.S. (2015). Dynamichniy alhorytm upravlinnia pererozpodilom merezhevoho trafiku v bezprovodovykh informatsiinykh systemakh [Dynamic algorithm for controlling the redistribution of network traffic in wireless information systems]. *Problemy modernizatsii Ukrainy: tezy dopov. vseukrainskoi nauk.-prakt. konf., m. Kyiv, 12 lystopada 2015 r.* Kyiv, p. 126. [in Ukrainian].

УДК 519.6:504.064

DOI <https://doi.org/10.32689/maup.it.2021.2.2>

Валерія КОВАЧ

доктор наук з державного управління, старший дослідник, провідний науковий співробітник відділу цивільного захисту та інноваційної діяльності, Державна установа «Інститут геохімії навколишнього середовища Національної академії наук України», просп. Академіка Палладіна 34а, Київ, Україна, індекс 03142 (valeriiakovach@gmail.com)

ORCID: <https://orcid.org/0000-0002-1014-8979>

Володимир КУЦЕНКО

молодший науковий співробітник відділу технологій захисту довкілля та радіаційної безпеки, Державна установа «Інститут геохімії навколишнього середовища Національної академії наук України», просп. Академіка Палладіна 34а, Київ, Україна, індекс 03142 (kuts.vo@gmail.com)

ORCID: <https://orcid.org/0000-0002-0577-2056>

Ірина МАРТИНЮК

молодший науковий співробітник відділу технологій захисту довкілля та радіаційної безпеки, Державна установа «Інститут геохімії навколишнього середовища Національної академії наук України», просп. Академіка Палладіна 34а, Київ, Україна, індекс 03142 (mira_27@ukr.net)

ORCID: <https://orcid.org/0000-0003-4654-7071>

Олександр КОВАЛЕНКО

молодший науковий співробітник відділу технологій захисту довкілля та радіаційної безпеки, Державна установа «Інститут геохімії навколишнього середовища Національної академії наук України», просп. Академіка Палладіна 34а, Київ, Україна, індекс 03142 (whitewavex@gmail.com)

ORCID: <https://orcid.org/0000-0003-4798-7722>

Valeriia KOVACH

Doctor of Science in Public Administration, Senior researcher, Leading Researcher of the Department of Civil Protection and Innovation, State Institution "The Institute of Environmental Geochemistry of National Academy of Sciences of Ukraine", 34a Palladin Ave., Kyiv, Ukraine, postal code 03142 (valeriiakovach@gmail.com)

Volodymyr KUTSENKO

Junior Researcher of the Department of environmental protection technologies and radiation safety, State Institution "The Institute of Environmental Geochemistry of National Academy of Sciences of Ukraine", 34a Palladin Ave., Kyiv, Ukraine, postal code 03142 (kuts.vo@gmail.com)

Iryna MARTYNIUK

Junior Researcher of the Department of environmental protection technologies and radiation safety, State Institution "The Institute of Environmental Geochemistry of National Academy of Sciences of Ukraine", 34a Palladin Ave., Kyiv, Ukraine, postal code 03142 (mira_27@ukr.net)

Oleksandr KOVALENKO

Junior Researcher of the Department of environmental protection technologies and radiation safety, State Institution "The Institute of Environmental Geochemistry of National Academy of Sciences of Ukraine", 34a Palladin Ave., Kyiv, Ukraine, postal code 03142 (whitewavex@gmail.com)

Бібліографічний опис статті: Ковач В., Куценко В., Мартинюк І., Коваленко О. Особливості використання методів інтелектуального аналізу даних для вирішення задач екологічної безпеки атмосферного повітря. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 15–25. DOI: <https://doi.org/10.32689/maup.it.2021.2.2>

Bibliographic description of the article: Kovach, V., Kutsenko, V., Martyniuk, I., Kovalenko, O. (2021). Osoblyvosti vykorystannia metodiv intelektualnoho analizu danykh dlia vyrishennia zadach ekolohichnoi bezpeky atmosferneho povitria [Peculiarities of using methods of intellectual data analysis to solve the problems

of environmental safety of atmospheric air]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 15–25. DOI: <https://doi.org/10.32689/maup.it.2021.2.2>

ОСОБЛИВОСТІ ВИКОРИСТАННЯ МЕТОДІВ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ ДАНИХ ДЛЯ ВИРІШЕННЯ ЗАДАЧ ЕКОЛОГІЧНОЇ БЕЗПЕКИ АТМОСФЕРНОГО ПОВІТРЯ

Анотація. Описано процес управління екологічною безпекою в контексті зменшення негативних впливів на навколишнє середовище. Запропоновано узагальнену структурну модель процесу управління екологічною безпекою, засновану на методах та технологіях інтелектуального аналізу даних моніторингу. Досліджено можливості адаптації та удосконалення ряду найбільш відомих алгоритмів інтелектуального аналізу даних, таких як C4.5, K-means, методу опорних векторів (SVM), kNN, наївного байєсового класифікатора, алгоритму Apriori, для задач аналізу даних мережі моніторингу атмосферного повітря. На прикладі даних щодо концентрацій забруднюючих речовин в атмосферному повітрі міста Кривий Ріг (Дніпропетровська область) побудовано діаграми: розсіювання концентрацій пилу та діоксиду азоту; розсіювання температури повітря і концентрації сірчаного ангідриду. Наведено приклади практичного використання окремих методів з метою виявлення небезпечних ситуацій. **Метою** статті є аналіз принципів та методів управління екологічною безпекою на основі інтелектуального аналізу даних мережі моніторингу атмосферного повітря. **Наукова новизна.** Запропоновано концептуальну модель управління екологічною безпекою урбанізованих територій за даними екологічного моніторингу, що відрізняється від своїх аналогів новими можливостями, базованими на використанні сучасних інформаційних технологій інтелектуального аналізу даних. У **висновках** наголошено, що авторами публікації визначено перспективи застосування методів та засобів інтелектуального аналізу даних щодо інформаційної підтримки прийняття рішень, спрямованих на оцінювання наслідків техногенного впливу і зменшення обсягу навантажень на довкілля. Запропоновано ряд методів та алгоритмів, які дають можливість оцінювати значення невідомих характеристик і параметрів за відомими даними, та показано приклади їх використання.

Ключові слова: екологічна безпека, управління, інтелектуальний аналіз даних, атмосферне повітря.

PECULIARITIES OF USING METHODS OF INTELLECTUAL DATA ANALYSIS TO SOLVE THE PROBLEMS OF ENVIRONMENTAL SAFETY OF ATMOSPHERIC AIR

Abstract. The process of environmental safety management in the context of reducing negative impacts on the environment is described. A generalized structural model of the environmental safety management process based on methods and technologies of intellectual analysis of monitoring data is proposed. Possibilities of adaptation and improvement of a number of the most known algorithms of data mining, such as C4.5, K-means, method of reference vectors (SVM), kNN, naive Bayesian classifier, Apriori algorithm, for data analysis of atmospheric air monitoring network are investigated. On the example of data on concentrations of pollutants in the air of Kryvyi Rih city (Dnipropetrovsk region) diagrams are constructed: scattering of concentrations of dust and nitrogen dioxide; dissipation of air temperature and concentration of sulfur dioxide. Examples of practical use of separate methods for the purpose of dangerous situations detection are resulted. **The aim** of the article is to analyze the principles and methods of environmental safety management based on the intellectual analysis of atmospheric air monitoring network data. **Scientific novelty.** A conceptual model of ecological safety management of urban areas based on ecological monitoring data is proposed, which differs from its analogues by new possibilities based on the use of modern information technologies of data mining. **The conclusions** emphasize that the authors of the publication identify prospects for the application of methods and tools of data mining for information support of decision-making aimed at assessing the effects of man-made impact and reducing the burden on the environment. A number of methods and algorithms are proposed, which make it possible to estimate the values of unknown characteristics and parameters according to known data, and examples of their use are shown.

Key words: ecological safety, management, data mining, atmospheric air.

Актуальність проблеми. Інтенсивний розвиток науково-технічного прогресу в останні десятиріччя супроводжується збільшенням техногенного впливу на атмосферу, гідросферу та літосферу, зростанням обсягів відходів виробничої діяльності й транспортних засобів. Серед глобальних екологічних проблем найбільш пріоритетними є забруднення повітря, дефіцит прісної води, зменшення видового й ландшафтного різноманіття планети, парниковий ефект, озонові діри, кислотні дощі, масові захворювання людей, загибель лісів тощо. Зменшення рівня антропогенного впливу на біосферу можна досягти якісним управлінням екологічною безпекою на всіх рівнях, забезпечивши стратегічну орієнтацію на принципи стійкого (гармонійного) розвитку.

На даний час існує багато визначень і трактувань щодо поняття «екологічної безпеки». Згідно ст. 50 Закону України «Про охорону навколишнього природного середовища», термін «екологічна безпека» визначає такий стан навколишнього природного середовища, у якому реалізовано реальні можливості для попередження погіршення екологічної обстановки та виникнення небезпеки для здоров'я людей. Як зазначено в Енциклопедії Сучасної України [esu.com.ua], екологічну безпеку можна розглядати як рівень захищеності життєво важливих інтересів людини, а також суспільства, довкілля та держави від реальних або потенційних загроз, зумовлених антропогенними або природними чинниками.

Головним критерієм для оцінювання екологічної безпеки певної екосистеми є якість життя і здоров'я населення. Отже, виникає необхідність цілеспрямованого управління екологічною системою з

метою підвищення її організованості та досягнення певного ефекту щодо зменшення негативного впливу на населення.

Аналіз останніх досліджень і публікацій. За визначенням американських учених Д. Кліланда і В. Кінга управління можна розглядати як процес, орієнтований на досягнення певних цілей [1]. Для вирішення проблеми необхідно змінити існуючий стан речей (або подій), щоб досягти своєї мети. При всьому різноманітті їх форм вплив можна поділити на два протилежні класи: впливи, що призводять до деградації, руйнування екосистеми, зменшення ступеня її стабільності, та впливи, які сприятимуть розвитку екосистеми, збільшенню ступеня її організованості.

Отже, процес управління екологічною безпекою на будь-якому рівні включає певну сукупність послідовних дій, а саме:

- збір інформації;
- передача інформації в пункти збереження та її обробка;
- аналіз збереженої та довідкової інформації;
- прийняття рішення на основі проведеного аналізу;
- створення відповідного керуючого впливу;
- доведення цього впливу до об'єкта управління;
- перевірка нового стану об'єкта управління.

Управління екологічною безпекою та раціональним використанням природних ресурсів передбачає аналіз потреб людини в природних ресурсах і з'ясування можливостей природи щодо задоволення цих потреб. Однак необхідно не тільки виявити оптимальне поєднання потреб людини та можливостей природи, але й забезпечити оптимальні умови для переходу від нинішнього далеко не ідеального стану до більш ефективного варіанту такого поєднання.

На підставі власного досвіду роботи в галузі екологічної безпеки та на основі аналізу наукових джерел [2; 3, с. 85–102; 4, с. 96–115; 5] визначено, що основною метою досліджень в даному напрямі є розробка і впровадження сучасних механізмів управління екологічною безпекою та станом навколишнього середовища. Під словами «механізми управління» будемо розуміти сукупність певних методів та засобів управління екологічною безпекою й природокористуванням. У свою чергу, метод управління визначає набір способів, прийомів, засобів впливу на керований об'єкт.

Метою статті є аналіз принципів та методів управління екологічною безпекою на основі інтелектуального аналізу даних мережі моніторингу атмосферного повітря.

Результати дослідження. У нашій країні прийняття управлінських рішень та реалізація конкретних регулюючих заходів, які необхідно здійснити щодо збереження навколишнього середовища, відбуваються із затримкою. Час затримки $\tau_{\text{ЗАТ}}$ може бути значним і залежить від багатьох факторів. За цей час система може вийти за межі області допустимих станів, які відповідають прийнятному рівню якості навколишнього середовища, тобто нормальним умовам для проживання та діяльності людей на досліджуваній території. Більше того, після досягнення $\tau_{\text{ЗАТ}}$ певного критичного значення екологічна система може перейти в область критичного відхилення, де незначні зміни призведуть до незворотних процесів деградації.

Представимо об'єкт управління екологічною безпекою і станом оточуючого середовища у вигляді двох взаємодіючих підсистем:

- 1) природно-територіальний комплекс, тобто компоненти, на які впливають об'єкти промисловості й транспорту на досліджуваній території, де поширюються та накопичуються забруднюючі речовини;
- 2) безпосередньо промислові об'єкти і транспорт (ПОТ), які функціонують на досліджуваній території і належать підсистемі стаціонарних джерел забруднення та підсистемі пересувних джерел забруднення.

На рис. 1 наведено узагальнену структурну модель об'єкта управління екологічною безпекою [3, с. 85–102], побудовану з урахуванням підходу, викладеного в роботі [4, с. 96–115]. На схемі позначено окремі інформаційні потоки, які визначають взаємодію між параметрами: X – вектор стану (якості) компонентів природного середовища (концентрації забруднень, рівень енергетичних впливів тощо), тобто це контрольовані показники якості компонентів природного середовища; Z – параметри ПОТ, які визначають умови і результат функціонування окремих об'єктів як з економічного (обсяги, якість продукції та послуг), так і з екологічного погляду (потужність негативного впливу на природне середовище). Для даної моделі $Z = \{Z_c, Z_n\}$, де Z_c, Z_n – множина станів відповідно стаціонарних та пересувних об'єктів ПОТ. При цьому x_c, x_n – множина впливів природного середовища на підсистеми ПОТ ($x_c, x_n \subset X$); z'_c, z'_n – множина результатів економічної діяльності об'єктів ПОТ ($z'_c \subset Z_c, z'_n \subset Z_n$); $\omega_{\text{ПС}}$ – множина зовнішніх впливів на компоненти природного середовища (ПС). У такому вигляді мож-

на врахувати природні та кліматичні параметри, особливості інфраструктури регіону тощо. Враховано також додаткові зовнішні впливи, пов'язані з ПОТ, а саме: $\omega_{\text{ПОТ}}$ – множина зовнішніх впливів на ПОТ, які впливають на результати їх діяльності, виробничі й транспортні параметри, що визначають рівень екологічної небезпеки; $\omega_{\text{ЗС}}$ – множина впливів на зовнішнє середовище через компоненти природного середовища, що надходять від ПОТ; $U_c, U_{\text{П}}$ – множина керуючих впливів відповідно на стаціонарні та пересувні ПОТ.



Рис. 1. Узагальнена модель управління екологічною безпекою територій, що зазнають впливу промислових об'єктів і транспорту

Особам, що приймають рішення (ОПР) щодо виникнення в момент t несприятливої екологічної ситуації, необхідно завчасно надати інформацію для розробки керуючих впливів $U_c(t)$ або $U_{\text{П}}(t)$, за допомогою яких можна контролювати та зменшувати негативний вплив на компоненти природного середовища: $\Delta X(t) < \varepsilon$, $\varepsilon \rightarrow 0$.

Розглянемо детальніше поняття стратегічного і тактичного управління екологічною безпекою, оскільки вони мають вирішальне значення в процесах управління. Стратегічне і тактичне управління екологічною безпекою мають різні цілі та засоби їх досягнення. Якщо стратегічне управління являє собою пошук і реалізацію нових можливостей при змінах зовнішнього середовища, то тактичне управління є процесом створення передумов для реалізації певних нових можливостей.

Тактичне управління можна визначити як засіб реалізації та деталізації відповідної стратегії. Воно розраховано на короткостроковий та середньостроковий періоди. Головною метою тактичного управління екологічною безпекою є вплив на життєдіяльність досліджуваної екосистеми, спрямований на досягнення стратегічних цілей та завдань при найбільш повному і найбільш раціональному використанні доступних ресурсів. Зрештою, засобом реалізації стратегічних планів є тактичне планування. Якщо основна мета стратегічного плану полягає у визначенні щодо перспективи на майбутнє, то тактичне планування допомагає знайти конкретні відповіді на питання, як поетапно досягти бажаного стану.

На рис. 2 наведено концептуальну схему системи управління екологічною безпекою урбанізованої території за даними екологічного моніторингу з використанням сучасних методів і технологій інтелектуального аналізу даних [3, с. 85–102]. Дано пояснення основних блоків, зазначених на схемі.

Блок «Бази даних та алгоритми оцінювання ситуації і ухвалення управлінських рішень в галузі екологічної безпеки» складається з блоків «Управління даними» та «Система підготовки рішень (на основі економічної оцінки ризик-ціна-ефект)». У свою чергу, блок «Система підготовки рішень» включає «Інтелектуальний аналіз даних» і «Систему підтримки прийняття рішень (СППР)».

Блок «Управління даними» забезпечує збереження даних і доступ до них. До таких даних відносять дані моніторингу, розподілені бази даних (БД) аналітичної та звітної інформації та розподілені БД дистанційного зондування Землі (ДЗЗ), які перебувають у сховищі даних. Під розподіленими БД розуміємо сукупність взаємопов'язаних на логічному рівні БД, розподілених у комп'ютерній мережі. Інтелекту-



альний аналіз даних у наведеній схемі включає експертні системи, геоінформаційні та інтеграційні технології (ДЗЗ, Інтернет).

В цілому інтелектуальний аналіз даних у задачах екологічної безпеки можна визначити як метод підтримки прийняття рішень, заснований на дослідженні певних закономірностей щодо вихідних даних. Він включає методики та засоби, які на основі будь-яких моделей, алгоритмів або математичних теорем дають можливість оцінювати значення невідомих характеристик і параметрів за відомими даними.

Алгоритми інтелектуального аналізу. На основі публікацій [6, с. 1–21; 7, с. 3558–3571; 8, с. 1–15; 9; 10; 11, с. 62–75; 12; 13, с. 139–151] визначено найпоширеніші алгоритми інтелектуального аналізу даних, спрямовані на виявлення прихованих закономірностей або взаємозв'язків між змінними у великих масивах необроблених даних.

Класифікатор C4.5 створює класифікацію у вигляді дерева рішень. Для цього задається набір даних, який являє собою вже класифіковані речі. Алгоритм C4.5, розроблений J.R. Quinlan [10], є вдосконаленою версією алгоритму ID3 того ж автора. Зокрема, в C4.5 додано відсікання гілок, можливість роботи з числовими атрибутами, а також можливість побудови дерева з неповною навчальною вибіркою, в якій відсутні значення деяких атрибутів.

Для того щоб за допомогою C4.5 побудувати дерево рішень і застосовувати його, дані повинні задовольняти декільком умовам, а саме:

- інформація про об'єкти, що необхідно класифікувати, повинна бути представлена у вигляді кінцевого набору ознак (атрибутів), кожен з яких має дискретне або числове значення. Якщо такий набір атрибутів назвати прикладом, тоді для всіх прикладів кількість атрибутів і їх склад повинні бути постійними;
- кожен із класів, на які будуть розбиватися приклади, повинен мати кінцеве число елементів, а кожен приклад повинен однозначно відноситись до конкретного класу;
- для випадків з нечіткою логікою, тобто коли приклади можна віднести до одного з класів з певною ймовірністю, C4.5 непридатний;
- у навчальній вибірці кількість прикладів має бути значно більше кількості класів і кожен приклад повинен бути заздалегідь асоційований зі своїм класом. За цими ознаками C4.5 віднесено до методів машинного навчання з учителем.

Алгоритм кластерного аналізу K-means виділяє певну кількість груп k з набору об'єктів таким чином, щоб члени кожної групи були максимально схожими між собою. Це досить популярний метод кластерного аналізу для вивчення набору даних. Нагадаємо, що кластерний аналіз – це сімейство алгоритмів, призначених для формування груп, де в кожній групі члени даної групи мають більше схожих рис з представниками своєї групи, ніж з тими, хто в цій групі не перебуває (в даному контексті кластери та групи є синонімами) [3, с. 85–102].

Метод K-means – це метод, метою якого є поділ m спостережень (з простору даних) на k кластерів таким чином, що кожне спостереження буде віднесено до того кластеру, до центру (центроїду) якого воно буде найближчим. Як міра близькості використовується Евклідова відстань:

$$\|x - y\| = \sqrt{\sum_{p=1}^n (x_p - y_p)^2},$$

де $x, y \in R^n$.

Для ряду спостережень $(x^1, x^2, \dots, x^m)$, $x_j \in R^n$, метод K-means мінімізує сумарне квадратичне відхилення точок кластерів від центроїдів цих кластерів:

$$\min \left[\sum_{i=1}^k \sum_{x^{(j)} \in S_i} \|x^{(j)} - \mu_i\|^2 \right],$$

де $x^{(j)} \in R^n$, $\mu_i \in R^n$, μ_i – центр ваги для кластера S_i .

Зауважимо, що при використанні даного методу неправильний вибір початкового числа кластерів k може призвести до некоректних результатів. Саме тому при використанні методу K-means на початковому етапі необхідно визначити відповідне число кластерів для визначеного набору даних.

Метод опорних векторів (SVM) знаходить гіперплощину для класифікації даних на два класи. На відміну від C4.5 в ньому не використовуються дерева рішень.

SVM – це лінійний алгоритм, який використовується в задачах класифікації та регресії. Він широко застосовується на практиці і за його допомогою можна вирішувати як лінійні, так і нелінійні задачі. Суть цього методу полягає у створенні лінії або гіперплощини, яка розділяє дані на класи. Якщо є два

класи спостережень і передбачається лінійна форма кордону між класами, то можливі два випадки.

Перший з них пов'язаний з можливістю ідеального поділу даних за допомогою деякої гіперплощини $z_k(x) = \sum_{i=1}^p \beta_i x_i + \beta_0$.

На рис. 3, а наведено двовимірний варіант. Оскільки таких гіперплощин може бути безліч, оптимальною буде та з них, яка максимально віддалена від навчальних точок, тобто має максимальний проміжок (зазор) M [12].

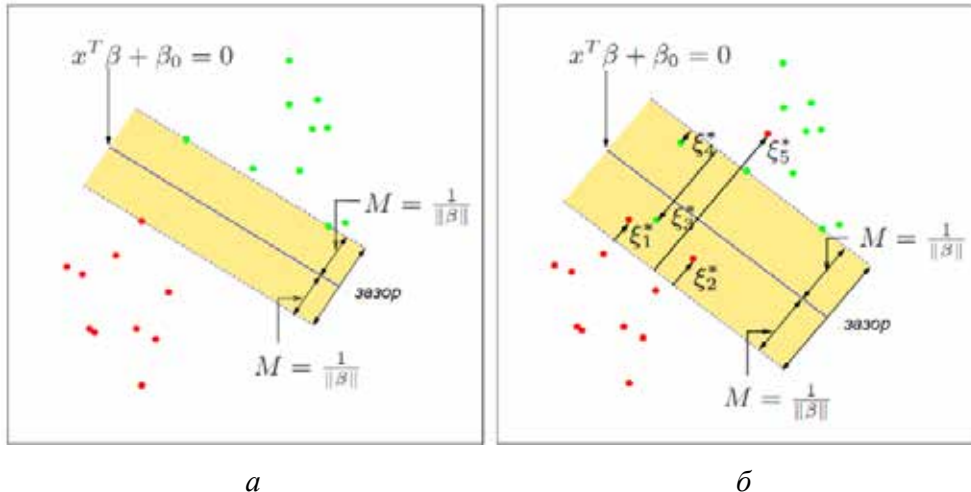


Рис. 3. Класифікатори з мінімальним зазором (а) і на опорних векторах (б)

Другий випадок показано на рис. 3, б, коли множина точок перекривається і обидва класи лінійно нероздільні. Власне опорними векторами називаються значення, які лежать безпосередньо на кордоні, що розділяє смуги, або на неправильній для свого класу стороні щодо кордонів зазору (такі точки позначені ξ_j^*). Для граничних і всіх інших точок прийнято $\xi_j^* = 0$.

Оптимальну розділяючу гіперплощину такого класифікатора $z_k(x)$ можна визначити за умови максимізації ширини зазору M , але при цьому дозволено невірно класифікувати деяку невелику групу спостережень, що відносяться до опорних векторів. Математично пошук рішення зведено до задачі квадратичної оптимізації з лінійними обмеженнями, яка гарантовано сходиться до одного глобального мінімуму.

Алгоритм класифікації kNN (k найближчих сусідів) відрізняється від раніше описаних тим, що він нагадує «лінивого учня». В основі методу kNN-класифікатора лежить гіпотеза компактності, яка передбачає, що тестований об'єкт d матиме таку ж мітку класу, як і навчальні об'єкти в локальній області його найближчого оточення. У варіанті kNN кожен об'єкт відносимо до пріоритетного класу найближчих сусідів, де k – параметр алгоритму.

Вирішальні правила в методі kNN визначаються межами суміжних сегментів діаграми Вороного, що розділяє площину на n опуклих багатокутників, кожен з яких містить один і тільки один об'єкт навчальної вибірки (рис. 4). В p -мірних просторах границі розв'язків складаються вже з сегментів $(p-1)$ -мірних напівплощин, утворених опуклими многогранниками Вороного [12].

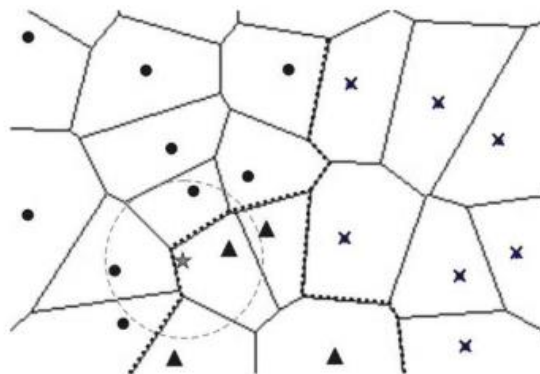


Рис. 4. Приклад роботи алгоритму класифікації найближчих сусідів

Алгоритм передбачення будується за принципом «більшості голосів», тобто за результатами голосування вибирається мітка класу-переможця. На рис. 4 тестований об'єкт «зірочка» потрапляє в осередок об'єктів класу «трикутників» і при $k = 1$ буде віднесений до цього класу. Однак при $k = 3$ за «голосами» двох найближчих сусідів з трьох екзаменованих цей об'єкт буде віднесено до класу «кружечків». Імовірнісний варіант методу kNN використовує для ранжирування передбачуваних класів суму голосів сусідів з урахуванням їх ваг, зокрема, евклідової міри відстані між тестованим об'єктом і кожним із сусідів.

Варіант 1NN завжди забезпечує 100% правильного розпізнавання прикладів навчальної вибірки (самий найближчий сусід – він сам), проте часто помиляється на невідомих йому даних. При збільшенні k від одиниці до деяких меж якості розпізнавання на контрольній вибірці буде зростати. Оптимальне щодо точності прогнозів значення k може бути знайдено з використанням перехресної перевірки. Для цього по фіксованому значенню k будується модель k найближчих сусідів і оцінюється CV-помилка класифікації. Ці дії повторюються для різних значень k і значення з найменшою помилкою розпізнавання приймається як оптимальне.

Наївний байєсів класифікатор включає сімейство алгоритмів класифікації, які поділяють одне спільне припущення. В основі байєсівської класифікації лежить гіпотеза максимальної ймовірності, тобто об'єкт d належить класу c_j ($c_j \in C$), і при цьому досягається найбільша апостеріорна ймовірність $\max P(c_j|d)$. За формулою Байєса:

$$P(c_j|d) = \frac{P(c_j)P(d|c_j)}{P(d)} \approx P(c_j)P(d|c_j),$$

де $P(d|c_j)$ – ймовірність зустріти об'єкт d серед об'єктів класу c_j ; $P(c_j)$ і $P(d)$ – апіорні ймовірності класу c_j і об'єкта d (остання не впливає на вибір класу і її можна не враховувати).

Якщо зробити «наївне» припущення, що всі ознаки, за якими класифікуються об'єкти, абсолютно рівноправні між собою і не пов'язані одна з одною, то $P(d|c_j)$ можна обчислити як добуток ймовірностей зустріти ознаку x_i ($x_i \in X$) серед об'єктів класу c_j :

$$P(d|c_j) = \prod_{i=1}^{|X|} P(x_i|c_j),$$

де $P(x_i|c_j)$ – ймовірнісна оцінка вкладу ознаки x_i в те, що $d \in c_j$.

На практиці при перемноженні дуже малих умовних ймовірностей може спостерігатися втрата значущих розрядів. У зв'язку з цим замість оцінок ймовірностей $P(x_i|c_j)$ застосовують логарифми цих ймовірностей. Оскільки логарифм є монотонно зростаючою функцією, то клас c_j з найбільшим значенням логарифма ймовірності залишиться найбільш ймовірним. Тоді вирішальне правило наївного байєсового класифікатора (Naive Bayes Classifier) приймає наступний остаточний вигляд [12]:

$$c^* = \arg_{c_j \in C} \max \left[\log P(c_j) + \sum_{i=1}^X P(x_i|c_j) \right].$$

Алгоритм Apriori належить до масштабованих алгоритмів, спрямованих на пошук асоціативних правил в БД. Сучасні БД мають дуже великі розміри, а також тенденцію щодо подальшого їх зростання. Пошук асоціативних правил в таких БД спрямований на отримання даних для вивчення кореляцій і взаємозв'язків між змінними. Алгоритм Apriori реалізує такий пошук в два етапи. На першому етапі здійснюється пошук вибірок, які часто зустрічаються в даній базі, а на другому – виявляються асоціативні правила, що є достовірними для цих вибірок.

Приклади інтелектуального аналізу даних. Інтелектуальний аналіз даних моніторингу стану атмосферного повітря (АП) урбанізованих територій проведено в рамках інформаційно-аналітичної системи еколого-енергетичного моніторингу AISEEM [3, с. 85–102] та програми Orange [14], яка є сучасним інструментом для візуалізації даних, машинного навчання та інтелектуального аналізу даних. Взаємодія між цими програмами здійснювалась за допомогою інтеграції методів та засобів Orange в систему AISEEM.

Дані щодо забруднюючих речовин завантажувались з автоматизованої системи моніторингу за станом АП м. Кривий Ріг (Дніпропетровська область). Основною метою системи моніторингу є безперервні виміри концентрацій забруднюючих речовин і метеорологічних параметрів АП жилої зони м. Кривий Ріг [15].

На рис. 5 темно-зеленим кольором позначені міські автоматизовані пости, темно-синім – пости підприємства ПАТ «Південний гірничо-збагачувальний комбінат», темно-червоним – ПРАТ «СУХА БАЛКА» та помаранчевим – ПАТ «Кривбасзалізрудком». Спостереження здійснено за шістьма забруднюючими речовинами: діоксид азоту (NO_2), оксид азоту (NO), сірчаний ангідрид (SO_2), оксид вуглецю (CO), аміак (NH_3), сірководень (H_2S), пил. Крім того, на міських автоматизованих постах спостережень (ПАС) № 3

та № 5 здійснено спостереження за етаном (C_2H_6) та озоном (O_3), а метеорологічними параметрами є середня швидкість вітру, напрямок вітру, температура, відносна вологість та тиск.

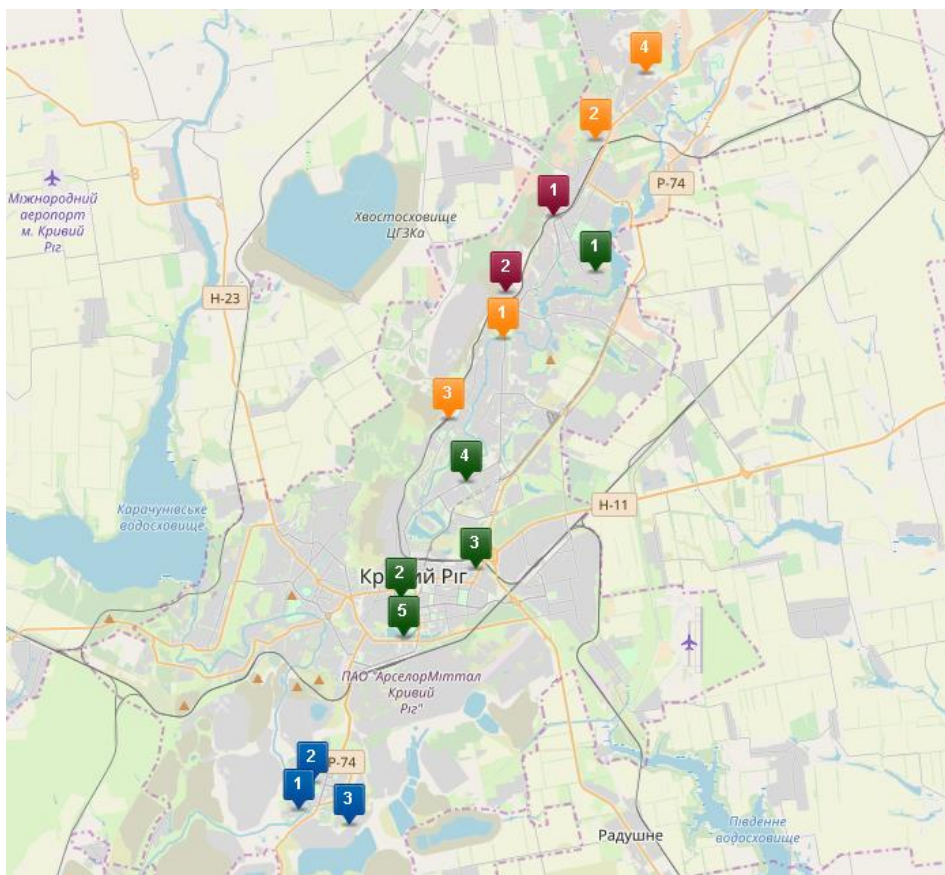


Рис. 5. Автоматизовані пости спостереження за станом АП

Варто також зазначити, що на території міста виробництво теплової енергії здійснює ПАТ «Криворізька теплоцентраль», на балансі якої є шість опалювальних котелень, обладнаних 31 паровими та водогрійними котлами з установленою та наявною потужностями відповідно 1406 Гкал/год та 1081 Гкал/год. Найшкідливіші викиди цього енергетичного об'єкта – це сірчаний ангідрид, оксиди азоту та пил.

При аналізі даних було виключено інформацію щодо напрямку вітру, оскільки напрямок вітру змінюється від 0° до 360° , і для 358° та 2° фактично отримаємо північний напрямок вітру згідно з даними [15] з зовсім різними значеннями, що ускладнює їх подальше трактування.

Наведемо приклади інтелектуального аналізу даних для погодинних максимальних разових концентрацій забруднюючих речовин та метеорологічних даних на міському ПАС № 1 з 10 по 16 грудня 2018 року.

На рис. 6 зображено діаграму розсіювання, де на осі абсцис показано рівні концентрацій пилу, а на осі ординат – діоксиду азоту. Як бачимо, концентрації пилу на цьому часовому інтервалі змінювались від 0,156 до 0,284, а концентрації діоксиду азоту – від 0,027 до 0,091. На рис. 7 на осі абсцис показано температуру повітря ($^\circ C$), а на осі ординат – концентрації сірчаного ангідриду. Температура на цьому часовому інтервалі змінювалась від $-4,1^\circ C$ до $+4,6^\circ C$, а концентрації сірчаного ангідриду – від 0,018 до 0,021 mg/m^3 .

Підкреслимо, що на даному етапі дослідження наведено лише попередні результати аналізу даних моніторингу. Адже застосування інтелектуальних технологій для виявлення нових знань та закономірностей потребує значно більшої кількості даних щодо забруднення атмосфери, викидів транспорту та промислових підприємств, захворюваності населення тощо. Маємо надію, що такі дані будуть доступні після реалізації відповідних загальнодержавних програм, що дасть можливість здійснювати більш ґрунтовний аналіз даних моніторингу стану атмосферного повітря.

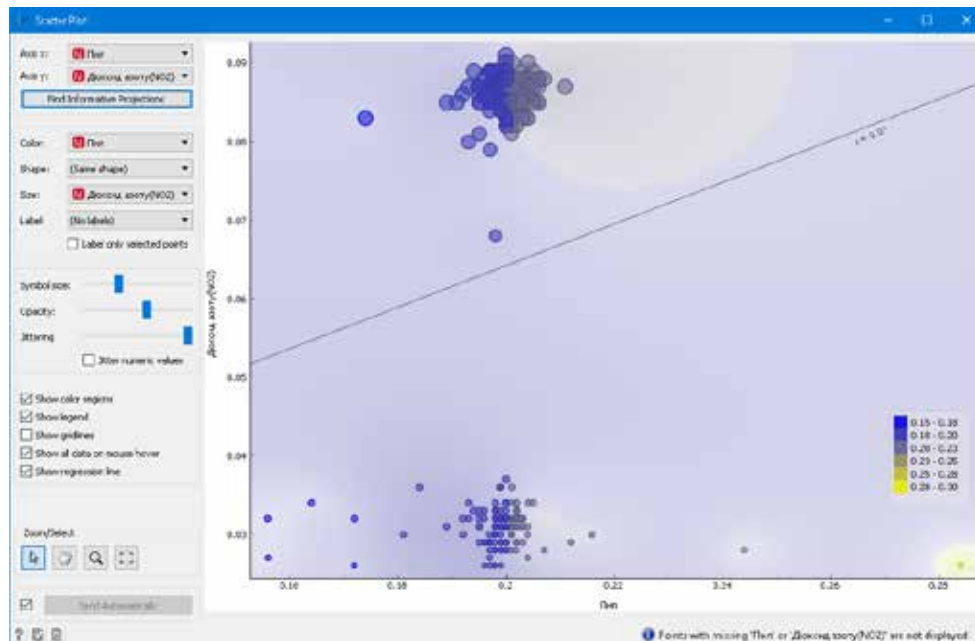


Рис. 6. Діаграма розсіювання концентрацій пилу та діоксиду азоту, (мг/м³)

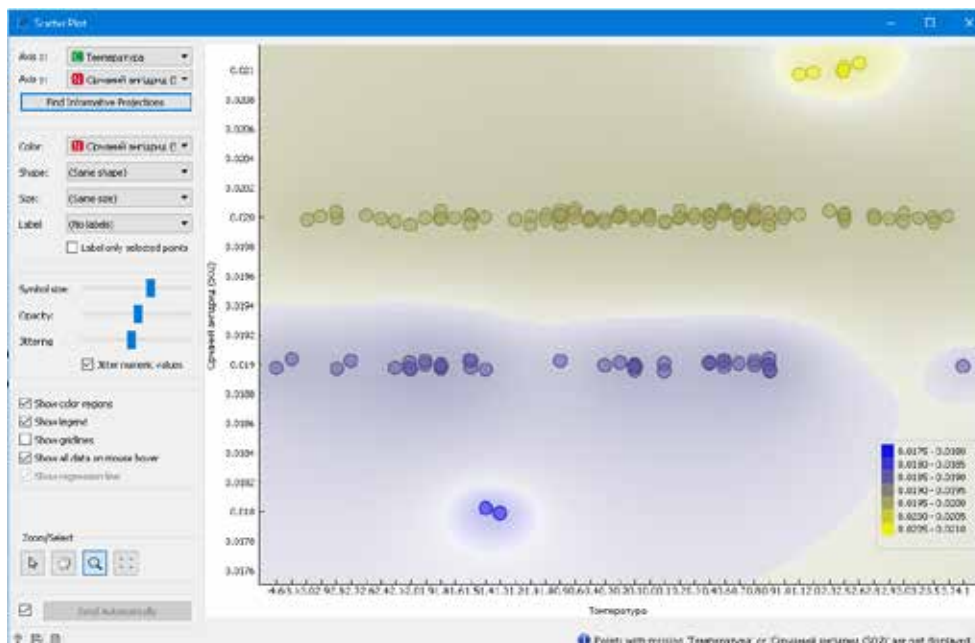


Рис. 7. Діаграма розсіювання температури повітря (°C) і концентрації сірчаного ангідриду

Висновки. Запропонована концептуальна модель управління екологічною безпекою урбанізованих територій за даними екологічного моніторингу, що відрізняється від своїх аналогів новими можливостями, базованими на використанні сучасних інформаційних технологій інтелектуального аналізу даних. Визначено перспективи застосування методів та засобів інтелектуального аналізу даних щодо інформаційної підтримки прийняття рішень, спрямованих на оцінювання наслідків техногенного впливу і зменшення обсягу навантажень на довкілля. Запропоновано ряд методів та алгоритмів, які дають можливість оцінювати значення невідомих характеристик і параметрів за відомими даними, та показано приклади їх використання.

Список використаних джерел:

1. Клиланд Д., Кинг В. Системный анализ и целевое управление. Москва : Сов. радио, 1974. 280 с.
2. Пампура В.И. Оптимальное управление безопасностью экологически опасных объектов. К. : Наукова думка, 2012. 599 с.
3. Яцишин А.В., Куцан Ю.Г., Артемчук В.О. та ін. Принципи та методи управління екологічною безпекою на основі інтелектуального аналізу даних мережі моніторингу атмосферного повітря. *Електронне моделювання*. 2019. № 4(41). С. 85–102.
4. Ивашук О.А., Константинов И.С. Обеспечение адаптивного управления экологической безопасностью промышленно-транспортного комплекса. *Управление большими системами*. 2009. Вып. 25. С. 96–115.
5. Капралов Е.Г., Кошкарев А.В., Тикунов В.С. Геоинформатика : учебник для вузов. М. : Академия, 2008. 384 с.
6. Chen C., Chuang C., Jiang J. Ecological Monitoring Using Wireless Sensor Networks. Overview, Challenges, and Opportunities. *Advancement in Sensing Technology. Smart Sensors, Measurement and Instrumentation*. 2013. Vol. 1. P. 1–21.
7. Dias D., Tchepel O. Modelling of Human Exposure to Air Pollution in the Urban Environment: a GPS-based approach. *Environmental Science and Pollution Research*. 2014. Vol. 5. P. 3558–3571.
8. Peters D.P.C., Havstad K.M., Cushing J. et al. Harnessing the Power of Big Data: Infusing the Scientific Method With Machine Learning to Transform Ecology. *Ecosphere*. 2014. Vol. 5(6). P. 1–15.
9. Алгоритмы интеллектуального анализа данных. 2015. URL: <https://tproger.ru/translations/top-10-data-mining-algorithms/>.
10. Деревья решений – C4.5 математический аппарат. Часть 1. 2019. URL: <https://basegroup.ru/community/articles/math-c45-part1>.
11. Згуровский М.З., Болдак А.А., Ефремов К.В. Интеллектуальный анализ и системное согласование научных данных в междисциплинарных исследованиях. *Кибернетика и системный анализ*. 2013. № 4. С. 62–75.
12. Шитиков В.К., Мاستицкий С.Э. Классификация, регрессия, алгоритмы Data Mining с использованием R. 2017. URL: <https://github.com/ranalytics/data-mining>.
13. Ashouri M., Haghighat F., Fung B.C.M. et al. Development of Building Energy Saving Advisory: A Data Mining Approach. *Energy and Buildings*. 2018. Vol. 172. P. 139–151.
14. Orange. 2021. URL: <http://orange.biolab.si/download/>.
15. Кривий Ріг. Автоматизовані пости спостереження. 2018. URL: <https://krmisto.gov.ua/ua/rc/ecomon.html>.

References:

1. Cleland, D., King, V. (1974). *Sistemnyy analiz i tselevoye upravleniye [System Analysis and Target Management]*. Moscow: Sov. Radio [in Russian].
2. Pampuro, V.I. (2012). *Optimal'noye upravleniye bezopasnost'yu ekologicheskikh opasnykh ob'yektov [Optimal safety management of environmentally hazardous facilities]*. Kiev, Ukraine: Naukova Dumka [in Russian].
3. Iatsyshyn, A.V., Kutsan, Yu.H., Artemchuk, V.O. et al. (2019). Pryntsypy ta metody upravlinnya ekolohichnoyu bezpekoyu na osnovi intelektual'noho analizu danykh merezhi monitorynhu atmosfernoho povitrya. [The principles and methods of ecological safety management through the data of air monitoring network analysis]. *Elektron. Model*, no. 41(4), pp. 85–101 [in Ukrainian].
4. Ivashchuk, O.A., Konstantinov, I.S. (2009). Obespecheniye adaptivnogo upravleniya ekologicheskoy bezopasnost'yu promyshlenno-transportnogo kompleksa. [Guarding of adaptive management by ecological safety of an industrial and transport complex]. *Large-Scale Systems Control*, no. 25, pp. 96–115 [in Russian].
5. Kapralov, E.G., Koshkarev, A.V., Tykunov, V.S. (2008). *Geoinformatika: uchebnik dlya vuzov [Geoinformatics: a textbook for universities]*. Moscow : Academy [in Russian].
6. Chen, C., Chuang, C., Jiang, J. (2013). Ecological Monitoring Using Wireless Sensor Networks. Overview, Challenges, and Opportunities. *Advancement in Sensing Technology. Smart Sensors, Measurement and Instrumentation*, no. 1, pp. 1–21.
7. Dias, D., Tchepel, O. (2014). Modelling of human exposure to air pollution in the urban environment: a GPS-based approach. *Environmental Science and Pollution Research*, no. 5, pp. 3558–3571.
8. Peters, D.P.C., Havstad, K.M., Cushing, J. et al. (2014). Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology. *Ecosphere*, no. 5(6), pp. 1–15.
9. Algoritmy intelektual'nogo analiza dannykh. [Algorithms of the intellectual data analysis] (2015). URL: <https://tproger.ru/translations/top-10-data-mining-algorithms/> [in Russian].
10. Derev'ya resheniy – C4.5 matematicheskii apparat. Chast' 1. [Decision trees – C4.5 mathematical apparatus. Part 1]. (2019). URL: <https://basegroup.ru/community/articles/math-c45-part1> [in Russian].
11. Zgurovsky, M.Z., Boldak, A.A., Yefremov, K.V. (2013). Intellektual'nyy analiz i sistemnoye soglasovaniye nauchnykh dannykh v mezhdistsiplinarnykh issledovaniyakh. [Intelligent analysis and systemic adjustment of scientific data in interdisciplinary research]. *Cybernetics and Systems Analysis*, no. 49, pp. 62–75 [in Russian].
12. Shitikov, V.K., Mastitsky, S.E. (2017). Klassifikatsiya, regressiya, algoritmy Data Mining s ispol'zovaniyem R. [Classification, regression, Data Mining algorithms using R]. URL: <https://github.com/ranalytics/data-mining> [in Russian].
13. Ashouri, M., Haghighat, F., Fung, B.C.M. et al. (2018). Development of building energy saving advisory: A data mining approach. *Energy and Buildings*, no. 172, pp. 139–151.
14. Orange. (2021). URL: <http://orange.biolab.si/download/>.
15. Kryvyy Rih. Avtomatyzovani posty sposterezhennya. (2018). [Kryviy Rih. Automated surveillance posts]. URL: <https://krmisto.gov.ua/ua/rc/ecomon.html> [in Ukrainian].

УДК 519.652:519.254 (045)

DOI <https://doi.org/10.32689/maup.it.2021.2.3>

Пилип ПРИСТАВКА

доктор технічних наук, професор, завідувач кафедри прикладної математики, Національний авіаційний університет, вул. Любомира Гузара 1, Київ, Україна, індекс 03058 (chindakor37@gmail.com)

ORCID: <https://orcid.org/0000-0002-0360-2459>

Ольга ЧОЛИШКІНА

кандидат технічних наук, доцент, директор, Інститут комп'ютерно-інформаційних технологій та дизайну, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом», вул. Фрометівська 2, Київ, Україна, індекс 03039 (greenhelga5@gmail.com)

ORCID: <https://orcid.org/0000-0002-0681-0413>

Рулуп PRYSTAVKA

Doctor of Technical Sciences, Professor, Head of the Department of Applied Mathematics, National Aviation University, 1 Liubomyra Huzara Street, Kyiv, Ukraine, postal code 03058 (chindakor37@gmail.com)

Olha CHOLYSHKINA

Candidate of Technical Sciences, Associate Professor, Director, Institute of Computer Information Technology and Design, Interregional Academy of personnel management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039 (greenhelga5@gmail.com)

Бібліографічний опис статті: Приставка П., Чолишкіна О. Часткові випадки локальних поліноміальних сплайнів, близьких до інтерполяційних у середньому другого та третього порядків. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 26–33. DOI: <https://doi.org/10.32689/maup.it.2021.2.3>

Bibliographic description of the article: Prystavka, P., Cholyshkina, O. (2021). Chastkovi wypadky lokalnykh polinomialnykh splainiv, blyzkykh do interpolatsiinykh u serednomu druhohto ta tretoho poriadkiv [Partial cases of local polynomial splines close to interpolations in the middle of the second and third orders]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 26–33. DOI: <https://doi.org/10.32689/maup.it.2021.2.3>

**ЧАСТКОВІ ВИПАДКИ ЛОКАЛЬНИХ ПОЛІНОМІАЛЬНИХ СПЛАЙНІВ,
БЛИЗЬКИХ ДО ІНТЕРПОЛЯЦІЙНИХ У СЕРЕДНЬОМУ ДРУГОГО ТА ТРЕТЬОГО ПОРЯДКІВ**

Анотація. Розглянуто задачі цифрової обробки сигналів та послідовностей. Для даного типу задач отримано лінійні оператори, які є частковими випадками локальних поліноміальних сплайнів другого порядку першого та другого ступеня уточнення та третього порядку першого ступеня уточнення, що є близькими до інтерполяційних у середньому. Отримані в статті формули можуть мати застосування при програмуванні, коли критичним є час обробки великих масивів даних. Також коли є потреба в оптимізації розрахунків в режимі реального часу. В статті наведено приклад застосування при реалізації обчислювальних процедур небінарного subdivision.

Ключові слова: локальні поліноміальні сплайни, subdivision-метод, оптимізація розрахунків.

**PARTIAL CASES OF LOCAL POLYNOMIAL SPLINES CLOSE TO INTERPOLATIONS
IN THE MIDDLE OF THE SECOND AND THIRD ORDERS**

Abstract. The problems of digital signal processing and sequences are considered. For this type of problems, linear operators are obtained, which are partial cases of local polynomial splines of the second order of the first and second degree of refinement and the third order of the first degree of refinement, which are close to interpolation on average. The formulas obtained in the paper can be used in programming when the processing time of large data sets is critical. Also when there is a need to optimize calculations in real time. The paper gives an example of the application of non-binary subdivision in the implementation of computational procedures.

Key words: local polynomial splines, subdivision method, calculation optimization.

Постановка проблеми. На сьогодні задачі обробки цифрових сигналів, зображень та відео характеризуються, в першу чергу, великим обсягом даних, що за вимогою мають оброблятися в режимі реального часу. Найвищу швидкість серед методів, що забезпечують операції фільтрації, масштабування, стиснення та які реалізуються при автоматизації розрахунків, показують ті, в основу яких покладено використання лінійних операторів, як таких, що мають найменшу обчислювальну складність. По суті застосування таких операторів являє собою дискретну згортку цифрової послідовності з симетричною або несиметричною маскою, коефіцієнти якої часто отримані, як частковий випадок деякого неперервного наближення. Наприклад, часткові випадки гаусіана та лапласіана використовують в задачі субполосної фільтрації, відповідно, як низько- та високочастотні фільтри [1].

Зростання обсягів інформації при обробці цифрових сигналів характеризується не просто формальним збільшенням кількості даних, а ще й зміною їх властивостей. Зокрема, при обробці цифрових фотографій спостерігається тенденція до збільшення розрішення систем фіксації, що може вимагати збільшення ширини масок операторів згортки. Тож актуальним є дослідження методів апроксимації, що мають високі апроксимативні властивості та низьку обчислювальну складність водночас, задля отримання на їх основі нових лінійних операторів для цифрової обробки сигналів.

Аналіз досліджень та постановка задачі. Однією із поширених задач при обробці одновимірних цифрових послідовностей є кратне або некрратне їх масштабування. Дана операція поширена при стисненні сигналів, виділення інформативних складових або при дослідженні локальних особливостей. Варто окремо звернути увагу на дві близькі, але все ж таки різні за суттю області обробки, що базуються з одного боку на методах кратного масштабування аналізу (КМА) [2], а з другого – *subdivision*-методи [3]. При багатьох точках дотику цих двох обчислювальних технологій можливим є відзначити, що в більшості випадків процедури КМА базуються на операторах, що згладжують, а процедури *subdivision* – на операторах, що близькі до інтерполяційних.

Якщо говорити про наближення, близькі до інтерполяційних, то з точки зору якості неперервних апроксимацій функцій, заданих послідовностями відліків у вузлових точках, з урахуванням вимоги низької обчислювальної складності відповідних процедур, позиції лідера займають методи, засновані на використанні лінійних комбінацій *B*-сплайнів [4–8]. В роботі [9] обґрунтовано лінійні комбінації *B*-сплайнів, близькі до інтерполяційних у середньому в якості моделі цифрових сигналів з кінцевою енергією, показано можливість їх реалізації у програмному забезпеченні обробки цифрових сигналів, в тому числі для систем, що функціонують у режимі реального часу.

Вперше такі сплайни на основі *B*-сплайнів другого порядку були введені в роботах А.О. Лигуна та В.В. Кармазіної [10]. Обчислювальний аспект їх застосування в задачі бінарного *subdivision* подано в роботах [8; 11]. Частковий випадок прикладу застосування уточнюючого локального сплайну на основі *B*-сплайнів другого порядку в задачі не бінарного *subdivision* наведено в роботі [12].

Поставимо за мету даної роботи отримати часткові випадки локальних поліноміальних сплайнів, близьких до інтерполяційних у середньому на основі *B*-сплайнів другого та третього порядків, що можуть мати застосування при розробці автоматизованих систем обробки цифрових сигналів та послідовностей.

Виклад основного матеріалу. Нехай з кроком $h > 0$ задано розбиття дійсної вісі $\Delta_h: t_i = ih, i \in \mathbb{Z}$, у кожній точці якого отримано значення деякої неперервної функції $p(t)$, визначеної на $R_1(-\infty; \infty)$. Будемо вважати, що інформація про функцію $p(t)$, яка підлягає відтворенню, задано у вузлах розбиття Δ_h у вигляді інтеграла

$$\bar{p}_i = \frac{1}{h} \int_{(i-0,5)h}^{(i+0,5)h} p(t) dt,$$

при цьому, істинне значення функції $p(t)$ у вузлах будемо визначати

$$p_i = \bar{p}_i + \varepsilon_i, i \in \mathbb{Z},$$

де ε_i – похибка.

Згідно роботи [7] уточнюючі сплайни на основі *B*-сплайнів, близькі до інтерполяційних у середньому, що найчастіше мають використання в практичній діяльності, такі:

$$S_{2,1}(p, t) = \sum_{i \in \mathbb{Z}} \left(p_i - \frac{1}{6} \Delta^2 p_i \right) B_{2,h}(t - (i + 0,5)h),$$

$$S_{2,2}(p, t) = \sum_{i \in \mathbb{Z}} \left(p_i - \frac{1}{6} \Delta^2 p_i + \frac{1}{36} \Delta^4 p_i \right) B_{2,h}(t - (i + 0,5)h),$$

$$S_{3,1}(p, t) = \sum_{i \in Z} \left(p_i - \frac{5}{24} \Delta^2 p_i \right) B_{3,h}(t - ih),$$

де

$$\Delta^{2u} p_i = \Delta^{2u-2} p_{i+1} - 2\Delta^{2u-2} p_i + \Delta^{2u-2} p_{i-1}, \quad u = 1, 2, \dots,$$

$B_{r,h}(t)$, $r = 2, 3, 4$ – B-сплайн, що з точністю до аргументу визначається так:

$$B_{2,h}(t) = \begin{cases} 0, & t \notin [-3h/2; 3h/2], \\ (3+2t/h)^2/8, & t \in [-3h/2; -h/2], \\ 3/4 - (2t/h)^2/4, & t \in [-h/2; h/2], \\ (3-2t/h)^2/8, & t \in [h/2; 3h/2]; \end{cases}$$

$$B_{3,h}(t) = \frac{1}{48} \begin{cases} 0, & t \notin [-2h; 2h], \\ (4+2t/h)^3, & t \in [-2h; -h], \\ -3(2t/h)^3 - 12(2t/h)^2 + 32, & t \in [-h; 0], \\ 3(2t/h)^3 - 12(2t/h)^2 + 32, & t \in [0; h], \\ (4-2t/h)^3, & t \in [h; 2h]. \end{cases}$$

Наведені сплайни мають високі апроксимативні властивості. Зокрема щодо норм сплайн-операторів справедливий наступний твердження [7]:

$$\|S_{2,1}(p, t)\| = \frac{4}{3} \|p(t)\|, \quad \|S_{2,2}(p, t)\| = \frac{3}{2} \|p(t)\|, \quad \|S_{3,1}(p, t)\| = \frac{41}{32} \|p(t)\|,$$

де

$$\|S_{r,u}(p, t)\| = \sup_{|\varepsilon_i|} \max_t |S_{r,u}(\varepsilon, t)|,$$

$$|\varepsilon_i| < \varepsilon, \quad i \in Z.$$

Зауважимо, що значення норми оператора $S_{r,u}(p, t)$ – це величина, яка характеризує в скільки разів може зрости похибка при відтворенні функції за допомогою сплайну, якщо значення p_i задані з похибкою.

Про похибку відтворення функції $p(t)$ за використанням сплайнів $S_{r,u}(\bar{p}, t)$, $r = 2, 3$, $u = 1, 2$ свідчать наступні оцінки [7].

При $h \rightarrow 0$ для довільної функції $p(t) \in C^3$ буде вірно наступне

$$\|p(t) - S_{2,u}(\bar{p}, t)\| = \frac{h^3}{12\sqrt{3}} \|p^{(3)}(t)\| + o(h^3), \quad u = 1, 2,$$

і для довільної функції $p(t) \in C^4$ справедливо

$$\|p(t) - S_{3,1}(\bar{p}, t)\| = \frac{1221h^4}{2880} \|p^{(4)}(t)\| + o(h^4),$$

Подання зазначених сплайнів у вигляді лінійної комбінації B-сплайнів не зовсім зручне для реалізації в обчислювальній середовищі, тож якщо ввести заміну

$$x = \frac{2}{h}(t - (i + 0,5)h), \quad |x| \leq 1, \quad \text{при } r = 2$$

та

$$x = \frac{2}{h}(t - ih), \quad |x| \leq 1, \quad \text{при } r = 3,$$

то сплайни $S_{r,u}(p, t)$, $r = 2, 3$, $u = 1, 2$ можна навести в розгорнутому представленні:

$$S_{2,1}(p, t) = \frac{1}{48} \left(-(1-x)^2 p_{i-2} + (2-16x+10x^2) p_{i-1} + (46-18x^2) p_i + \right. \\ \left. + (2+16x+10x^2) p_{i+1} - (1+x)^2 p_{i+2} \right), \quad (1)$$

$$S_{2,2}(p,t) = \frac{1}{288} \left((1-x)^2 p_{i-3} + (-4+20x-12x^2) p_{i-2} + \right. \\ \left. + (-5-106x+75x^2) p_{i-1} + (304-128x^2) p_i + (-5+106x+75x^2) p_{i+1} + \right. \\ \left. + (-4-20x-12x^2) p_{i+2} + (1+x)^2 p_{i+3} \right), \quad (2)$$

$$S_{3,1}(p,t) = \frac{1}{1152} \left(-5(1-x)^3 p_{i-2} + (-81-27x+117x^2-49x^3) p_{i-1} + \right. \\ \left. + (662-570x-102x^2+122x^3) p_i + (662+570x-102x^2-122x^3) p_{i+1} + \right. \\ \left. + (-81+27x+117x^2+49x^3) p_{i+2} - 5(1+x)^3 p_{i+3} \right). \quad (3)$$

На основі виразів (1)-(3) нескладно забезпечити зміну кількості відліків в послідовності $\{p_i\}_{i \in \mathbb{Z}}$ в довільну кількість раз (необов'язково у цілочисельну). Але, якщо зміна масштабу послідовності здійснюється на певний наперед відомий коефіцієнт, то в цьому разі для апроксимацій, що мають явний вигляд як наведені вище, можна отримати обчислювальні процедури з меншою обчислювальною складністю, як часткові випадки (*subdivision*-процедури).

Введемо позначення

$$\gamma_{(x)}^{(r,k)}$$

– вектор стовпець координати якого після операції дискретної згортки з послідовністю відліків $\{p_i\}_{i \in \mathbb{Z}}$, функції $p(t)$ надають лінійний функціонал $S_{r,k}^{(x)}$, що є частковим випадком сплайнів (1)-(4) в точці x , тобто:

$$S_{r,k}^{(x)} = \sum_{i \in \mathbb{Z}} \gamma_{(x)}^{(r,k)} \cdot p_i.$$

Нехай для визначеності x може набувати значення із множини

$$\left\{ -1, -\frac{4}{5}, -\frac{3}{4}, -\frac{2}{3}, -\frac{1}{2}, -\frac{2}{5}, -\frac{1}{3}, -\frac{1}{4}, -\frac{1}{5}, 0, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{4}, \frac{4}{5}, 1 \right\}. \quad (4)$$

Тоді для сплайну $S_{2,1}(p,t)$ має місце:

$$S_{2,1}^{(x)} = \sum_{j=i-2}^{i+2} \gamma_{(x)}^{(2,1)} \cdot p_j,$$

де

$$\begin{aligned} \gamma_{(-1)}^{(2,1)} &= \frac{1}{12} (-1 \ 7 \ 7 \ -1 \ 0)^T; \\ \gamma_{(-\frac{4}{5})}^{(2,1)} &= \frac{1}{1200} (-81 \ 530 \ 862 \ -110 \ -1)^T; \\ \gamma_{(-\frac{3}{4})}^{(2,1)} &= \frac{1}{768} (-49 \ 314 \ 574 \ -70 \ -1)^T; \\ \gamma_{(-\frac{2}{3})}^{(2,1)} &= \frac{1}{432} (-25 \ 154 \ 342 \ -38 \ -1)^T; \\ \gamma_{(-\frac{1}{2})}^{(2,1)} &= \frac{1}{300} (-16 \ 95 \ 247 \ -25 \ -1)^T; \\ \gamma_{(-\frac{1}{5})}^{(2,1)} &= \frac{1}{192} (-9 \ 50 \ 166 \ -14 \ -1)^T; \\ \gamma_{(-\frac{2}{5})}^{(2,1)} &= \frac{1}{1200} (-49 \ 250 \ 1078 \ -70 \ -9)^T; \\ \gamma_{(-\frac{1}{3})}^{(2,1)} &= \frac{1}{108} (-4 \ 19 \ 99 \ -5 \ -1)^T; \\ \gamma_{(-\frac{1}{4})}^{(2,1)} &= \frac{1}{768} (-25 \ 106 \ 718 \ -22 \ -9)^T; \\ \gamma_{(-\frac{1}{5})}^{(2,1)} &= \frac{1}{300} (-9 \ 35 \ 283 \ -5 \ -4)^T; \\ \gamma_{(0)}^{(2,1)} &= \frac{1}{48} (-1 \ 2 \ 46 \ 2 \ -1)^T; \end{aligned}$$

$$\begin{aligned}
 \gamma_{(\frac{1}{5})}^{(2,1)} &= \frac{1}{300}(-4 \quad -5 \quad 283 \quad 35 \quad -9)^T; \\
 \gamma_{(\frac{1}{4})}^{(2,1)} &= \frac{1}{768}(-9 \quad -22 \quad 718 \quad 106 \quad -25)^T; \\
 \gamma_{(\frac{1}{3})}^{(2,1)} &= \frac{1}{108}(-1 \quad -5 \quad 99 \quad 19 \quad -4)^T; \\
 \gamma_{(\frac{2}{5})}^{(2,1)} &= \frac{1}{1200}(-9 \quad -70 \quad 1078 \quad 250 \quad -49)^T; \\
 \gamma_{(\frac{1}{2})}^{(2,1)} &= \frac{1}{192}(-1 \quad -14 \quad 166 \quad 50 \quad -9)^T; \\
 \gamma_{(\frac{2}{3})}^{(2,1)} &= \frac{1}{300}(-1 \quad 25 \quad 247 \quad 95 \quad -16)^T; \\
 \gamma_{(\frac{2}{5})}^{(2,1)} &= \frac{1}{432}(-1 \quad -38 \quad 342 \quad 154 \quad -25)^T; \\
 \gamma_{(\frac{3}{4})}^{(2,1)} &= \frac{1}{768}(-1 \quad -70 \quad 574 \quad 314 \quad -49)^T; \\
 \gamma_{(\frac{4}{5})}^{(2,1)} &= \frac{1}{1200}(-1 \quad -110 \quad 862 \quad 530 \quad -81)^T; \\
 \gamma_{(1)}^{(2,1)} &= \frac{1}{12}(0 \quad -1 \quad 7 \quad 7 \quad -1)^T.
 \end{aligned}$$

Для сплайну $S_{2,2}(p, t)$ виконється:

$$S_{2,2}^{(x)} = \sum_{j=i-3}^{i+3} \gamma_{(x)}^{(2,2)} \cdot p_j,$$

де

$$\begin{aligned}
 \gamma_{(-1)}^{(2,2)} &= \frac{1}{72}(1 \quad -9 \quad 44 \quad 44 \quad -9 \quad 1 \quad 0)^T; \\
 \gamma_{(-\frac{4}{5})}^{(2,2)} &= \frac{1}{7200}(81 \quad -692 \quad 3195 \quad 5552 \quad -1045 \quad 108 \quad 1)^T; \\
 \gamma_{(-\frac{3}{4})}^{(2,2)} &= \frac{1}{4608}(49 \quad -412 \quad 1867 \quad 3712 \quad -677 \quad 68 \quad 1)^T; \\
 \gamma_{(-\frac{2}{3})}^{(2,2)} &= \frac{1}{2592}(25 \quad -204 \quad 891 \quad 2224 \quad -381 \quad 36 \quad 1)^T; \\
 \gamma_{(-\frac{1}{5})}^{(2,2)} &= \frac{1}{1800}(16 \quad -127 \quad 535 \quad 1612 \quad -260 \quad 23 \quad 1)^T; \\
 \gamma_{(-\frac{1}{2})}^{(2,2)} &= \frac{1}{1152}(9 \quad -68 \quad 267 \quad 1088 \quad -157 \quad 12 \quad 1)^T; \\
 \gamma_{(-\frac{2}{5})}^{(2,2)} &= \frac{1}{7200}(49 \quad -348 \quad 1235 \quad 7088 \quad -885 \quad 52 \quad 9)^T; \\
 \gamma_{(-\frac{1}{3})}^{(2,2)} &= \frac{1}{648}(4 \quad -27 \quad 87 \quad 652 \quad -72 \quad 3 \quad 1)^T; \\
 \gamma_{(-\frac{1}{4})}^{(2,2)} &= \frac{1}{4608}(25 \quad -156 \quad 419 \quad 4736 \quad -429 \quad 4 \quad 9)^T; \\
 \gamma_{(-\frac{1}{5})}^{(2,2)} &= \frac{1}{1800}(9 \quad -53 \quad 120 \quad 1868 \quad -145 \quad -3 \quad 4)^T; \\
 \gamma_{(0)}^{(2,2)} &= \frac{1}{288}(1 \quad -4 \quad -5 \quad 304 \quad -5 \quad -4 \quad 1)^T; \\
 \gamma_{(\frac{1}{5})}^{(2,2)} &= \frac{1}{1800}(4 \quad -3 \quad -145 \quad 1868 \quad 120 \quad -53 \quad 9)^T; \\
 \gamma_{(\frac{1}{4})}^{(2,2)} &= \frac{1}{4608}(9 \quad 4 \quad -429 \quad 4736 \quad 419 \quad -156 \quad 25)^T; \\
 \gamma_{(\frac{1}{3})}^{(2,2)} &= \frac{1}{648}(1 \quad 3 \quad -72 \quad 652 \quad 87 \quad -27 \quad 4)^T;
 \end{aligned}$$

$$\begin{aligned}\gamma_{(-\frac{2}{5})}^{(2,2)} &= \frac{1}{7200} (9 \ 52 \ -885 \ 7088 \ 1235 \ -348 \ 49)^T; \\ \gamma_{(\frac{1}{2})}^{(2,2)} &= \frac{1}{1152} (1 \ 12 \ -157 \ 1088 \ 267 \ -68 \ 9)^T; \\ \gamma_{(\frac{2}{3})}^{(2,2)} &= \frac{1}{1800} (1 \ 23 \ -260 \ 1612 \ 535 \ -127 \ 16)^T; \\ \gamma_{(\frac{2}{5})}^{(2,2)} &= \frac{1}{2592} (1 \ 36 \ -381 \ 2224 \ 891 \ -204 \ 25)^T; \\ \gamma_{(\frac{3}{4})}^{(2,2)} &= \frac{1}{4608} (1 \ 68 \ -677 \ 3712 \ 1867 \ -412 \ 49)^T; \\ \gamma_{(\frac{4}{5})}^{(2,2)} &= \frac{1}{7200} (1 \ 108 \ -1045 \ 5552 \ 3195 \ -692 \ 81)^T; \\ \gamma_{(1)}^{(2,2)} &= \frac{1}{72} (0 \ 1 \ -9 \ 44 \ 44 \ 9 \ 1)^T.\end{aligned}$$

Для сплайну $S_{3,1}(p, t)$ буде:

$$S_{3,1}^{(x)} = \sum_{j=i-2}^{i+3} \gamma_{(x)}^{(3,1)} \cdot p_j,$$

де

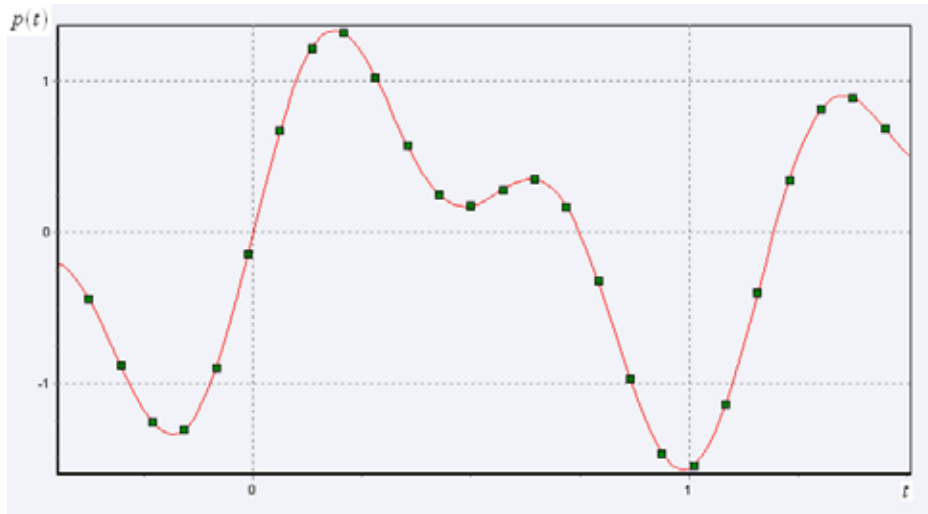
$$\begin{aligned}\gamma_{(-1)}^{(3,1)} &= \frac{1}{144} (-5 \ 14 \ 126 \ 14 \ -5 \ 0)^T; \\ \gamma_{(-\frac{4}{5})}^{(3,1)} &= \frac{1}{144000} (-3645 \ 5071 \ 123782 \ 25398 \ -6601 \ -5)^T; \\ \gamma_{(-\frac{3}{4})}^{(3,1)} &= \frac{1}{73728} (-1715 \ 1647 \ 62762 \ 14630 \ -3591 \ -5)^T; \\ \gamma_{(-\frac{2}{3})}^{(3,1)} &= \frac{1}{31104} (-625 \ 95 \ 25934 \ 7366 \ -1661 \ -5)^T; \\ \gamma_{(-\frac{2}{5})}^{(3,1)} &= \frac{1}{18000} (-320 \ -189 \ 14702 \ 4838 \ -1026 \ -5)^T; \\ \gamma_{(-\frac{1}{2})}^{(3,1)} &= \frac{1}{9216} (-135 \ -257 \ 7250 \ 2934 \ -571 \ -5)^T; \\ \gamma_{(-\frac{2}{5})}^{(3,1)} &= \frac{1}{144000} (-1715 \ -6043 \ 108234 \ 53186 \ -9527 \ -135)^T; \\ \gamma_{(-\frac{1}{3})}^{(3,1)} &= \frac{1}{3888} (-40 \ -193 \ 2822 \ 1570 \ -266 \ -5)^T; \\ \gamma_{(-\frac{1}{4})}^{(3,1)} &= \frac{1}{73728} (-625 \ -4235 \ 50958 \ 32962 \ -5197 \ -135)^T; \\ \gamma_{(-\frac{1}{5})}^{(3,1)} &= \frac{1}{18000} (-135 \ -1102 \ 12046 \ 8514 \ -1283 \ -40)^T; \\ \gamma_{(0)}^{(3,1)} &= \frac{1}{1152} (-5 \ -81 \ 662 \ 662 \ -81 \ -5)^T; \\ \gamma_{(\frac{1}{5})}^{(3,1)} &= \frac{1}{18000} (-40 \ -1283 \ 8514 \ 12046 \ -1102 \ -135)^T; \\ \gamma_{(\frac{1}{4})}^{(3,1)} &= \frac{1}{73728} (-135 \ -5197 \ 32962 \ 50958 \ -4235 \ -625)^T; \\ \gamma_{(\frac{1}{3})}^{(3,1)} &= \frac{1}{3888} (-5 \ -266 \ 1570 \ 2822 \ -193 \ -40)^T; \\ \gamma_{(\frac{2}{5})}^{(3,1)} &= \frac{1}{144000} (-135 \ -9527 \ 53186 \ 108234 \ -6043 \ -1715)^T; \\ \gamma_{(\frac{1}{2})}^{(3,1)} &= \frac{1}{9216} (-5 \ -571 \ 2934 \ 7250 \ -257 \ -135)^T; \\ \gamma_{(\frac{2}{5})}^{(3,1)} &= \frac{1}{18000} (-5 \ -1026 \ 4838 \ 14702 \ -189 \ -320)^T;\end{aligned}$$

$$\gamma_{(-\frac{2}{3})}^{(3,1)} = \frac{1}{31104}(-5 \quad -1661 \quad 7366 \quad 25934 \quad 95 \quad -625)^T;$$

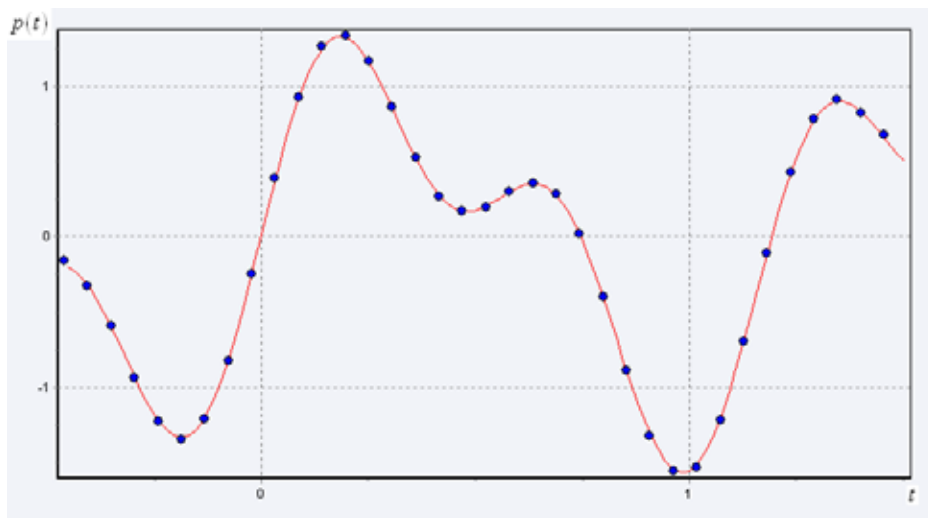
$$\gamma_{(\frac{3}{4})}^{(3,1)} = \frac{1}{73728}(-5 \quad -3591 \quad 14630 \quad 62762 \quad 1647 \quad -1715)^T;$$

$$\gamma_{(\frac{5}{8})}^{(3,1)} = \frac{1}{144000}(-5 \quad -6601 \quad 25398 \quad 123782 \quad 5071 \quad -3645)^T;$$

$$\gamma_{(1)}^{(3,1)} = \frac{1}{144}(0 \quad -5 \quad 14 \quad 126 \quad 14 \quad -5)^T.$$



а)



б)

Рис. 1. Небінарний subdivision «3 в 4» на основі сплайну (1):

а) функція $p(t) = \sin(5t) + 0,6\sin(11t)$, $t \in [-0,45; 1,52]$ та її початкові відліки;

б) результат subdivision після однієї ітерації.

На графіках (рис. 1) подано приклад небінарного subdivision при проектуванні кожних 2-х послідовних відліків функції $p(t) = \sin(5t) + 0,6\sin(11t)$, $t \in [-0,45; 1,52]$ (суцільна крива) в 3-и відліки. В якості лінійних операторів масштабування брались часткові випадки сплайну (1) відповідно в точках $x = -\frac{1}{4}$, $x = -\frac{3}{4}$, $x = \frac{3}{4}$, $x = \frac{3}{4}$ з множини (5), котрим, відповідають маски $\gamma_{(-\frac{1}{4})}^{(2,1)}$, $\gamma_{(-\frac{3}{4})}^{(2,1)}$, $\gamma_{(\frac{3}{4})}^{(2,1)}$,

$\gamma_{\frac{1}{4}}^{(2,1)}$. Як видно з графіків, при низькій обчислювальній складності, що надають часткові випадки сплайну (1), одержано наближення прийнятно високої якості апроксимації.

Висновки. В роботі отримано понад п'ятдесят часткових випадків локальних поліноміальних сплайнів другого порядку першого та другого ступеня уточнення та третього порядку першого ступеня уточнення, що є близькими до інтерполяційних у середньому. Нові лінійні оператори можуть мати використання при розробці програмного забезпечення опрацювання числових послідовностей, цифрових сигналів та зображень, що функціонує у режимі реального часу. Швидкодія розрахунків забезпечується низькою обчислювальною складністю запропонованих лінійних операторів.

Подальші дослідження можуть полягати в отриманні нових операторів обробки дво- та тривимірних послідовностей на основі наведених в роботі функціоналів.

Список використаних джерел:

1. Роджерс Д., Адамс Дж. Математические основы машинной графики. М. : Мир, 2001. 604 с.
2. Holschneider M. Wavelets. An analysis Tool. Oxford : Oxford University Press, 1995.
3. Andersson L.-E., Stewart N. Introduction to the mathematics of subdivision surfaces. Philadelphia : Society for Industrial and Applied Mathematics, 2010. 356 p.
4. Де Бор К. Практическое руководство по сплайнам. М. : Радио и связь, 1985. 303 с.
5. Чуи Ч. Введение в вэйвлеты. М. : Мир, 2001. 412 с.
6. Unser M. Splines: A Perfect Fit for Signal and Image Processing. *IEEE Signal Processing Magazine*. 1999. Vol. 16. № 6. P. 22–38.
7. Лигун А.А., Шумейко А.А. Асимптотические методы восстановления кривых. К. : ІМ НАН України, 1996. 358 с.
8. Приставка П.О. Поліномаїльні сплайни при обробці даних. Д. : Вид-во Дніпропетр. ун-ту, 2004. 236 с.
9. Приставка П.О. Лінійні комбінації В-сплайнів, близькі до інтерполяційних у середньому, в задачі моделювання аналогових сигналів. *Актуальні проблеми автоматизації та інформаційних технологій* : зб. наук. праць. Д. : Вид-во Дніпропетр. ун-ту. 2011. Т. 15. С. 4–17.
10. Лигун А.А., Кармазина В.В. О восстановлении эмпирической функции плотности распределения с помощью гистосплайнов второго порядка. Днепропетровск : Днепропетровский индустр. ин-т, 1989. 30 с.
11. Приставка П.О. Поліноміальні сплайни в задачах бінарного поповнення. *Актуальні проблеми автоматизації та інформаційних технологій*. 2003. Т. 7. С. 39–53.
12. Приставка П.О. Небінарне поповнення послідовностей відліків гладких функцій лінійними операторами на основі поліноміальних сплайнів. *Вісн. НАУ*. 2008. № 3. С. 85–89.

References:

1. Rogers, D., Adams, J. (2001). *Matematicheskie osnovy mashinnoy grafiki [Mathematical foundations of computer graphics]*. Moscow: Mir [in Russian].
2. Holschneider, M. Wavelets (1995) *An analysis Tool*. Oxford: Oxford University Press.
3. Andersson, L.-E., Stewart, N. (2010). *Introduction to the mathematics of subdivision surfaces*. Philadelphia: Society for Industrial and Applied Mathematics.
4. De Bohr, K. (1985). *Prakticheskoe rukovodstvo po splajnam [A Practical Guide to Splines]*. Moscow: Radio and communication [in Russian].
5. Chui, Ch. (2001). *Vvedenie v vehvlety [An Introduction to Wavelets]*. Moscow: Mir [in Russian].
6. Unser, M. (1999). Splines: A Perfect Fit for Signal and Image Processing. *IEEE Signal Processing Magazine*, vol. 16, no 6, pp. 22–38.
7. Ligon, A.A., Shumeiko, A.A. (1996). *Asimptoticheskie metody vosstanovleniya krivyykh [Asymptotic methods for restoring curves]*. Kyiv: IM of the NAS of Ukraine [in Russian].
8. Prystavka, P.O. (2004). *Polinomialni splainy pry obrobsi danykh [Polynomial splines in data processing]*. Dnipro: Dnipropetrovsk University Publishing House [in Ukrainian].
9. Prystavka, P.O. (2011). Liniini kombinatsii V-splainiv, blyzki do interpolatsiinykh u serednomu, v zadachi modeliuvannia analohovykh syhnaliv [Linear combinations of B-splines, close to interpolation on average, in the problem of modeling analog signals]. *Aktualni problemy avtomatyzatsii ta informatsiinykh tekhnolohii – Actual problems of automation and information technology*, vol. 15, pp. 4–17 [in Ukrainian].
10. Ligon, A.A., Karmazina, V.V. (1989). O vosstanovlenii ehmpiricheskoy funktsii plotnosti-raspredeleniya s pomoshchyu gistosplajnov vtorogo poryadka [Reconstruction of the empirical distribution density function using second-order histosplines]. Dneprodzerzhinsk: Dneprodzerzhinsk Industrial Institute in Russian].
11. Prystavka, P.O. (2003). Polinomialni splainy v zadachakh binarnoho popovnnennia [Polynomial splines in binary improvement problems]. *Aktualni problemy avtomatyzatsii ta informatsiinykh tekhnolohii – Actual problems of automation and information technology*, vol. 7, pp. 39–53 [in Ukrainian].
12. Prystavka, P.O. (2008). Nebinarne popovnnennia poslidovnostei vidlikiv hladkykh funktsii liniinymy operatoramy na osnovi polinomialnykh splainiv [Non-binary reimplementations of poslidities in smooth functions by line operators based on polynomial splines]. *Visnyk NAU – Proceedings of the National Aviation University*, no 3, pp. 85–89 [in Ukrainian].

УДК 004.932

DOI <https://doi.org/10.32689/maup.it.2021.2.4>

Микола РУДНІЧЕНКО

кандидат технічних наук, доцент кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (nickolay.rud@gmail.com)

ORCID: <https://orcid.org/0000-0002-7343-8076>

Сергій ГРИШИН

кандидат технічних наук, доцент кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (grishin_si@ukr.net)

ORCID: <https://orcid.org/0000-0002-1933-9552>

Денис ШИБАЄВ

аспірант кафедри технічної кібернетики та інформаційних технологій, Одеський національний морський університет, вул. Мечникова 34, Одеса, Україна, індекс 65029 (denscreamer@gmail.com)

ORCID: <https://orcid.org/0000-0002-3260-5843>

Ігор ПЕТРОВ

доктор технічних наук, професор кафедри морських перевезень, Національний університет «Одеська морська академія», вул. Дідріхсона 8, Одеса, Україна, індекс 65029 (firmness@list.ru)

ORCID: <https://orcid.org/0000-0002-8740-6198>

Максим НОСОВ

студент, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом», вул. Фрометівська 2, Київ, Україна, індекс 03039 (maxnosovgenie@gmail.com)

ORCID: <https://orcid.org/0000-0213-8322-1311>

Mykola RUDNICHENKO

PhD in Technical Sciences, Associate Professor of Information Technology, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (nickolay.rud@gmail.com)

Sergey GRISHIN

PhD in Technical Sciences, Associate Professor of Information Technology, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (grishin_si@ukr.net)

Denis SHIBAYEV

Graduate student of the Department of Technical Cybernetics and Information Technologies, Odessa National Maritime University, 34 Mechnikova Street, Odessa, Ukraine, postal code 65029 (denscreamer@gmail.com)

Igor PETROV

Doctor of Technical Sciences, Professor of the Department of Maritime Transport, National University "Odessa Maritime Academy", 8 Didrichson Street, Odessa, Ukraine, postal code 65029 (firmness@list.ru)

Maxim NOSOV

Student, Interregional Academy of Personnel Management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039 (maxnosovgenie@gmail.com)

Бібліографічний опис статті: Рудніченко М., Гришин С., Шibaєв Д., Петров І., Носов М. Проект інформаційної системи для збору, обробки та аналізу метеоданих. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 34–41. DOI: <https://doi.org/10.32689/maup.it.2021.2.4>

Bibliographic description of the article: Rudnichenko, M., Hryshyn, S., Shybaiev, D., Petrov, I., Nosov, M. (2021). Proekt informatsiinoi systemy dlia zboru, obrobky ta analizu meteo danykh [Project of information system for collection, processing and data analysis]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 34–41. DOI: <https://doi.org/10.32689/maup.it.2021.2.4>

ПРОЕКТ ІНФОРМАЦІЙНОЇ СИСТЕМИ ДЛЯ ЗБОРУ, ОБРОБКИ ТА АНАЛІЗУ МЕТЕОДАНИХ

Анотація. Завдання точного та оперативного прогнозу погодних умов у різних географічних регіонах держав набуває все більшої актуальності та затребуваності у різних прикладних сферах людської діяльності, зокрема у аграрному, господарському, будівельному, виробничому, торговому та туристичному секторах сучасного бізнесу. **Метою** статті є розробка проекту інформаційної системи для збору, обробки та аналізу метеоданих великого обсягу. Реалізація поставленої мети передбачає вирішення низки **завдань**: 1) визначення та характеристика специфіки побудови метеорологічних прогнозів та обробки статистичних даних; 2) розробка концепції роботи системи та її архітектури; 3) формалізація та проведення тестовою обробки даних для виявлення сильних сторін у створеній інформаційній системі. **Наукова новизна.** У статті розглядається можливість використання низки алгоритмів машинного навчання, зокрема моделей штучних нейронних мереж різних архітектурних типів, для автоматизації процесу вирішення завдання регресії та при роботі з часовими рядами даних. Як **висновок**, у статті наголошується, що розроблена інформаційна система для збору, обробки та аналізу метеоданих великого обсягу реалізує повний цикл функціональних процесів з побудови прогнозів на погодний стан на декілька днів наперед. Суттєвою перевагою пропонованої системи є її веб-орієнтованість, що дозволяє зробити кінцевий програмний продукт кросплатформним, масштабованим і розподіленим програмним рішенням для використання різними типами користувачів для виконання дослідних чи прикладних завдань з обробки метеоданих. Подальшим шляхом розвитку системи може бути додавання підтримки нових типів вихідних даних про досліджувані об'єкти за різними показниками та інтеграція електронних засобів та платформ для безпосереднього отримання даних з джерел (датчиків та сенсорів).

Ключові слова: інформаційні системи, аналіз даних, обробка даних, збір даних, метеоданих.

PROJECT OF INFORMATION SYSTEM FOR COLLECTION, PROCESSING AND DATA ANALYSIS

Abstract. The task of accurate and efficient weather forecasting in different geographical regions of the state is becoming increasingly important and in demand in various applied areas of human activity, including the agricultural, economic, construction, manufacturing, trade and tourism sectors of modern business. **The aim** of the article is to develop a project of an information system for the collection, processing and analysis of large-scale meteorological data. Realization of the set purpose provides the decision of a number of problems: 1) definition and the characteristic of specificity of construction of meteorological forecasts and processing of statistical data; 2) development of the concept of system operation and its architecture; 3) formalization and testing of data processing to identify strengths in the created information system. **Scientific novelty.** The article considers the possibility of using a number of machine learning algorithms, in particular models of artificial neural networks of different architectural types, to automate the process of solving the regression problem and when working with time series. In **conclusion**, the article emphasizes that the developed information system for the collection, processing and analysis of large-scale meteorological data implements a full cycle of functional processes for building weather forecasts for several days in advance. A significant advantage of the proposed system is its web orientation, which allows to make the final software product cross-platform, scalable and distributed software solutions for use by different types of users to perform research or application tasks for meteorological data processing. A further way to develop the system may be to add support for new types of source data on the studied objects on various indicators and the integration of electronic tools and platforms for direct data acquisition from sources (sensors and sensors).

Key words: information systems, data analysis, data processing, data collection, meteorological data.

Актуальність проблеми. Інформаційні технології активно використовуються в метеорології і роботі синоптиків, через те, що дозволяють не тільки автоматизувати їх роботу, а також підвищити зручність її виконання і значно підвищити точність результату [1].

Сучасна метеорологія не змогла б розвинутися без впровадження в неї інформаційних технологій і особливо суперкомп'ютерів, це пов'язано з тим, що робота синоптиків, зокрема побудова прогнозу погоди, вимагає великих обчислювальних і людських ресурсів. Починаючи від простого побудови прогнозу погоди закінчуючи попередженням про насування погодних катастроф, які могли б привести до великих людських жертв і масштабних руйнувань [2].

Використання обчислювальної техніки і відповідних програм дозволяє вирішувати завдання метеорології з великою точністю, здійснювати розрахунки для великих територій за більш короткий термін. У зв'язку з цим дана галузь вкрай сильно потребує розробки зручних і високопродуктивних програмних засобів. Розвинена метеорологія дозволяє поліпшити важливі аспекти життя людей: збільшує ефективність сільськогосподарської галузі, дозволяє визначати безпечні для життя людей місця, прокласти більш оптимальні транспортні маршрути, мінімізувати наслідки природних катастроф [3].

Аналіз останніх досліджень і публікацій. Сучасних метеорологів можна розділити по виду діяльності на наступні категорії або наукові напрямки [4]:

- Фізична метеорологія, вивчає розробку радіолокаційних і космічних методів дослідження атмосферних явищ.
- Динамічна метеорологія, вивчає фізичні механізми атмосферних процесів.
- Синоптична метеорологія, наука про закономірності зміни погоди.
- Кліматологія, вивчає клімат як сукупність погодних характеристик за багаторічний період і їх зміну.
- Аерологія, вивчає верхні шари атмосфери для декількох десятків кілометрів від поверхні землі.

Для виконання аналітичної роботи необхідно взаємодіяти злагоджено і задіяти безліч фахівців по всьому світу, які кожні три години збирають інформацію про погоду зі всіх доступних джерел, далі проводять їх обробку і будують кліматичні карти, описують графічним методом поточний погодний стан атмосфери в різних регіонах [5].

Далі за отриманими картами проводиться аналіз, що виявляє необхідну інформацію, наприклад аналіз погодних аномалій для розуміння їх потенціалу до перетворення у природні катастрофи.

Прикладами метеорологічних даних, що збираються для досліджень, є: місце та час коли були зняті метеодані, базова інформація яка потрібна для прив'язки метеоданих на карті, висота хмар від рівня моря, дальність видимості, вид хмар у секторі, напрям та швидкість вітру, температура повітря в час збору даних, точка роси, атмосферний тиск на рівні метеостанції, атмосферний тиск на рівні моря, максимальна та мінімальна температура за добу, погодне явище та кількість опадів [6].

На базі цих метеоданих будуються прогнози погоди та більшості з усіх метеорологічних розрахунків.

Основою для прогнозу погоди є облік періодичних і неперіодичних змін метеорологічних величин і явищ погоди [7].

Періодичні зміни тієї чи іншої метеорологічної величини обумовлюються добовим і річним ходом цієї величини, неперіодичні – еволюцією і переміщенням синоптичних об'єктів: циклонів і антициклонів, повітряних мас і атмосферних фронтів [8].

Синоптичний метод в даний час є основним при розробці довгострокового прогнозу погоди. Його суть полягає у тому, що на підставі аналізу карт погоди за кілька послідовних термінів складається прогноз синоптичного положення, який полягає в прогнозуванні виникнення, переміщення і еволюції повітряних мас, атмосферних фронтів та баричних систем.

Прогностичну карту складають на певний момент часу наступної доби. На підставі даної карти формується прогноз погоди, який є логічним продовженням прогнозу синоптичного положення і виходить з основного принципу: переміщення і еволюція синоптичних об'єктів здійснюють певні зміни та властиві умови погоди [9].

Тому за прогностичні значення метеорологічної величини в першому наближенні приймаються їх значення від району, звідки очікується переміщення синоптичного об'єкту, до району, для якого складається прогноз погоди. При такому великому обсязі роботи та даних необхідних до зберігання дана галузь не може обійтися без використання інформаційних систем (ІС) та технологій.

З кожним днем необхідних метеорологічних розрахунків стає все більше, через те, що все більше галузей виявилися залежні від інформації про стан атмосфери або про її вплив, що обумовлює необхідність в аналізі метеоданих [10]. Для ефективного виконання завдань метеорології сучасні програмні застосування повинні мати наступні властивості: високу швидкодію, гнучкість та зручність використання, достатній рівень точності розрахунків, простий процес налаштування та інтеграції.

Метою статті є розробка проекту інформаційної системи для збору, обробки та аналізу метеоданих.

Виклад основного матеріалу. Початковим етапом при розробці ІС є проектування, метою даного процесу є визначення головних функціональних можливостей програмного забезпечення, його архітектуру, складові, а також компоненти. В більшості випадків проект ІС складається з розробки головних діаграм, які в графічному вигляді дозволяють визначити усі складові. В якості засобу проектування використовується мова UML, яка є спеціалізованою мовою моделювання, яку використовують розробники для графічного представлення програмного забезпечення. Згідно проекту ІС буде мати архітектуру типового веб-застосування, що дозволяє користувачам отримувати усі функціональні можливості без інсталяції. Розроблена діаграма варіантів використання ІС наведена на рис. 1. Основними варіантами використання є:

1. Перегляд знайдених аномалій, користувачу на формі ІС необхідно вибрати пункт «перегляд аномалій». Після переходу на цю сторінку користувач бачить таблицю з усіма знайденими аномаліями в метеоданих, доступними є функції фільтрації, пошуку та сортування.

2. Перегляд зібраних метеоданих, користувачу на формі ІС необхідно вибрати пункт «метеодані».

3. Перегляд розрахованих прогнозів.

4. Перегляд даних в таблицях.

5. Пошук. Дозволяє відшукувати данні в таблицях за будь-яким ключем. Ключем може бути будь-який текст, цифри, символи.

6. Оновлення метеоданих. Після оновлення ІС активує алгоритм пошуку аномалій в даних, далі на вилучених та «чистих» вибірках система робить прогнозування метеоданих на наступні 3 доби, що використовується у якості бази прогнозування погодних явищ. Усі данні які оброблюються на даному етапі записуються в базу даних (БД).

7. Оновлення прогнозу та запуск алгоритму пошуку аномалій.

8. Можливість проведення тестування ефективності навчання штучних нейронних мереж (ШНМ) на сірих та чистих метеоданих. Процес тестування починається з того, що модель для пошуку аномалій оброблює данні, формує два списки, на яких ШНМ проходять навчання для прогнозування метеоданих та погодних явищ, після цього ці моделі проходять тестування на сирих та чистих даних для оцінки їх точності.

ІС можна розділити на два головних компоненти: інтерфейс та контролер для роботи з алгоритмами обробки моделей штучних нейронних мереж. Діаграма класів компонента інтерфейсу та контролеру, зображено на рис. 2. На діаграмі зображено функціональну залежність компонентів. У зв'язку з тим, що це web-орієнтована ІС, то простір класів розроблюється з точки зору окремих компонентів. На цій діаграмі можна розглянути усі задіяні класи. Класи *DataReader*, *DB* та *DataWriter* необхідні для роботи з БД, зокрема для запису та зчитування даних. Кожен з цих класів є обгорткою над інтегрованими в класах та працює з даними через спеціальний контейнер для більш зручної роботи з ними. Клас *DataReader* тісно пов'язаний з інтерфейсом сторінок, він витягує потрібні данні з БД та передає їх до головного класу контролера *DataManager*. Клас *DataReader* реалізує функціонал фільтрації, пошуку, сортування та зміни сторінок. Клас *DataWriter* реалізує функціонал запису даних в базу. Він бере данні з двох джерел, з компонента обробки та аналізу метеоданих та компонента парсингу даних з зовнішніх джерел. Клас *DB* реалізує базову логіку роботи з БД.

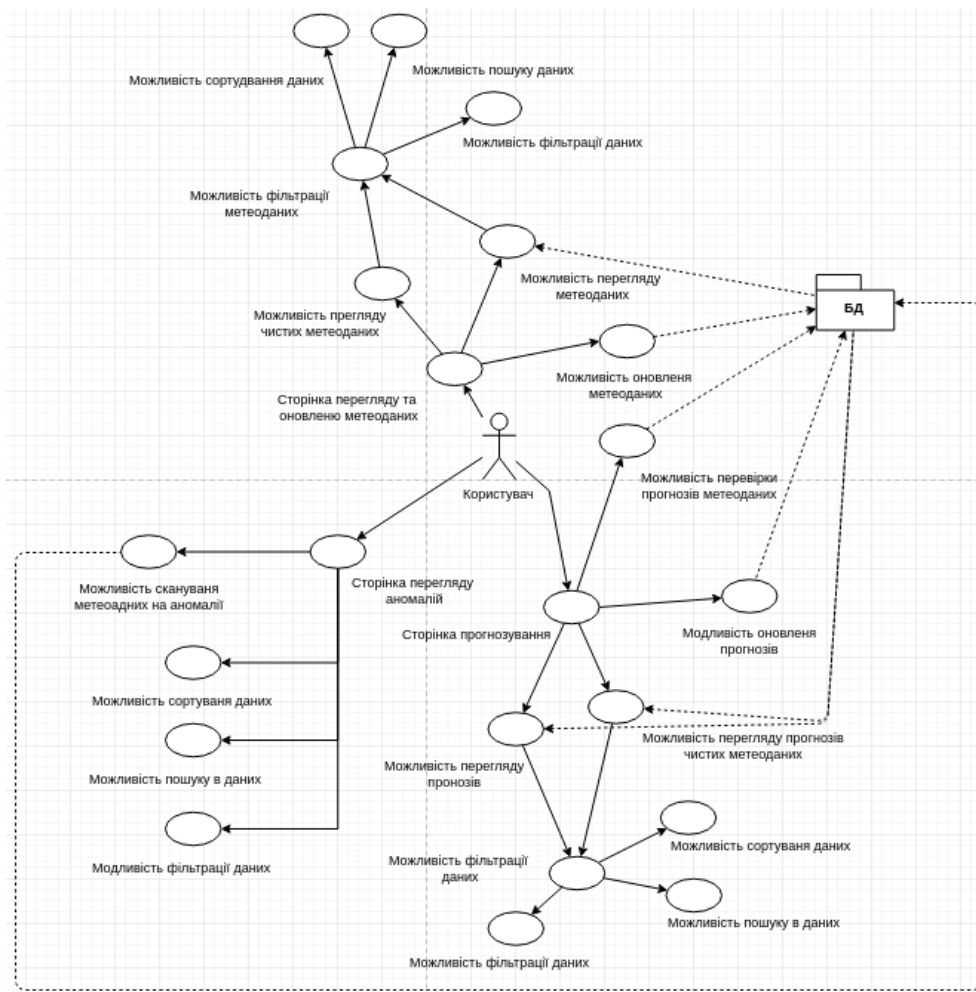


Рис. 1. Діаграма варіантів використання ІС

Підкомпонент для парсингу даних з зовнішніх джерел складається з двох класів DataMiner і Parser, останній реалізує функціонал збору даних з зовнішніх джерел, використовуючи HTTPS запити та бібліотеку BS4 для парсингу. Клас DataMiner використовується для об'єднання усіх вилучених даних в єдиний список для подальшого завантаження в БД.

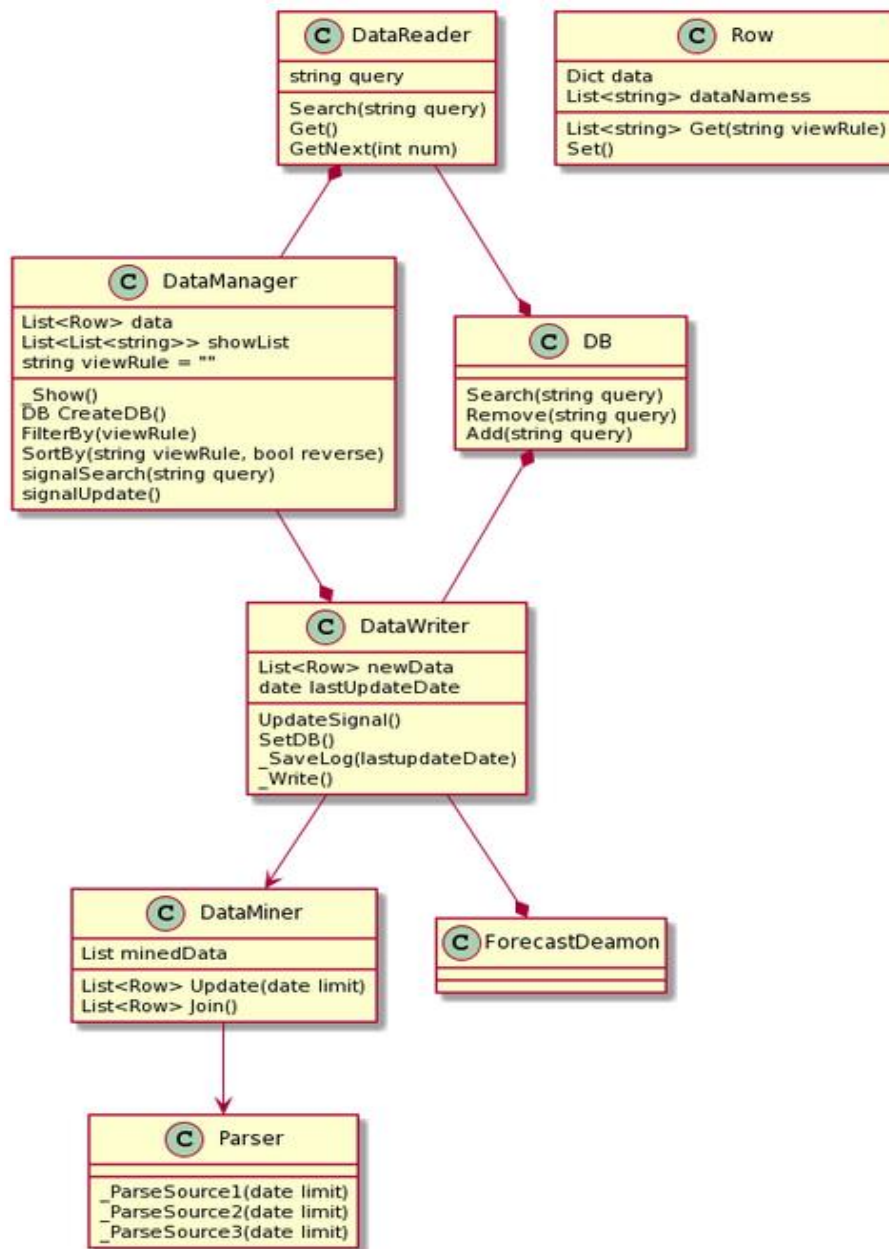


Рис. 2. Діаграма класів IC

Блок-схема алгоритму обробки отриманих даних і побудови прогнозу на три доби наведена на рис. 3, кожен запис прогнозу містить інформацію про погоду і прогнозоване за цими даними погодні явище, різниця між записами дорівнює 3 годинам. Кроки по блок-схемі наступні:

- читання даних з обох списків;
- підготовка моделі ШНМ під ці дані;
- складання прогнозу метеоданих;
- виконання прогнозу погодного явища моделлю ШНМ;
- аналіз точність роботи всіх моделей та їх оцінка.

Базова сторінка є сторінкою перегляду метеоданих, вона складається з декількох елементів. Перший елемент, необхідний для роботи з системою, це верхня панель переходу між сторінками.

На цій панелі розташовані три кнопки для переходу між сторінками, перегляду метеоданих, перегляду прогнозів і перегляду аномалій шляхом натискання відповідних кнопок.

Кнопка «Оновити дані» запускає основний алгоритм обробки даних, це основна кнопка по роботі з системою.

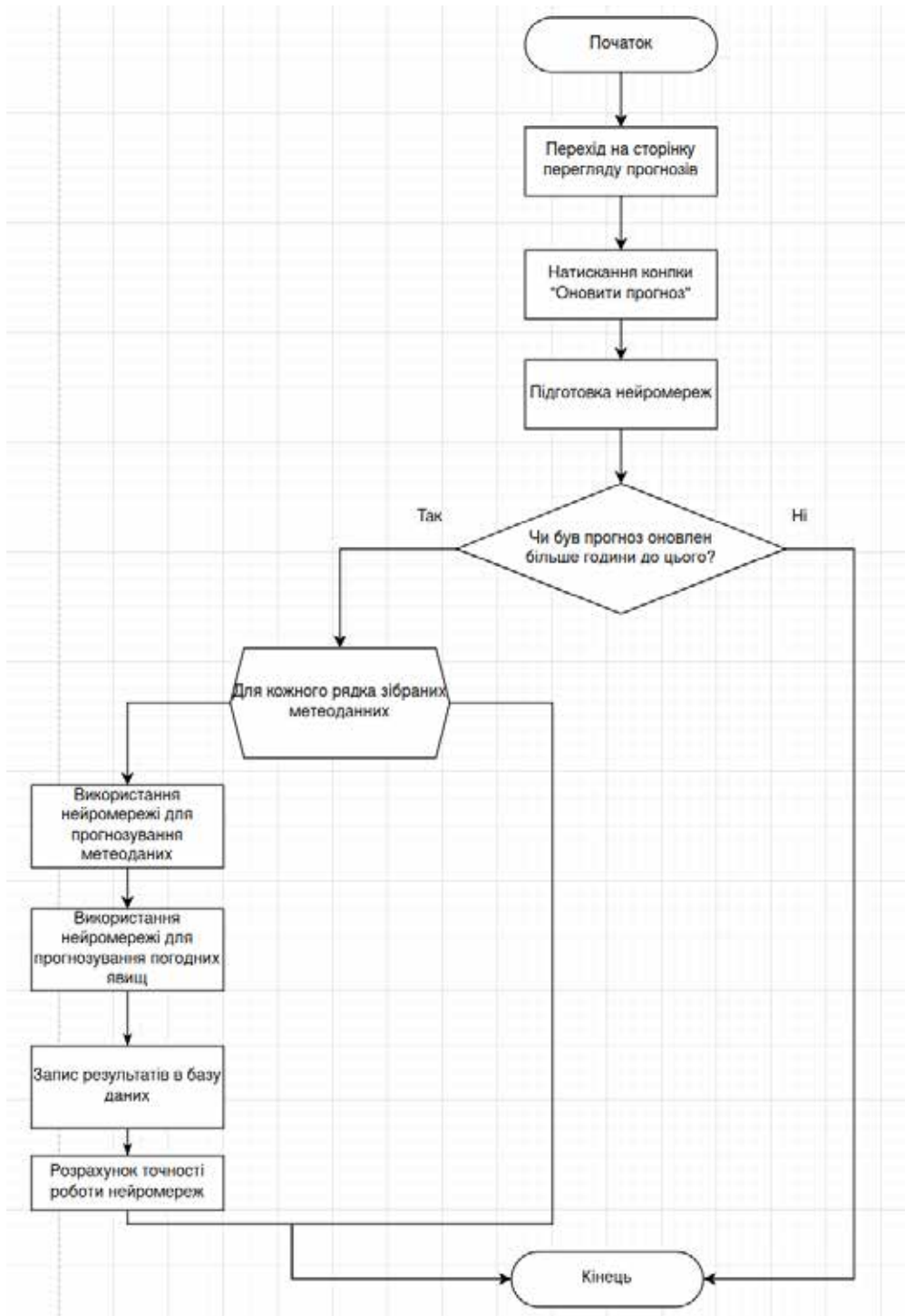


Рис. 3. Блок-схема алгоритму обробки отриманих даних і побудова прогнозу
Структура головної таблиці БД системи наведена на рис. 4.

id	datetime	place	placeName	wind_way	wind_speed	air_pressure	water_pressure	weather	temperature
id	timestamp with time zone	place	storage: string (30)	layer	double precision	double precision	double precision	text	double precision
1	4060709	2017-11-01 02:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
2	4060710	2017-11-01 05:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
3	4060711	2017-11-01 08:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
4	4060712	2017-11-01 11:00:00+02	31267	Руса Рускан	2	0	760	760 (Морось (чм))	3.8521391333916076
5	4060713	2017-11-01 14:00:00+02	31267	Руса Рускан	2	0	760	760 (Колониста...)	3.8521391333916076
6	4060714	2017-11-01 17:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
7	4060715	2017-11-01 20:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
8	4060716	2017-11-01 23:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
9	4060717	2017-11-02 02:00:00+02	31267	Руса Рускан	2	0	760	760 (Ланчови...)	3.8521391333916076
10	4060718	2017-11-02 05:00:00+02	31267	Руса Рускан	2	0	760	760 (Ланчови...)	3.8521391333916076
11	4060719	2017-11-02 08:00:00+02	31267	Руса Рускан	2	0	760	760 (Колониста...)	3.8521391333916076
12	4060720	2017-11-02 11:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
13	4060721	2017-11-02 14:00:00+02	31267	Руса Рускан	2	0	760	760 (чм)	3.8521391333916076
14	4060722	2017-11-02 17:00:00+02	31267	Руса Рускан	2	0	760	760 (Ланчови...)	3.8521391333916076
15	4060723	2017-11-02 20:00:00+02	31267	Руса Рускан	2	0	760	760 (Ланчови...)	3.8521391333916076
16	4060724	2017-11-02 23:00:00+02	31267	Руса Рускан	2	0	760	760 (Ланчови...)	3.8521391333916076
17	4060725	2017-11-03 02:00:00+02	31267	Руса Рускан	2	0	760	760 (Ланчови...)	3.8521391333916076

Рис. 4. Структура таблиці БД по зібраних метеоданих

Форма додавання нового запису метеоданих включає в себе наступні поля:

- Id, поле цілочисельного типу, зберігає унікальний номер запису;
- Datetime, зберігає дату та час виконання запису;
- Place, зберігає назву регіону в якому був зроблений запис;
- placeNumber, зберігає номер регіону, в якому був зроблений запис;
- temperature, зберігає значення температури, що була зафіксована в момент запису;
- wind_way, зберігає напрямок вітру;
- wind_speed, зберігає значення швидкості вітру;
- air_pressure, зберігає значення атмосферного тиску;
- water_pressure, зберігає значення атмосферного тиску на рівні моря.

Концепція роботи системи є наступною. Першим етапом ІС завантажує інформацію про погоду, в разі якщо дані частково були завантажені раніше то система оновить їх, вибравши тільки ті, що з'явилися з дати останнього запису.

Далі дані передаються на обробку моделі ШНМ з пошуку аномалій, яка записує їх в два списки, «чистий» (в якому знаходяться всі записи, які пройшли перевірку) і «сирий» (в нього входять всі отримані дані про погоду), а усі знайдені аномалії заносяться в окрему таблицю в БД.

Наступним етапом є підготовка моделі ШНМ для прогнозування, після чого обидві пари проходять свої списки і будують прогноз метеоданих і погодних явищ. В процесі обробки даних система збирає статистику алгоритмів прогнозування для подальшого порівняння ефективності моделей.

По завершенні цього процесу відбувається запис розрахованих прогнозів до БД, прогноз пари, яка робила розрахунок на чистих даних, записується в список чистого прогнозу.

Прогноз другої пари записується у повний список прогнозів. Після завершення процесів обробки даних користувач переходить на відповідну сторінку, де зображена стовпчаста діаграма з коефіцієнтами оцінки точності алгоритмів, виводиться кількість отриманих записів і число знайдених в них аномалій, а також загальні витрати часу на процеси обчислень.

Висновки та перспективи подальших досліджень. Розроблена інформаційна система для метеорологічних центрів дозволяє робити автоматичний короткостроковий прогноз погоди та проводити аналіз метеоданих завдяки інструментам створення графіків. Головна перевага даної системи в тому, що вона дозволяє задіяти та оцінити різні алгоритми прогнозування і пошуку аномалій, зокрема різних моделей ШНМ. Подальшим шляхом розвитку системи є підвищення якості прогнозування та збільшення підтримуваних обсягів даних для обробки.

Список використаних джерел:

1. Бондар О.І. Вступ до кліматології і метеорології: екологічні аспекти. Херсон : Олді-плюс, 2011. 220 с.
2. Братков В.В., Воронин А.П. Метеорологія і кліматологія. М. : МИИГАиК, 2015. 209 с.
3. Гринченко Н.Н., Потапова В.Ю., Тарасов А.С. Алгоритми прогнозування погодних умов в системах збору і обробки метеорологічних даних. *Известия Тульского государственного университета. Технические науки*. 2018. № 2. С. 113–119.
4. Костров Б.В., Гринченко Н.Н., Потапова В.Ю., Тарасов А.С. Разработка сетевой версии системы сбора метеоданных для прогнозирования ночных заморозков. *Известия Тульского государственного университета. Технические науки*. 2017. № 2. С. 89–95.
5. Глаголев В.А. Разработка информационной системы оценки и прогнозирования пожарной опасности растительной территории по метеорологическим условиям. *Современные научные исследования и инновации*. 2014. № 4. С. 56–67.

6. Алексеев Г.В. Метеорологические и геофизические исследования. М. : Paulsen, 2011. 352 с.
7. Арсенюк С.Ю. Метеорология и климатология. Алчевск : ДонГТУ, 2005. 162 с.
8. Барашкова Н.К., Кижнер Л.И., Кузевская И.В. Атмосферные процессы: динамика, численный анализ, моделирование. Томск : Томский государственный университет, 2012. 312 с.
9. Богаткин О.Г., Тараканов Г.Г., Основы метеорологии. СПб. : РГГМУ, 2006. 232 с.
10. Боголюбова Е.В., Сулейменова Г.Т. Основы динамической метеорологии. Алматы : Қазақ университеті, 2017. 174 с.

References:

1. Bondar, O.I. (2011). *Vstup do klimatologii i meteorologii: ekologichni aspekti [Introduction to climatology and meteorology: environmental aspects]*. Herson: Oldi-pljus [In Ukrainian].
2. Bratkov, V.V., Voronin, A.P. (2015). *Meteorologija i klimatologija [Meteorology and climatology]*. M: MIIGAIK [In Russian].
3. Grinchenko, N.N., Potapova, V.Ju., Tarasov, A.S. (2018). Algoritmy prognozirovanija pogodnyh uslovij v sistemah sbora i obrabotki meteorologicheskikh dannyh [Algorithms for forecasting weather conditions in systems for collecting and processing meteorological data]. *Izvestija Tul'skogo gosudarstvennogo universiteta. Tehnicheskie nauki – Proceedings of Tula State University. Technical sciences*, no. 2, pp. 113–119 [In Russian].
4. Kostrov, B.V., Grinchenko, N.N., Potapova, V.Ju., Tarasov, A.S. (2017). Razrabotka setевой versii sistemy sbora meteoannyh dlja prognozirovanija nochnyh zamorozkov [Development of a network version of the meteorological data collection system for forecasting night frosts]. *Izvestija Tul'skogo gosudarstvennogo universiteta. Tehnicheskie nauki – Proceedings of Tula State University. Technical sciences*, no. 2, pp. 89–95 [In Russian].
5. Glagolev, V.A. (2014). Razrabotka informacionnoj sistemy ocenki i prognozirovanija pozharnej opasnosti rastitel'noj territorii po meteorologicheskim uslovijam [Development of an information system for assessing and forecasting the fire hazard of the vegetation area according to meteorological conditions]. *Sovremennye nauchnye issledovanija i innovacii – Modern scientific researches and innovations*, no. 4, pp. 56–67 [In Russian].
6. Alekseev, G.V. (2011). *Meteorologicheskie i geofizicheskie issledovanija [Meteorological and geophysical research]*. M.:Paulsen [In Russian].
7. Arsenjuk, S.Ju. (2005). *Meteorologija i klimatologija [Meteorology and climatology]*. Alchevsk: DonGTU [In Russian].
8. Barashkova, N.K., Kizhner, L.I., Kuzhevskaja, I.V. (2012). *Atmosfernye processy: dinamika, chislennyj analiz, modelirovanie [Atmospheric processes: dynamics, numerical analysis, modeling]*. Tomsk: Tomskij gosudarstvennyj universitet [In Russian].
9. Bogatkin, O.G., Tarakanov, G.G. (2006). *Osnovy meteorologii [Meteorology and climatology]*. SPb.: RGGMU [In Russian].
10. Bogoljubova, E.V., Sulejmenova, G.T. (2017). *Osnovy dinamicheskoy meteorologii [Auleymenova GT Fundamentals of dynamic meteorology]*. Almaty: Kazakh University [In Russian].

УДК 004

DOI <https://doi.org/10.32689/maup.it.2021.2.5>

Микола РУДНІЧЕНКО

кандидат технічних наук, доцент кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (nickolay.rud@gmail.com)

ORCID: <https://orcid.org/0000-0002-7343-8076>

Сергій ГРИШИН

кандидат технічних наук, доцент кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (grishin_si@ukr.net)

ORCID: <https://orcid.org/0000-0002-1933-9552>

Наталя ШИБАЄВА

кандидат технічних наук, доцент кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (nati.shibaeva@gmail.com)

ORCID: <https://orcid.org/0000-0002-7869-9953>

Володимир ВИЧУЖАНІН

доктор технічних наук, професор, завідувач кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (vint532@yandex.ua)

ORCID: <https://orcid.org/0000-0002-5244-5808>

В'ячеслав РИЧКО

студент, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом», вул. Фрометівська 2, Київ, Україна, індекс 03039 (ruchkoliveviacheslav@gmail.com)

ORCID: <https://orcid.org/0000-1167-1312-9081>

Mykola RUDNICHENKO

PhD in Technical Sciences, Associate Professor of Information Technology, Odessa Polytechnic State University, 1 ave. Shevchenko, 1, Odessa, Ukraine, postal code 65001 (nickolay.rud@gmail.com)

Sergey GRISHIN

PhD in Technical Sciences, Associate Professor of Information Technology, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (grishin_si@ukr.net)

Natalia SHIBAYEVA

Candidate of Technical Sciences, Associate Professor of Information Technology, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (nati.shibaeva@gmail.com)

Vladimir VYCHUZHANIN

Doctor of Technical Sciences, Professor, Head of the Department of Information Technologies, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (vint532@yandex.ua)

Vyacheslav RYCHKO

Student, Interregional Academy of Personnel Management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039 (ruchkoliveviacheslav@gmail.com)

Бібліографічний опис статті: Рудніченко М., Гришин С., Шибаяєва Н., Вичужанін В., Ричко В. Аналіз особливостей та засобів розробки інтерактивних систем з підтримкою технологій доповненої реальності для сфери освіти. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 42–49. DOI: <https://doi.org/10.32689/maup.it.2021.2.5>

Bibliographic description of the article: Rudnichenko, M., Hryshyn, S., Shybaieva, N., Vychuzhanin, V., Rychko, V. (2021). Analiz osoblyvostei ta zasobiv rozrobky interaktyvnykh system z pidtrymkoiu tekhnolohii dopovnenoi realnosti dlia sfery osvity [Analysis of features and development potential of interactive systems supported by auxiliary reality technologies for education field]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 42–49. DOI: <https://doi.org/10.32689/maup.it.2021.2.5>

АНАЛІЗ ОСОБЛИВОСТЕЙ ТА ЗАСОБІВ РОЗРОБКИ ІНТЕРАКТИВНИХ СИСТЕМ З ПІДТРИМКОЮ ТЕХНОЛОГІЙ ДОПОВНЕНОЇ РЕАЛЬНОСТІ ДЛЯ СФЕРИ ОСВІТИ

Анотація. В даний час на ринку сучасних програмних систем та інформаційних технологій особливим чином розвивається напрям віртуальної реальності як одного з зручних підходів до поліпшення процесу інтерактивного навчання та візуалізації складних систем, пристроїв та окремих елементів, важливою галуззю якого є технології доповненої реальності. **Метою** статті є проведення аналізу особливостей та засобів розробки інтерактивних систем з підтримкою технологій доповненої реальності для сфери освіти для формування концепції подальшої розробки повнофункціональної інформаційної системи підтримки вивчення студентами навчальних дисциплін. Реалізація поставленої мети передбачає вирішення низки **завдань**: 1) аналіз особливостей застосування мультимедійних технологій у сфері науки та освіти; 2) аналіз технологій та засобів розробки систем з елементами віртуальної та доповненої реальності; 3) розробка концепції проекту системи підтримки процесів освіти. **Наукова новизна.** У статті наведено підхід з використання технологій доповненої реальності на мобільних пристроях, задіяючи можливості модуля камери під час її активізації, для набуття процесу вивчення теоретичних положень з учбових дисциплін більшого рівня інтерактивності засобами середовища Unity. Як **висновок**, у статті наголошується, що технології віртуальної та додаткової реальності дозволяють значним чином спростити вирішення складних завдань з наочної та динамічної візуалізації окремих досліджуваних об'єктів чи систем. Гнучкість та універсальність системи Unity дозволяє реалізувати повноцінних клієнт-серверних програмних застосувань, що мають тонкий клієнт у вигляді кросплатформеного мобільного застосування та серверної частини, що несе відповідальність по процесах обробки даних для забезпечення працездатності усієї системи.

Ключові слова: інтерактивні системи, доповнена реальність, освітні процеси.

ANALYSIS OF FEATURES AND DEVELOPMENT POTENTIAL OF INTERACTIVE SYSTEMS SUPPORTED BY AUXILIARY REALITY TECHNOLOGIES FOR EDUCATION FIELD

Abstract. Currently, the market of modern software systems and information technology is developing a special direction of virtual reality as one of the convenient approaches to improve the process of interactive learning and visualization of complex systems, devices and individual elements, an important area of which is augmented reality technology. **The aim** of the article is to analyze the features and means of developing interactive systems with the support of augmented reality technologies for education to form the concept of further development of a full-featured information system to support the study of academic disciplines. Realization of the set purpose provides the decision of a number of **tasks**: 1) the analysis of features of application of multimedia technologies in the field of science and education; 2) analysis of technologies and tools for developing systems with elements of virtual and augmented reality; 3) development of the concept of the project of system of support of processes of education. **Scientific novelty.** The article considers an approach to the use of augmented reality technologies on mobile devices, using the capabilities of the camera module during its activation, to acquire the process of studying theoretical positions in academic disciplines of a higher level of interactivity by means of Unity. In **conclusion**, the article emphasizes that virtual and augmented reality technologies can greatly simplify the solution of complex problems of visual and dynamic visualization of individual objects or systems. The flexibility and versatility of the Unity system allows you to implement full-fledged client-server software applications that have a thin client in the form of a cross-platform mobile application and a server part that is responsible for data processing to ensure the efficiency of the entire system.

Key words: interactive systems, augmented reality, educational processes.

Актуальність проблеми. Інформаційна революція і формування інформаційного суспільства висунули інформацію і знання на передній план соціального та економічного розвитку. В даний час в сучасному світі повсюдно спостерігається активна інформатизація як виробничих, так і навчальних процесів [1]. Це обумовлено необхідністю забезпечення віддаленого режиму роботи фахівців і учнів, а також постійним розвитком і вдосконаленням ринку інформаційних технологій (ІТ) і систем, чільне місце в якому займають мобільні застосування [2].

З метою забезпечення максимального ступеня гнучкості у виборі необхідного користувачам функціоналу з'явилася і отримала активний розвиток модель освітніх послуг і технологій, яка використовується низкою великих компаній у сфері ІТ. Тенденції розвитку сучасної системи освіти нерозривно пов'язані з широким впровадженням в навчальний процес різних форм, методів і засобів активного навчання [3].

Однією з провідних тенденцій інформатизації суспільства є розвиток мультимедійних технологій, їх проникнення в різні сфери соціального життя: виробництво, бізнес, науку, освіту, масову споживачку

культуру [4]. Забезпечуючи багатство змісту і форми, поєднання різних видів текстової, графічної, мовної, музичної, відео-, фото- інформації і різноманітність способів їх вилучення, ці технології формують мультимедійне сприйняття світу [5].

Актуальність досліджуваної тематики обумовлена активним і масовим використанням технологій доповненої реальності (AR) в практиці освіти, як нового витка в розвитку мультимедійних технологій, що відкриває нові можливості в організації навчального процесу, а також розвитку творчих здібностей учнів. Використання технологій доповненої реальності відкриває нові можливості в організації навчального процесу, а також у розвитку творчих здібностей учнів. Такий вид навчання сприяє підвищенню залученості, комунікацій між тими, хто навчається і інтересу до предмета [6].

Аналіз останніх досліджень і публікацій. Існуюча система отримання професійної освіти в її класичному вигляді, сформована раніше сьогодні не є ефективною для повноцінного вирішення проблеми підготовки майбутніх кадрів і їх виховання. У процесі переходу багатьох початкових спеціальних освітніх установ в режим організаційно-правової автономної форми одним з ключових факторів є активний розвиток цільової спрямованості впроваджуваних освітніх програм. За допомогою даного аспекту проводиться підвищення обсягу вибору учнів різних рівнів і типів освіти, що сприяє формуванню вироблення акцентів на вимоги сучасного ринку праці. Ключовим при цьому є розробка і впровадження підходів до складання змісту і засобів навчання учнів з урахуванням використання сучасної науки [3].

Слід зазначити, що одним з найбільш ефективних напрямків підвищення якості проведення освітнього процесу є інтеграція мультимедійних засобів і технологій. Поряд з цим, невід'ємною частиною процесу навчання стають комп'ютерні інформаційні технології [5]. Сучасні мультимедійні технології (MT) підтримки процесів утворення активно використовуються в якості основи для формування нового напрямку навчання, тому що MT засновані на використанні інтерактивних і обчислювальних можливостей комп'ютерів, із задіянням мов і технологій програмування. MT навчання концептуально можуть бути виражені у вигляді набору технологій підтримки процесів утворення, що сприяють ефективно виконувати проектування і розробки змісту, форм і методів навчання для підвищення якості забезпечення навчального процесу, які передбачають використання програмних технологій інформаційних засобів [2]. Мультимедіа як явище та інструмент освіти являє собою комплекс пов'язаних і взаємодіючих апаратних і програмних засобів і технологій, який здійснює підтримку користувачів в їх роботі з різнорідними даними великих обсягів, які організовані у вигляді цілісної інформаційного середовища. Також, мультимедіа як явище, означає якийсь спектр інформаційних технологій для найбільш ефективного впливу на кінцевого користувача [4; 5].

Метою статті є проведення аналізу особливостей та засобів розробки інтерактивних систем з підтримкою технологій доповненої реальності для сфери освіти для формування концепції подальшої розробки повнофункціональної інформаційної системи підтримки вивчення студентами навчальних дисциплін.

Виклад основного матеріалу. У зв'язку з тим, що стрімкий розвиток ІТ знаходить своє відображення в освітньому процесі, все більшої актуальності в даній сфері набувають MT у вигляді віртуальної реальності (VR), хоча вони стали застосовуватися відносно недавно.

В даний час VR технології застосовні в різних освітніх процесах, в тому числі [6]:

1. У класичному середню або вищу освіту за рахунок можливостей доповнення навчальних занять VR-симуляцією на кілька хвилин, що вносить більше розуміння процесів, що розглядаються.
2. При дистанційному навчанні VR дозволяє об'єднати учнів і викладачів в єдиному віртуальному просторі, віртуальної класній кімнаті, в режимі реального часу забезпечуючи інтерактивний освітній процес.
3. Для самостійного освіти VR може використовуватися для освоєння мов, проходження екскурсій або відвідування віртуальних музеїв, використовуючи спеціальні окуляри і встановивши мобільні додатки на смартфони.

Популярним напрямом VR є технології розширеної реальності (AR), що відрізняються від них тим, що додають реальному світу деякі інтерактивні шари або елементи тривимірного простору. Таким чином, користувачі систем з підтримкою AR можуть як і раніше взаємодіяти з фізичним середовищем, отримуючи при це додаткову інформацію від своїх пристроїв або додатків. Використання AR технологій в навчанні, зокрема, дозволяє: підвищити інтерактивність навчальних матеріалів для учнів, що робить їх більш зрозумілими і цікавими; забезпечити більш глибоке занурення учнів в освітній процес за допомогою використання 3D-візуалізації і елементів гейміфікації [7].

Застосування технологій AR сприяє більш якісному наданню учням практичного досвіду, знижує вплив різних відволікаючих чинників, які ускладнюють процес сприйняття інформації і спрощує складні явища і процеси для їх усвідомлення [8].

В даний час на ринку рішень в області AR існує кілька платформ і програмних рішень, що підтримують технології та засоби розробки додатків, розглянемо далі найбільш актуальні з них.

1. Unity 3D – являє собою комплексну середу для створення різних комп'ютерних ігор та інтерактивних програм для різних цільових платформ і операційних систем, імплементує набір різних програмних засобів і технологій, що використовуються при розробці програм, в тому числі текстовий редактор коду, компілятор, модуль налагодження з підтримкою засобів AR / VR.

Завдяки модульності і гнучкості використання, Unity спрощує виконання рутинних операцій при розробці додатків, в тому числі завдяки багатому числу платформ, що підтримує движок системи.

Серед Unity мінімізує процес написання коду, роблячи акцент на налаштування інтерфейсу і об'єктів додатки в інтерактивному вигляді, шляхом редагування сцени і її параметризації [9]. Компонентний склад системи на базі Unity наведено на рис. 1.

В системі використовується компонентно-орієнтований підхід, коли розробник починає процес створення програми з окремих об'єктів, виставляючи потрібні параметри відображення і описуючи порядок взаємодії між ними, в міру необхідності доповнюючи сцену різними компонентами (відображення або управління об'єктами).

Завдяки підтримці функцій Drag & Drop в інтерфейсі графічного редактора середовища ставати можливою зображення карти координат і расстановки об'єктів на сцені з подальшим тестуванням роботи програми [10].

Суттєва перевага Unity 3D полягає в підтримці і постійному оновленні бібліотеки Ассет і плагінів, що розширюють і прискорюють процес розробки програми.

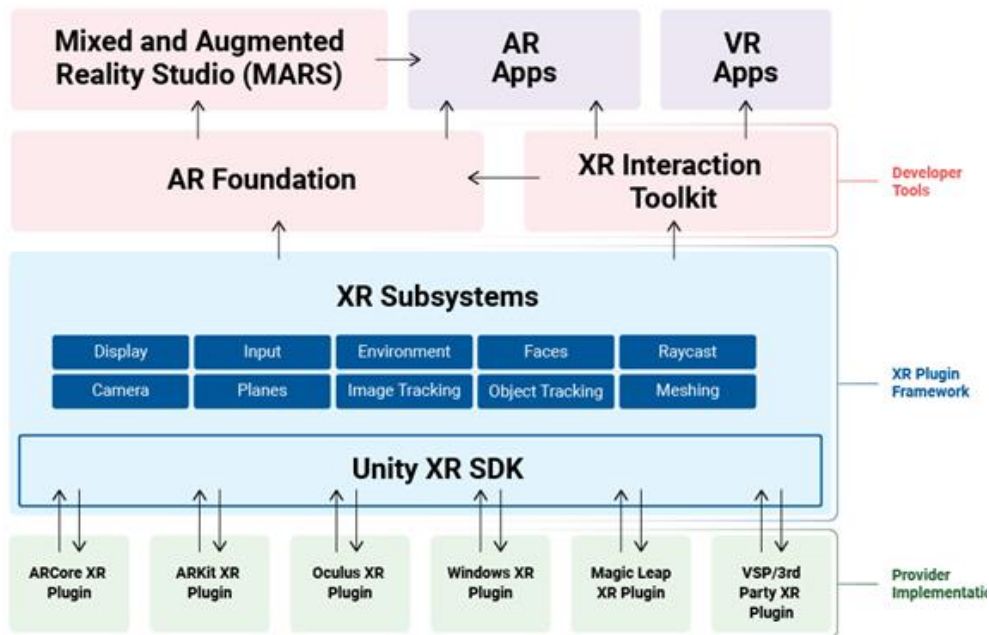


Рис. 1. Компонентний склад системи на базі Unity

Складністю Unity є високе навантаження на обчислювальні можливості комп'ютера, через що на не надто сучасному апаратному забезпеченні робота по рендерингу і розробці додатків сповільнюється. Розробка масштабних, насичених і складних сцен може негативно позначитися на загальній продуктивності додатка, в зв'язку з чим необхідно проведення процедури оптимізації його роботи, що не завжди ефективно і рентабельно.

Ще одним недоліком є досить високий обсяг займаного простору додатків на базі Unity.

2. Vuforia. Однією з найбільш популярних технологій підтримки AR для Unity є Vuforia, що представляє собою платформу для підтримки процесів створення AR додатків для портативних пристроїв (смартфонів і планшетів на різних операційних системах).

Дана платформа дозволяє застосовувати технічно ефективні і швидкодіючі технології розпізнавання графічних зображень на основі імплементації алгоритмів і методів комп'ютерного зору в фізичних пристроях для з'єднання реального світу з цифровою віртуальним середовищем. Фактично, весь цей функціонал знаходиться в наборі для (SDK), який надає розробникам інструменти для створення

AR додатків. Зокрема, Vuforia SDK надає наступні можливості: зворотні виклики для подій; високо-рівневий доступ до апаратних блокам; кілька об'єктів відстеження: зображення, множинні мішені, циліндричні мішені, маркери кадру і об'єктна мішень; взаємодії з реальним світом за допомогою, наприклад, віртуальних кнопок [11].

Ключові компоненти AR в Vuforia це: дані камери пристрою; дані про місцезнаходження (GPS); сенсорні дані, отримані через датчики пристрої; засоби накладення тривимірної комп'ютерної графіки на активну сцену [8].

Компонентний склад Vuforia SDK наведено на рис. 2.

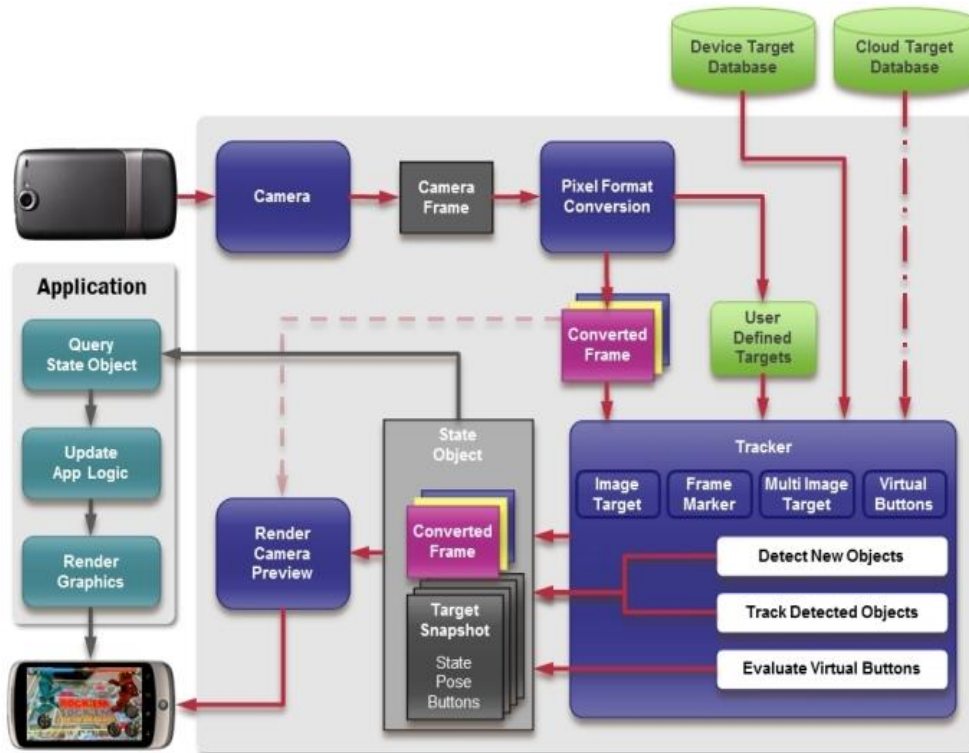


Рис. 2. Компонентний склад Vuforia SDK

3. Unreal Engine, система є затребуваною для створення контенту в VR і AR, працює на базі використання мови C++, а так само середовища розробки Blueprint.

Підтримувані Unreal Engine платформи AR (Android, iOS), VR (Google VR, Oculus, Gear VR, SteamVR, Windows MR) [9].

Серед безлічі поліпшень, передбачених системою, доданий фреймворком доповненої реальності (Augmented Reality Framework), реалізує підтримку додатків для iOS і Android, використовуючи єдиний код, з підтримкою функції Light Estimation, Pinning, Alignment, Trace Results, Session State і Tracking. Функція візуалізації контролерів руху компонентів дозволяє команді оперативно додавати модель візуалізації Static Mesh. За замовчуванням модель буде відповідати Oculus Touch, з підтримкою завантаження призначених для користувача 3D-проектів. Для спрощення процесу розгортання Unreal Engine використовує систему залежних компонентів, підтримує системи рендеринга (Direct3D, OpenGL, Pixomatic), відтворення звуку, засоби відтворення голосового тексту, модулі для роботи з комп'ютерною мережею. Додатково, користувачі Unreal Engine можуть з'єднувати Maya і переглядати зроблені зміни в редакторі UE4 в режимі реального часу.

4. ARCore є комплексний набір засобів (SDK) для розробки прикладних програм для сімейства ОС Android під смартфони і планшети. Розробники використовують функціонал ARCore при розробці модулів і додатків, що володіють інтерактивністю, з використанням можливостей AR.

Приклад архітектури системи на базі ARCore наведено на рис. 3.

ARCore ідентифікує горизонтальні поверхні на зображенні і стежить за рухом камери пристрою. Система базується на таких технологіях: розпізнавання елементів навколишнього світу (пошук горизонтальних поверхонь і розрахунок їх фізичних розмірів); детекція руху об'єктів (фіксація віртуальних

об'єктів в рамках захопленої поверхні при русі смартфона); оцінка ступеня освітленості сцени (забезпечення реакції віртуальних об'єктів на інтенсивність освітлення шляхом внесення будь-яких шумів або змін його візуального стану) [8].

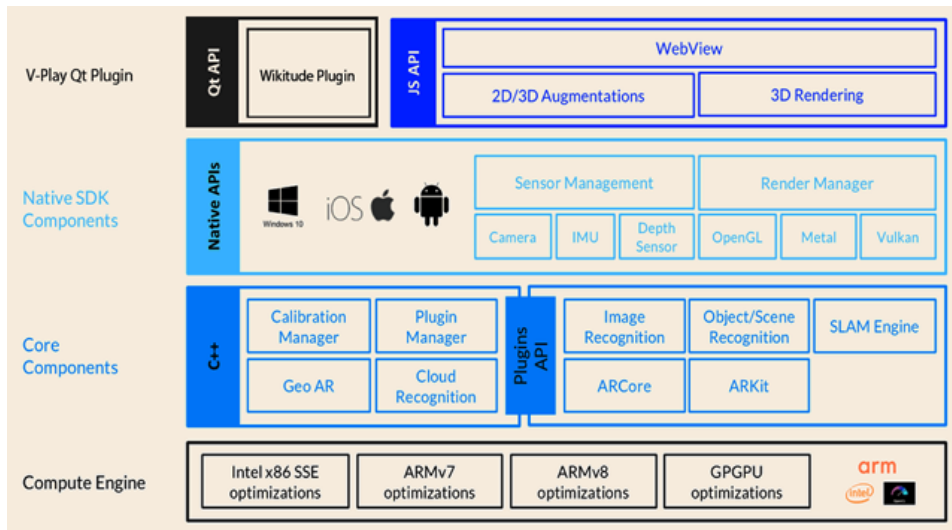


Рис. 3. Типова архітектура системи на базі ARCore

ARCore виконує пошук кластерів характерних точок на зображенні, які розташовані на горизонтальних поверхнях, на базі чого вони можуть використовуватися додатком при розмітці площин.

В основі всієї логіки ARCore розташована технологія SLAM (локалізація і розмітка), імплементація якої призводить до формування карти у вигляді графіка 3D точок. Вони являють собою певним чином розріджений хмара, в якому кожна окрема відмітка відповідає конкретним координатам віртуального оптичного об'єкта [9].

Розглянувши наведені засоби доцільним є обрання системи Unity, як найбільш технічно розвинутої та універсальної для завдань розробки інтерактивних систем з підтримкою розширеної реальності.

Типовою пропонованою концепцією проекту системи для підтримки процесів освіти може бути мобільне застосування (засноване на системі Unity), мета якого полягає у забезпечення можливостей інтерактивного вивчення ряду положень навчальних дисциплін для підвищення ефективності засвоєння матеріалів здобувачами вищої чи середньої освіти.

Завдання мобільного застосування:

- організація персоніфікованого доступу за логіном і паролем;
- реалізація підтримки процесу вивчення інформації в рамках класних занять з начальних дисциплін шляхом внесення в систему елементів розширеної реальності;
- забезпечення можливостей оцінки рівня теоретичної підготовки з навчальних дисциплін шляхом інтеграції тестів контролю знань.

Цільова аудиторія системи може складатися з учнів середніх і випускних класів шкіл і початкових курсів коледжів, а також для всіх бажаючих ознайомитися з теоретичними аспектами начальних дисциплін.

Ключові функціональні вимоги:

- реєстрація для доступу в систему;
- авторизація в системі за логіном і паролем (в захищеному режимі відображення);
- вибір режиму роботи з додатком (теорія або практика);
- вибір розділу теоретичних положень;
- вибір підрозділу активного розділу для вивчення конкретних навчальних матеріалів;
- проходження тестів з обраної тематики;
- оцінка результатів тестування;
- активація режиму AR через інтеграцію API камери мобільного пристрою;
- захоплення зображення з камери мобільного пристрою;
- розпізнавання маркера (зображення прив'язки об'єкта) на зображенні з потоку камери мобільного пристрою;

- завантаження і активація AR об'єкта в режимі активної камери мобільного пристрою;
- динамічне відображення (анімація) AR об'єкта в режимі активної камери мобільного пристрою.

Головні не функціональні вимоги:

- кросплатформеність (робота під різними операційними платформами, Android, iOS);
- обсяг підсумкового файлу білда (на прикладі Android – apk пакета) проекту не більше 100 Мб;
- обсяг займаної оперативної пам'яті не більше 400 Мб;
- рендеринг моделей AR на екрані активної камери по всій області захоплення;
- коректне відображення (масштабування) елементів інтерфейсу і робота програми на діагоналях екранів мобільних пристроїв від 5 дюймів і вище з дозволом від 800 на 600 пікселів і вище.

Висновки та перспективи подальших досліджень. В результаті проведених досліджень встановлено, що сучасні мобільні технології і системи розширеної реальності мають значний технічний потенціал, який може мати прикладне використання в рамках освітнього процесу в навчальних закладах, дозволяючи спростити процес імплементації як очного, так і дистанційного режиму навчання, що є вкрай актуальним завданням в умовах існуючої світової пандемії.

На базі виконаного аналізу існуючих технологій розробки додатків розширеної реальності для різних прикладних сфер можна зробити висновок про фактичне домінування на ринку системи Unity, яка імплементує велику кількість бібліотек, залежностей і артефактів, які розширюють можливості розробки зі створення систем з елементами доповненої реальності. Для подальшого створення концепції проекту інтерактивної системи можуть бути обрані наступні засоби розробки: для написання програмного коду – мова C #, для створення проектів класів і редагування коду – середовище Visual Studio, для створення структури і інтерфейсу додатку, а також підключення технологій доповненої реальності системи Unity.

Перспективним напрямком подальших досліджень є аналіз шляхом розробки корпоративних мобільних додатків і систем, що включають в свій склад найбільш цільові функції з підтримки навчального процесу для створення єдиного уніфікованого мобільного сервісу для інтеграції комплексу рішень в сфері освіти на державному рівні.

Список використаних джерел:

1. Гранин К. Образование в контексте глобализации. *Высшее образование*. 2014. № 12. С. 112–116.
2. Бондаренко О.В. Применение мультимедийных технологий в образовательном процессе высшего учебного заведения. *Современные проблемы науки и образования*. 2017. № 3. С. 55–71.
3. Князева Г.В. Применение мультимедийных технологий в образовательных учреждениях. *Вестник Волжского университета им. В.Н. Татищева*. 2018. № 16. С. 77–95.
4. Андресен Б.Б. Мультимедиа в образовании: специализированный. М.: Дрофа, 2017. 221 с.
5. Захарова И.Г. Информационные технологии в образовании. М.: Академия, 2013. 188 с.
6. Плотніков М.С., Голопотилук Є.А., Рудніченко М.Д. Специфіка розробки проекту навчального мобільного застосування з сучасними мультимедійними засобами розширеної реальності. *XXI Всеукраїнська науково-технічна конференція молодих вчених, аспірантів та студентів* (м. Одеса, 22–23 квітня 2021 р.). Одеса, 2021. С. 85–86.
7. Плотніков М.С., Рудніченко М.Д., Граб К.В. Інтерактивна система вивчення астрономії з використанням технології доповненої реальності. *Інформатика, інформаційні системи та технології: тези доповідей шістнадцятої всеукраїнської конференції студентів і молодих науковців* (м. Одеса, 23 квітня 2021 р.). Одеса, 2021. С. 115–116.
8. Уваров А. Ю. Технологии виртуальной реальности в образовании. *Наука и школа*. 2018. № 4. С. 108–117.
9. Аверкина А.В., Лазарева О.Ю. Системы дополненной реальности для мобильных платформ. *Вестник Московского государственного университета печати*. 2015. № 1. С. 8–12.
10. Хокинг Д. Unity в действии. Мультиплатформенная разработка на C#. СПб.: Питер, 2019. 352 с.
11. Торн А. Основы анимации в Unity. М.: ДМК Пресс, 2016. 176 с.

References:

1. Granin, K. (2014). *Obrazovanie v kontekste globalizacii* [Education in the context of globalization]. *Vysshee obrazovanie – Higher education*, no. 12, pp. 112–116 [In Russian].
2. Bondarenko, O.V. (2017). *Primenenie mul'timedijnyh tehnologij v obrazovatel'nom processe vysshego uchebnogo zavedenija* [Application of multimedia technologies in the educational process of a higher educational institution]. *Sovremennye problemy nauki i obrazovanija – Modern problems of science and education*, no. 3, pp. 55–71 [In Russian].
3. Knjazeva, G.V. (2018). *Primenenie mul'timedijnyh tehnologij v obrazovatel'nyh uchrezhdenijah* [Application of multimedia technologies in educational institutions]. *Vestnik Volzhskogo universiteta im. V.N. Tatishcheva – Bulletin of the Volga University named after V.N. Tatishchev*, no. 16, pp. 77–95 [In Russian].
4. Andresen, B.B. (2017). *Mul'timedia v obrazovanii: specializirovannyj* [Multimedia in education: specialized]. M.: Drofa [In Russian].
5. Zaharova, I.G. (2013). *Informacionnye tehnologii v obrazovanii* [Information technologies in education]. M.: Akademiya [In Russian].

6. Plotnikov, M.S., Holopotyliuk, Ye.A., Rudnichenko, M.D. (2021). Specifika rozrobki proektu navchal'nogo mobil'nogo zastosuvannja z suchasnimi mul'timedijnimi zasobami rozshirenoi real'nosti [The specifics of the project development of educational mobile application with modern multimedia means of augmented reality]. *XXI Vseukraïns'ka naukovo-tehnichna konferencija molodih vchenih, aspirantiv ta studentiv (22–23 kvitnja 2021 hoda) – XXI All-Ukrainian scientific and technical conference of young scientists, graduate students and students.* (pp. 85–86). Odessa [In Ukrainian].
7. Plotnikov, M.S., Rudnichenko, M.D., Hrab, K.V. (2021). Interaktivna sistema vivchennja astronomii z vikoristannjam tehnologii dopovnenoï real'nosti [An interactive system for studying astronomy using augmented reality technology]. *Informatika, informacijni sistemi ta tehnologii: tezi dopovidej shistnadcjatoï vseukraïns'koï konferencii studentiv i molodih naukovciv (23 kvitnja 2021 hoda) – Informatics, information systems and technologies: abstracts of the sixteenth All-Ukrainian Conference of Students and Young Scientists.* (pp. 115–116). Odessa [In Ukrainian].
8. Uvarov, A.Ju. (2018). Tehnologii virtual'noj real'nosti v obrazovanii [Technologies of virtual reality in education]. *Nauka i shkola – Science and school*, no. 4, pp. 108–117 [In Russian].
9. Averkina, A.V., Lazareva, O.Ju. (2015). Sistemy dopolnenoj real'nosti dlja mobil'nyh platform [Augmented reality systems for mobile platforms]. *Vestnik Moskovskogo gosudarstvennogo universiteta pečati – Bulletin of the Moscow State University of Printing Arts*, no. 1, pp. 8–12 [In Russian].
10. Hoking D. (2019). *Unity v dejstvii. Mul'tiplatformennaja razrabotka na C#* [Unity in action. Multi-platform development in C #]. SPb.: Piter [In Russian].
11. Torn A. (2016). *Osnovy animacii v Unity* [Basics of animation in Unity]. M.: DMK Press [In Russian].

УДК 004.9

DOI <https://doi.org/10.32689/maup.it.2021.2.6>

Микола РУДНІЧЕНКО

кандидат технічних наук, доцент кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (nickolay.rud@gmail.com)

ORCID: <https://orcid.org/0000-0002-7343-8076>

Тетяна ОТРАДСЬКА

кандидат технічних наук, доцент кафедри інформаційних систем, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (tv_61@ukr.net)

ORCID: <https://orcid.org/0000-0002-5808-5647>

Денис ШИБАЄВ

аспірант кафедри технічної кібернетики та інформаційних технологій, Одеський національний морський університет, вул. Мечникова 34, Одеса, Україна, індекс 65029 (denscreamer@gmail.com)

ORCID: <https://orcid.org/0000-0002-3260-5843>

Ігор ПЕТРОВ

доктор технічних наук, професор кафедри морських перевезень, Національний університет «Одеська морська академія», вул. Дідріхсона 8, Одеса, Україна, індекс 65029 (firminess@list.ru)

ORCID: <https://orcid.org/0000-0002-8740-6198>

Микита ПОЛІКАРПОВ

студент, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом», вул. Фрометівська 2, Київ, Україна, індекс 03039 (mikpolicarp@gmail.com)

ORCID: <https://orcid.org/0000-0122-4911-3779>

Mykola RUDNICHENKO

PhD in Technical Sciences, Associate Professor of Information Technology, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (nickolay.rud@gmail.com)

Tatiana OTRADSKA

PhD in Technical Sciences, Associate Professor of Information Systems Department, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (tv_61@ukr.net)

Denis SHIBAYEV

Graduate student of the Department of Technical Cybernetics and Information Technologies, Odessa National Maritime University, 34 Mechnikova Street, Odessa, Ukraine, postal code 65029 (denscreamer@gmail.com)

Igor PETROV

Doctor of Technical Sciences, Professor of the Department of Maritime Transport, National University "Odessa Maritime Academy", 8 Didrichson Street, Odessa, Ukraine, postal code 65029 (firminess@list.ru)

Nikita POLIKARPOV

Student, Interregional Academy of Personnel Management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039 (mikpolicarp@gmail.com)

Бібліографічний опис статті: Рудніченко М., Отрадська Т., Шibaєв Д., Петров І., Полікарпов М. Розробка системи підтримки прийняття рішень для менеджера з управління ІТ-проектами. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 50–57. DOI: <https://doi.org/10.32689/maup.it.2021.2.6>

Bibliographic description of the article: Rudnichenko, M., Otradska, T., Shybaiev, D., Petrov, I., Polikarpov, M. (2021). Rozrobka systemy pidtrymky pryiniattia rishen dlia menedzhera z upravlinnia IT-proektamy [Decision support system for the IT project management manager]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 50–57. DOI: <https://doi.org/10.32689/maup.it.2021.2.6>

РОЗРОБКА СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ МЕНЕДЖЕРА З УПРАВЛІННЯ ІТ-ПРОЕКТАМИ

Анотація. Проблема ефективного, автоматизованого та швидкого пошуку адекватних варіантів вирішення як операційних, так і стратегічних питань, сучасними менеджерами ІТ-проектів не є вирішеною у повній мірі та потребує застосування не тільки якісних, а й кількісних підходів до оцінки найбільш ймовірних альтернатив. **Метою** статті є опис особливостей розробки системи підтримки прийняття рішень для менеджера з управління ІТ-проектами для формалізації складних ситуацій з великою кількістю альтернатив. Реалізація поставленої мети передбачає вирішення низки **завдань**: 1) аналізі сучасних аспектів та проблематики процесу прийняття рішень менеджерами проектів із урахуванням їх виробничої діяльності; 2) аналізі методів оцінки ймовірних альтернатив; 3) розробці концепції, алгоритму та основних форм інтерфейсу для обробки даних користувачем. **Наукова новизна.** У статті розглядається можливість застосування методу аналізу ієрархій у сукупності з методами системного аналізу для вирішення завдання автоматизації оцінки можливих альтернативних сценаріїв розвитку ситуацій шляхом розрахунку складових векторів та чисельних показників. Як **висновок**, у статті наголошується, що системи підтримки прийняття рішень на базі використання методів та алгоритмів системного аналізу, зокрема метода аналізу ієрархій, є ефективним технічним засобом автоматизації процесу оцінки різних варіантів дій менеджерів ІТ проектів у різних сценаріях розвитку подій під час виконання виробничих проектів. Результати розробки системи підтверджують її затребуваність, функціональність та наочність для використання менеджерами ІТ проектів та іншими керуючими посадовцями для зменшення ймовірності помилки через людський фактор. Реалізовані у системі візуальні можливості перегляду та зіставлення різних альтернатив одну з одною дозволяють забезпечити достатній рівень порівняльно-аналітичних можливостей менеджера проекту під час прийняття рішення.

Ключові слова: системи підтримки прийняття рішень, управління ІТ проектами, метод аналізу ієрархій.

DECISION SUPPORT SYSTEM FOR THE IT PROJECT MANAGEMENT MANAGER

Abstract. The problem of effective, automated and rapid search for adequate solutions to both operational and strategic issues, modern IT project managers is not fully addressed and requires the use of not only qualitative but also quantitative approaches to assess the most likely alternatives. **The aim** of the article is to describe the features of developing a decision support system for an IT project management manager to formalize complex situations with many alternatives. Realization of the set purpose provides the decision of a number of tasks: 1) the analysis of modern aspects and problems of process of decision-making by managers of projects taking into account their production activity; 2) analysis of methods for assessing probable alternatives; 3) development of the concept, algorithm and basic forms of the interface for data processing by the user. **Scientific novelty.** The article considers the possibility of applying the method of hierarchy analysis in combination with methods of system analysis to solve the problem of automating the assessment of possible alternative scenarios by calculating the constituent vectors and numerical indicators. In **conclusion**, the article emphasizes that decision support systems based on the use of methods and algorithms of system analysis, including the method of hierarchy analysis, are an effective technical means of automating the process of evaluating different actions of IT project managers in different scenarios during production projects. The results of the system development confirm its demand, functionality and clarity for use by IT project managers and other management officials to reduce the probability of error due to the human factor. The visual capabilities of viewing and comparing different alternatives implemented in the system allow to ensure a sufficient level of comparative and analytical capabilities of the project manager during decision-making.

Key words: decision support systems, IT project management, hierarchy analysis method.

Актуальність проблеми. В даний час можна спостерігати зростання обсягів і високого ступеня нечіткості, що надходить до керівників і менеджерів, фінансової, організаційної, технічної та іншої видів інформації. Через це необхідним є використання інформаційних систем і прикладних програмних продуктів, що дозволяють полегшити процеси прийняття рішень шляхом автоматизації процесів перебору можливих і найбільш ймовірних сценаріїв з безлічі альтернатив [1]. Не дивлячись на наявність готових програмних продуктів не всі з них можуть бути швидко і просто інтегровані в реальні процеси праці в конкретних організаціях в через їхню складність, перенасиченості функціями і високою вартістю ліцензій. Тому, вивчення можливостей, засобів і технологій для розробки систем підтримки прийняття рішень (СППР) є актуальною і затребуваною завданням в сучасному бізнесі [2].

Сучасна команда розробників різних програмних рішень в мінімальному наборі (від 4-6 чоловік) найчастіше включає в себе наступні співробітників [3]:

1. Програміст – здійснює написання коду програмного продукту в рамках проекту. Може мати одну з двох спеціалізацій (або обидві відразу): front-end (інтерфейс системи) або back-end (серверна сторона).

2. Тестувальник – виконує перевірки та затвердження функціоналу програмного продукту, створеного розробниками.

3. Дизайнер – здійснює проектування інтерфейсу системи і окремих її компонентів, з урахуванням ефективного і інтуїтивного взаємодії користувача з програмним продуктом.

4. Тімлід – лідер команди, який здійснює технічну перевірку і розподіл завдань в рамках команди розробників і несе відповідальність за дотримання якості і термінів роботи.

5. Адміністратор – виконує конфігурацію і технічну підтримку роботи апаратного обладнання та робочого програмного оточення для членів команди.

6. Спеціаліст по базах даних – проектує логічну структуру і взаємодію між сутностями БД.

У зв'язку з різним характером діяльності кожного співробітника необхідним є координація і планування їх праці з метою виконання проекту в терміни і на належному рівні якості, а також пошук замовників. Дані завдання допомагає вирішити менеджер, який є особою приймають управлінські рішення (ЛПР) [4].

Аналіз останніх досліджень і публікацій. Сучасний менеджер (МП) представляє собою фахівця широкого кола компетенцій, який здійснює управлінську діяльність в конкретній предметній області, пов'язаної з виробничою діяльністю компанії.

Даний фахівець, як правило, займає постійну посаду в фірмі, тому він наділений необхідними повноваженнями для оперативного прийняття управлінських рішень різного ступеня відповідальності.

Ключовими операціями в діяльності типового МП є наступні [5]:

1. Установка цілей, тобто менеджер визначає і розподіляє завдання в існуючих групах цілей, приймаючи рішення про порядок дій, спрямованих для досягнення всіх цілей.

2. Організація процесу діяльності компанії. Це має на увазі аналіз ключових видів діяльності, заходів, що вживаються і рішень, розділяє їх на окремі керовані сукупності, які в свою чергу поділяються на конкретні трудові завдання. Після цього може здійснюватися підбір виконавців для контролю і виконання поставлених завдань.

3. Підтримка процесів мотивації і комунікації в організації. Дана операція здійснюється шляхом складання трудового команди з відповідальних людей. Потім проводиться розробка заходів підвищення продуктивності діяльності співробітників і засобів визначення якості праці.

4. Встановлення показників і одиниць вимірювання ефективності діяльності компанії.

Його завдання забезпечити розуміння і орієнтацію співробітників на результат, причому як на виконання оперативних завдань, так і на досягнення довгострокових перспектив.

Це досягається завдяки аналізу, оцінці та інтерпретації отриманих результатів за окремі періоди часу.

5. Підвищення рівня своєї компетенції і знань співробітників, що дозволяє забезпечити особистісний ріст і підвищити ефективність роботи фірми.

Специфіка виконуваних МП завдань зумовлює творчий інтелектуальний характер діяльності. Однією з концептуальних завдань менеджера є підвищення продуктивності використання людських або фінансових ресурси компанії [6].

Ключовими завданнями по автоматизації праці МП є наступні [3]:

- ефективна і своєчасна підтримка процесів оперативної діяльності, організація завдання контролю і обліку;

- формування різних типів електронних документів для партнерів, в тому числі рахунков-фактури, податкові накладні, ділові пропозиції, спеціальні пропозиції для замовників;

- оперативне отримання і формування звітів про поточний стан і значення фінансових показників в організації за обраний період часу;

- оптимізація витрачаються витрат на роботу співробітників команди за допомогою підвищення ступеня ефективності витрат робочого часу шляхом позбавлення від рутинної роботи;

- зменшення негативного впливу людського фактора на ключові бізнес-процеси;

- захищене зберігання даних;

- збільшення загального рівня якості обслуговування клієнтів.

В умовах існуючої економічної кризи значно змінюються пріоритети. До найбільш популярних трендів в сфері автоматизації діяльності сучасних МП слід віднести [7]:

- застосування хмарних технологій. Ключова відмінність такого методу полягає у відсутності власного сервера в компанії. У представників малого бізнесу подібні системи користуються високою популярністю, що обумовлено зниження витрат на обслуговування обладнання;

- автоматизація механізмів організації маркетингу шляхом інтеграції CRM-систем, систем управління і підбору контекстної реклами, систем бізнес-аналітики [8];

– розвиток технологій взаємодії. Активно розвивається напрямок M2M – Machine to Machine, що знижує участь людини у виробничих процесах, тому що ускладнення проведення бізнес-процесів та інтеграції обладнання не дозволяє співробітнику адекватно реагувати на модифікацію ситуації;

– використання інструментів Big Data. Важливим напрямком є адаптація засобів проведення бізнес-аналітики до збережених облікових даних [9];

– методи аналізу даних в online режимі. Системи підтримки діяльності МП організацій все частіше починають еволюціонувати в сфері обробки транзакцій в режимі real-time, що забезпечує синхронізацію операцій, що проводяться [10].

В рамках загальносистемної концепції аналізу та оцінки альтернативних сценарії управління проектами виконують за допомогою різних методів [4–8], які в загальному випадку діляться на детерміновані, імовірнісні, експертні, методи оцінки в умовах невизначеності нестатичних природи, комбіновані. Результаті порівняльного аналізу цих методів наведено у таблиці 1.

Розглянуті методи є доцільними у різних ситуаціях та проектах, однак усі вони потребують значних витрат часу на інтерпретацію результатів, тому більш доцільним для нашої роботи є використання ієрархічних методів прийняття рішень.

Метою статті є розробка системи підтримки прийняття рішень для менеджера з управління ІТ-проектами для формалізації складних ситуацій з великою кількістю альтернатив.

Таблиця 1

Методи оцінки альтернатив при прийнятті рішень

Метод	Назва	Переваги	Недоліки
Детермінований	Марківського аналізу	Можливість обчислення ймовірностей станів систем з відновленням і множинними станами деградації	Оперування тільки двома можливими станами елементів системи
	Монте-Карло	Адаптація до будь-якого розподілу вхідних даних. Відносна простота моделей і зручність розширення	Неможливо адекватно моделювати події з дуже високою або дуже низькою ймовірністю появи
Метод	Назва	Переваги	Недоліки
Ймовірнісний	Байесовських мереж	Для застосування методу досить знання апіорної інформації, логічно виведені затвердження легкі для розуміння, використовуються суб'єктивні ймовірнісні оцінки	Трудомісткість визначення всіх взаємодій в мережах Байеса через необхідність отримання умовних ймовірностей за допомогою експертних методів
Експертних оцінок	Нечітких множин	Можливість спрощення розрахункових математичних моделей	Недолік відомостей по функціям розподілу параметрів. Спрощеність розрахункових моделей знижує вірогідність оцінок
Ймовірнісний	Дерева несправностей	Дозволяє здійснити аналіз різноманітних факторів, врахувати впливу відмов безпосередньо пов'язаних з кінцевим подією	Невизначеність оцінок ймовірностей базисних подій впливає на оцінку ймовірності і вірогідність виникнення кінцевого події
	Логіко-ймовірнісний	Результати застосування методу, містять значення вагових значень елементів, що дозволяє оцінити їх значимість	Складність у визначенні точних значень ймовірностей через статичного характеру подій
Оцінки в умовах невизначеності нестатичної природи	Дослідження небезпеки і працездатності	Застосовність до широкого класу систем, процесів і процедур	Тривалий аналіз у часі. Обмеженість завданнями проекту, областю і цілями дослідження
Метод	Назва	Переваги	Недоліки
Оцінки в умовах невизначеності нестатичної природи	Аналізу дерева рішень	Точне графічне представлення деталей вирішення проблеми. Можливість розрахунку оптимального шляху вирішення проблеми	Застосування діаграми дерева рішень може привести до зайвого спрощення ситуації

Виклад основного матеріалу. СППР в нашому випадку являє собою інтерактивну інформаційну систему, яка надає МП інтерфейс для використання даних і моделей для аналізу і вирішення необхідних завдань з метою вироблення рекомендацій щодо прийняття управлінських рішень [11]. Концептуально СППР інтегрує в своєму складі функціонал роботи з інтерактивними запитами на простій мові.

Ключовими функціональними можливостями створеної СППР є: введення даних, фільтрація і агрегування даних, виключення дубльованих даних, зберігання і аналіз даних [12].

Основоположним компоненти розробленої СППР є наступні:

1. Інтерфейс користувача. ЛПР використовує форми інтерфейсу для введення предметної інформації і команд в систему, а також для отримання і відображення вихідної інформації. Команди по взаємодії з системою можуть включати в себе параметри, які регламентують процес обробки і використання збережених знань. Необхідна інформація видається у формі значень, які зберігаються в певних колекціях даних.

2. База знань (БЗ) містить набори виявлених фактів, які формалізують проблемну область, виявляючи логічний взаємозв'язок всіх фактів між собою. Найбільш критичними для БЗ є набори логічних правил. Окреме правило дозволяє однозначно визначити порядок дій при виникненні конкретної ситуації, і включає до свого складу: умова, яка може виконуватися або не виконав, і дія, вироблене при виконанні заданої умови.

3. Інтерпретатор, це модель, який здійснює в певному порядку обробку знань у БЗ. Логіка функціонування інтерпретатора зводиться до послідовного аналізу сукупності логічних правил. У тому випадку, коли умова, яке міститься в окремому правилі, дотримується в повній мірі, то виконується заданий вплив, а користувач отримує можливість ознайомитися з результатами підбору потрібних рішень задачі.

4. Модуль обробки даних СППР, необхідний створення логічної послідовності ієрархій правил в БЗ, на основі використання методу аналізу ієрархій (МАІ), який представляє математично обґрунтований підхід до вирішення проблем прийняття рішень МП на базі використання системного аналізу.

МАІ не регламентує МП жорстко задане, найбільш пріоритетне рішення, надаючи гнучкі можливості пошуку підходящої альтернативи в інтерактивному режимі, що дозволяє найкращим чином сформувати розуміння процесу рішення задачі. В рамках нашої системи реалізація МАІ розглядає відразу кілька варіантів вирішення задачі, базується на заданому чисельному критерії (обмеженні), за яким здійснюється перевірка і оцінка заходів рішення та оперує заданими логічними умовами, в яких відбувається аналіз і рішення поставленої проблеми, а також відомі причини, що впливають на вибір конкретного рішення.

Структурування даної задачі відбувається у вигляді складання ієрархічної схеми з декількома рівнями для опису цілей, оціночних критеріїв і допустимих альтернатив.

Створювана ієрархія будується з вершини, яка є відображенням головної мети розв'язуваної задачі, через проміжні рівні, які представляють собою уніфіковані критерії порівняння, до нижнього рівня, що має вигляд безлічі допустимих альтернативних варіантів дій.

На базі виконання парних порівнянь елементів системи кожного рівня отримані результати приводяться до числового значення за допомогою спеціальної таблиці.

Після ієрархічного відтворення вирішуваної проблеми виконується побудова матриці порівняння допустимих критеріїв і чисельний розрахунок всіх значень пріоритетів по кожному з критеріїв.

В системі всі елементи порівнюються попарно по відношенню до їх безпосереднього впливу на загальну характеристику (рішення).

Ступінь інтенсивності оцінюється на базі використання шкали в бальних оцінках від 1 до 9.

Після цього виконується розрахунок коефіцієнтів важливості для елементів кожного рівня. При цьому перевіряється узгодженість суджень МП. Порівнюються результати, отримані на нижньому рівні, тобто набір альтернатив, отриманий між собою окремо по кожному обраному критерію проміжного рівня. На базі цього здійснюється підрахунок кількісного індикатора якості кожної з альтернатив і визначення найкращої альтернативи. Визначається підсумкове значення ваги кожної альтернативи шляхом множення на пріоритет відповідного критерію на вищому рівні і подальшим підсумовуванням по кожному елементу відповідно до критеріїв, на які впливає елемент. Алгоритм роботи системи має наступний вигляд (рис. 1). Спочатку виконується завантаження головної форми системи, далі користувач вводить кількість елементів 2-го і 3-го рівнів для побудови відповідних матриць. На базі цього динамічно генеруються відповідні компоненти інтерфейсу користувача і здійснюється створення ієрархії і матриць для кожного рівня альтернатив. Наступним кроком є введення користувачем даних в матриць альтернатив, якщо введені дані надані в коректній формі, то виконується розрахунок підсумкової матриці альтернатив. Далі, якщо користувач вибирає побудову графічної візуалізації результатів розрахунку, то відбувається генерація відповідних графіків в кругової і стовбурових формах. Наступним кроком є можливість збереження отриманих результатів в окремий файл, з якого при необхідності можна знову завантажити дані в систему.



Рис. 1. Загальний алгоритм взаємодії користувача з системою

Приклад розрахунку показників та вектору пріоритетів за критерієм 2-го рівня наведено на рис. 2.

Критерій - Назва елемента, рівень 2						Вектор пріоритетів
Назва елемента 3.1	1	2	3	4	5	0.395860199327
Назва елемента 3.2	1/2	1	2	3	4	0.249770900153
Назва елемента 3.3	1/3	1/2	1	6	7	0.224246254437
Назва елемента 3.4	1/4	1/3	1/6	1	5	0.089131885022
Назва елемента 3.5	1/5	1/4	1/7	1/5	1	0.040990761062
Lmax=5.5061352 IIC=0.1265338 OC=11.29766127						
Розрахунок	Кругова діаграма	Гістограма				

Рис. 2. Приклад розрахунку показників та вектору пріоритетів за критерієм 2-го рівня

Після проходження усіх етапів оцінок за кожним критерієм по усіх рівнях ієрархії формується кінцева матриця розрахунків, приклад наведено на рис. 3.

Кінцева матриця				
	0.149373136132	0.474230146864	0.376396717004	Загальні ваги
Назва елемента 3.1	0.395860199327	0.564657911783	0.593472083145	0.550289627605
Назва елемента 3.2	0.249770900153	0.196161321485	0.224325835682	0.214770183057
Назва елемента 3.3	0.224246254437	0.101266934784	0.093909210233	0.116867318074
Назва елемента 3.4	0.089131885022	0.071751347114	0.045416099032	0.064435031649
Назва елемента 3.5	0.040990761062	0.066162484834	0.042876771908	0.053637839614

Рис. 3. Результуюча матриця розрахунку системою варіантів альтернатив

Система складається з наступних форм:

1. Головна форма – містить відповідні компоненти для вибору кількості елементів 2-го і 3-го рівнів, подальшого їх іменування та визначення в рамках системи. Також ця форма містить кнопки для створення матриць і ієрархій.

2. Форма створення матриці – дозволяє за допомогою таблиці здійснити введення даних в осередку для розрахунку вектора пріоритетів і відповідних значень. На базі отриманих розрахунків стає можливим побудова відповідних ієрархій альтернатив.

3. Опції матриці – дозволяє здійснювати зміну кольору налаштувань, максимальний пріоритет, м'якість зміни пріоритету на діаграмі або в розрахунках.

4. Опції ієрархії – дозволяє включати можливість називати елементи ієрархії в автоматичному режимі, виділяти рівні ієрархії окремо і активувати режим збереження ресурсів системи для оптимізації використання розрахункових етапів.

Висновки та перспективи подальших досліджень. Розроблена СППР дозволяє МП: автоматизувати процес оцінку обстановки, здійснити вибір критеріїв і оцінити їх відносну важливість; згенерувати можливі рішення; здійснити оцінку альтернативних сценаріїв прийняття рішень з управління процесом розробки; забезпечити постійний обмін інформацією про обстановку прийнятих рішень і допомогти узгодити групові рішення; здійснювати аналіз можливих наслідків прийнятих рішень; проводити збір даних і оцінку результатів аналізу.

Список використаних джерел:

1. Бенін Д.М. Системи підтримки прийняття рішень. М. : Тріада, 2019. 165 с.
2. Головіна Є.Ю. Інтелектуальні методи для створення систем підтримки прийняття рішень. М. : MEI, 2016. 235 с.
3. Грекул В.І. Методичні основи управління IT-проектами. М. : Інтуїт Біном, 2019. 392 с.
4. Двордіца Ю.С. Інтелектуальні системи підтримки прийняття рішень. Луганськ : СТУ, 2015. 122 с.
5. Баронів В.В. Інформаційні технології та управління підприємством. М. : АйТі, 2018. 512 с.
6. Іващенко Т.І. Системи підтримки прийняття рішень. Хабаровськ : ТОДУ, 2015. 100 с.
7. Оводенко А.А. Формування інформаційного забезпечення для підтримки прийняття рішень на підприємстві. СПб. : СПбГУАП, 2016. 411 с.
8. Пуха Г.П. Системи підтримки прийняття рішення. СПб. : СПбДЕУ, 2018. 386 с.
9. Рудніченко Н.Д., Вычужанин В.В., Козлов А.Е. Модель front-end прототипа системы поддержки принятия решений мониторинга и управления рисками сложных технических систем. *Інформаційні управляючі системи та технології* : матеріали міжнародної науково-практичної конференції (м. Одеса, 22–24 вересня 2015 р.) Одеса, 2015. С. 198–201.
10. Vychuzhanin V.V., Rudnichenko N.D., Shibaeva N.O., Boyko V.D. The development of user interface prototype of decision support system for risk management of complex technical systems. *Sustainability and Competitiveness in Business*. 2016. P. 162–172.
11. Рудніченко М.Д., Кривиленко І.Є. Розробка модульної структури інформаційної системи оцінки пріоритетів альтернатив управлінських рішень. *Сучасні інформ. технології та телекомунікаційні мережі : тези доп. 56-ої наук. конф. молодих дослідників ДУОП-бакалаврів* (м. Одеса, 2021 р.). Одеса, 2021. С. 55–59.
12. Рудніченко М.Д., Кривиленко І.Є. Проектування інтерфейсу інформаційної системи оцінки пріоритетів альтернатив управлінських рішень. *Сучасні інформ. технології та телекомунікаційні мережі : тези доп. 56-ої наук. конф. молодих дослідників ДУОП-бакалаврів*. (м. Одеса, 2021 р.). Одеса, 2021. С. 144–148.

References:

1. Benin, D.M. (2019). *Sistemi pidtrimki priynjattja rishen'* [Decision support systems]. M.: Triada [In Ukrainian].
2. Golovina, E.Yu. (2016). *Intelektual'ni metodi dlja stvorennja sistem pidtrimki priynjattja rishen'* [Intelligent methods for creating decision support systems]. M.: MEI [In Ukrainian].
3. Grekul, V.I. (2019). *Metodichni osnovi upravlinnja IT-proektami* [Methodical bases of management of IT projects]. M.: Intuit Binom [In Ukrainian].

4. Dvordica, Ju.S. (2015). *Intelektual'ni sistemi pidtrimki priynjattja rishen'* [Intelligent decision support systems]. Lugans'k: SNU [In Ukrainian].
5. Baroniv, V.V. (2018). *Informacijni tehnologii ta upravlinnja pidpriemstvom* [Information technology and enterprise management]. M.: AjTi [In Ukrainian].
6. Ivashhenko, T.I. (2015). *Sistemi pidtrimki priynjattja rishen'* [Decision support systems]. Habarovs'k: TODU [In Ukrainian].
7. Ovodenko, A.A. (2016). *Formuvannja informacijnogo zabezpechennja dlja pidtrimki priynjattja rishen' na pidpriemstvi* [Formation of information support to support decision-making in the enterprise]. SPb.: SPbGUAP [In Ukrainian].
8. Puha, G.P. (2018). *Sistemi pidtrimki priynjattja rishennja* [Decision support systems]. SPb.: SPBDEU [In Ukrainian].
9. Rudnichenko, N.D., Vychuzhanin, V.V., Kozlov, A.E. (2015). Model' front-end prototipa sistemy podderzhki prinjatij reshenij monitoringa i upravlennja riskami slozhnyh tehniceskikh system [The front-end model of the prototype of the decision support system for monitoring and risk management of complex technical systems]. *Information Control Systems and Technologies: Materiali Mizhnarodnoi naukovo-praktichnoi konferencii (23–24 veresnia 2015 hoda) – Proceedings of the International Scientific and Practical Conference.* (pp. 198–201). Odessa [In Russian].
10. Vychuzhanin, V.V., Rudnichenko, N.D., Shibaeva, N.O., Boyko, V.D. (2016). The development of user interface prototype of decision support system for risk management of complex technical systems. *Sustainability and Competitiveness in Business* [In English].
11. Rudnichenko, M.D., Krivilenko, I.Y. (2021). Rozrobka modul'noi strukturi informacijnoi sistemi ocinki prioritativ al'ternativ upravlins'kih rishen' [Development of a modular structure of the information system for assessing the priorities of alternatives to management decisions]. *Modern inform. technologies and telecommunication networks: tezi dop. 56-oї nauk. konf. molodih doslidnikiv DUOP-bakalavriv (2021 hod) – theses add. 56th Sciences. conf. young researchers of DUOP bachelors.* (pp. 55–59). Odessa [In Ukrainian].
12. Rudnichenko, M.D., Krivilenko, I.Y. (2021). Proektuvannja interfejsu informacijnoi sistemi ocinki prioritativ al'ternativ upravlins'kih rishen' [Designing the interface of the information system for assessing the priorities of alternatives to management decisions]. *Modern inform. technologies and telecommunication networks: tezi dop. 56-oї nauk. konf. molodih doslidnikiv DUOP-bakalavriv (2021 hod) – theses add. 56th Sciences. conf. young researchers of DUOP bachelors.* (pp. 144–148). Odessa [In Ukrainian].

UDC 619.12

DOI <https://doi.org/10.32689/maup.it.2021.2.7>

Ruslan SKURATOVSKIY

Lecturer at the Department of Computational Mathematics and Computer Modeling, Interregional Academy of personnel management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039

ORCID: <https://orcid.org/0000-0002-5692-6123>

Anna BAZARNA

Student, Institute for Applied System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37 Peremohy Avenue, Kyiv, Ukraine, postal code 03056

Руслан СКУРАТОВСЬКИЙ

викладач кафедри обчислювальної математики і комп'ютерного моделювання, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом», вул. Фрометівська 2, Київ, Україна, індекс 03039

Анна БАЗАРНА

студентка, Інститут прикладного системного аналізу, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», просп. Перемоги 37, Київ, Україна, індекс 03056

Bibliographic description of the article: Skuratovskiy, R., Bazarna, A. (2021). Analiz movlennievkykh osoblyvostei shkaly mel z vykorystanniam elektrohrfichnykh i movnykh syhnaliv parametryzovanyim metodom KNN i XGBOOST [Analysis of the mel scale features using electrography and speech signals by parameterized KNN and XGBOOST]. *Informatsiini tekhnologii ta suspilstvo – Information technology and society*, 2, 58–74. DOI: <https://doi.org/10.32689/maup.it.2021.2.7>

Бібліографічний опис статті: Скуратовський Р., Базарна А. Аналіз мовленнєвих особливостей шкали мел з використанням електрографічних і мовних сигналів параметризованим методом KNN і XGBOOST. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 58–74. DOI: <https://doi.org/10.32689/maup.it.2021.2.7>

**ANALYSIS OF THE MEL SCALE FEATURES USING ELECTROGRAPHY
AND SPEECH SIGNALS BY PARAMETERIZED KNN AND XGBOOST**

Abstract. Recognizing human emotions and speech is something that has always been an exciting research topic for scientists. In our work, we show that the parametrization of a vector is obtained and carried out from a sentence divided into an emotionally-informational containing part and an effectively used informational part. There are several characteristics of speech that distinguish it between utterances. Pitch, timbre, intensity, and vocal tone classify speech into multiple emotions. We have supplemented them with a new function of speech classification, which consists in dividing a sentence into an emotionally charged part of a sentence and a part that carries only informational load. Therefore, we can conclude that the speech pattern changes under the influence of different emotional environments. Since the identification of the emotional states of the speaker can be made on the basis of the Mel scale, MFCC is one of such options for studying the emotional aspects of the speaker's utterances. In this paper, we implement a model to identify multiple MFCC emotional states for two datasets, classify the emotions for them based on MFCC characteristics, and compare them accordingly. Overall, this work implements a dataset minification-based classification model that uses averaging functions to improve the level of classification accuracy in various machine learning algorithms. In addition to the static analysis of the author's tonal portrait, which is especially used in MFCC, we propose a new method of dynamic analysis of a sentence in processing and research as a new linguistic-emotional entity pronounced by the same author. Due to the ranking by importance of the characteristics of the MEL scale, we can parameterize the coordinates of the vectors being processed using the parameterized KNN method.

Language recognition is a multilevel pattern recognition task. Here, acoustic signals are analyzed and structured as a hierarchy of structural elements, words, phrases and sentences. Each level of such a hierarchy can provide time constants: possible word sequences or known pronunciation types that reduce recognition errors to a lower level.

The analysis of the dynamics of voice and speech is suitable for improving the quality of human perception and the formation of human speech by a machine. It is within the capabilities of artificial intelligence. Emotion results can be widely applied in e-learning platforms, automotive systems, medicine, etc.

Key words: machine learning, speech recognition, emotion recognition, MFCC, supervised learning, decision trees, MEL scale.

АНАЛІЗ МОВЛЕННЄВИХ ОСОБЛИВОСТЕЙ ШКАЛИ MEL З ВИКОРИСТАННЯМ ЕЛЕКТРОГРАФІЧНИХ І МОВНИХ СИГНАЛІВ ПАРАМЕТРИЗОВАНИМ МЕТОДОМ KNN І XGBOOST

Анотація. Розпізнавання емоцій та людської мови завжди було захоплюючим викликом для вчених. У нашій роботі для вектора даних, отриманого з речення, реалізується і ефективно застосовується параметризація цього вектора, розділеного на містить емоційно забарвлену частину і інформаційну частину. Виразність людської мови посилюється емоціями, які вона передає. Існує кілька характеристик та особливостей мови, які розрізняють її серед висловлювань, тобто різні провідні характеристики, такі як висота, тембр, гучність. Ми доповнили їх новим класифікаційними ознакою мови, що полягає в поділі пропозиції на емоційно навантажену частину пропозиції і частина, яка несе тільки інформаційне навантаження. Таким чином, еталонна фраза з даного класу емоцій змінюється оскільки вибір відбувається з відповідного класу емоцій для підрахунку її відстані методом параметризованого KNN / зразок мови змінюється, коли він піддається впливу різних емоційних середовищ. Оскільки визначення емоційних станів мовця може бути виконано на основі шкали MEL, MFCC є одним з таких варіантів вивчення емоційних аспектів висловлювань мовця. У цій роботі ми реалізуємо модель для визначення декількох емоційних станів з MFCC для декількох наборів даних, класифікуємо емоції для них на основі характеристик MFCC і даємо відповідне порівняння. Окрім статистичного аналізу тонального портрету автора, який застосовується, зокрема, у MFCC, ми запропонували новий метод динамічного аналізу лише те, що сказано фразами, як нова лінгвістично-емоційна сутність виробленої тем самим автором. Завдяки ранжуванню згідно важливості особливостей голосової шкали нам вдалося параметризувати координати векторів, які будуть оброблятися методом параметризованого KNN.

Розпізнавання мови – це багаторівнева задача розпізнавання образів. Тут акустичні сигнали аналізуються і структуруються в ієрархію структурних елементів, слів, фраз і пропозицій. Кожен рівень такої ієрархії може передбачати деякі тимчасові константи: можливі послідовності слів або відомі види вимови, які дозволяють зменшити кількість помилок розпізнавання на більш низькому рівні.

Аналіз голосу і динаміки мови доцільний для підвищення якості сприйняття людиною і формування людської мови машиною і знаходиться в межах можливостей штучного інтелекту. Результати визначення емоцій можуть широко застосовуватися в платформах електронного навчання, автомобільних бортових системах, медицині і т. д.

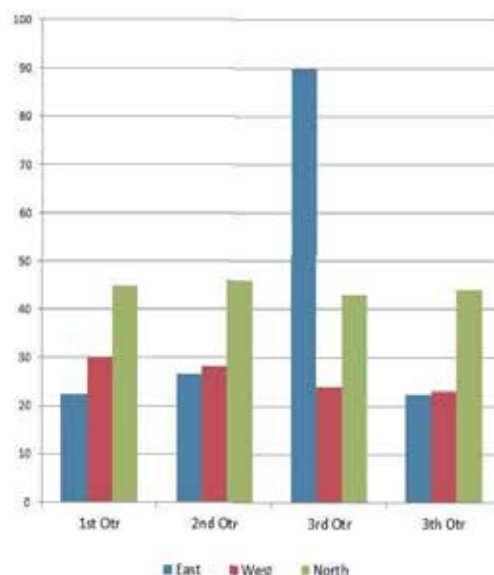
Ключові слова: машинне навчання, розпізнавання мови, розпізнавання емоцій, MFCC, контрольоване навчання, дерева рішень, шкала MEL.

Introduction. We use Miller function and scale in order to receive the emotion score differences. In fact, the vocal acoustics are full of emotional cues, so it is possible to analyze the speaker's emotional state. A sentence was divided by us into two parts as far as the expression of emotions frequently occurs either at the beginning or at the end of a sentence. We refer a beginning and an end of a sentence, referred to the emotional content expressing by an author to the first part. Then, to the second one we refer a middle of a sentence containing only the informational and narrative part. It serves to vector parameterization in KNN for speech emotions. Emotions can be recognized both from text and sound. Each of them has different approaches to identify the emotional state of the speaker. Moreover, the emotions are a communication bridge among the speaker and the listener. An interaction between individuals is way more clear and effective when the emotions are used in utterances. They play a pivotal role in engaging of a human being in a group discussion and can tell a lot about the one's mental state [1]. The information hidden in emotions ignited the process of the evolution of the speech recognition field commonly referred as automatic speech recognition. Several models of retrieval and interpretation of emotions from images of speaker's face and the recordings of his expressions, voice and tone during a conversation has been proposed by researchers. The utilization of physiological signals in the same manner has also been discussed [2]. The significance of emotions in communication can hardly be overestimated since they express the speaker's intentions to his listeners. There are several spoken language interfaces available today that support automatic speech recognition. By collecting the samples, such systems are providing a base for the speech recognition field [3]. The currently available speech systems are able of processing naturally spoken utterances with high accuracy, but the lack of emotional component makes the ASR systems less realistic and meaningful. There are several real-world fields that may benefit from the recognition of the emotional context of an utterance, such as entertainment, emotion-based audio file indexation, HCI-based systems etc. [4].

There are some of the selected features, which can be trained to classify, recognize and predict emotions. There are several emotions that can be extracted from the utterances. Few of the universally enlisted among them are Happiness, Fear, Sadness, Anger, Neutral, and Surprise. These emotions can be recognized by any intelligent system, constrained by computational resources. The implementation of the emotional sector of speech makes the human-computer interactions more real and efficient. The analysis of voice and speech for the sake of enhancing the quality of human conversations is reasonable and within the bounds of possibility. The results of emotion detection can be broadly applied in e-learning platforms, car-board systems, medical field etc.

The remaining sections have the following content: Section 2 contains literature observation on the topic, Section 3 is dedicated to the description of the problem, Section 4 carries the details of method implementation and results achieved under problem solution and finally Section 5 is the conclusion. In this paper we prove efficiency of the concept of using only 12 MFCC from 39, have identified which 12 MFCC to use for speech and emotion analysis. The dataset used for this experimentation is EMODB. Variants of supervised learning approaches have been implemented to classify emotions from two databases EMODB and SAVEE.

Literature Review. As well known KNN makes prognostications on time by quick calculating the similarity between an input sample and each training exemplar. Spectral analysis is a promising technique for detecting emotions from sample speech. Its purpose is to use a database in which the data points are separated into several classes to predict the classification of a new sample point. Spectral analysis is a promising method of emotion detection from samples of speeches. Prosodic features of speech signals can also be used for analyzing emotions since they contain emotional information. Researchers explored the role and context of emotions by using a set of 88 features called eGmaps [5]. Speech patterns can be obtained from combination of various speech features acquired from speaker's utterances. Feature selection plays a pivotal role in the differentiation of different emotions of the same speaker from his speech [6] and it relies on selecting the best features from the signal. Different human languages have different accents, structures of sentences, and speaking styles [7] thus making the identification of emotions from utterances challenging. Various aspects of spoken languages alter the extracted features of the sound signal. It is possible that a sample speech may have more than one emotion which means that each emotion corresponds to a different part of the same speech signal, which complicates the setting of boundaries between emotions. An attempt has been made to study models of the multilingual emotion classification in literature [8].



The substantial advancement in technology has boosted the development of the emotion recognition fields [9; 10]. Call-centers and remote education for example [11] Existing speech recognition systems can be improved by implementing spectral analysis [12].

Authors have identified classes of features extracted from speech electrography and signals. There is an important aspect of SER that includes characterization of emotional content of a speech [13]. Several speech features are obtained from speech acoustic analysis and can be used to detect and predict emotions [14]. The aim of selecting speech features is to determine properties that can improve the rate of classification emotions from a set [15]. The machine learning models are flexible enough to adapt themselves to any model that studies emotions and show good performance in predicting tasks based on selected features [16].

Authors identified classes of features extracted from emotional speech features electrography and speech signals. There is an important aspect of SER that includes characterization of emotional content of speech [13]. Several speech features are obtained from speech acoustic analysis and can be used to detect and predict emotions [14]. The aim of selecting speech features is determination of properties that can improve rate of classification for emotions from a feature set [15]. The machine learning approaches are flexible enough to adapt themselves to any model that studies emotion and show good performance perform well predicting tasks based on selected features [16].

Speech, Emotion and Classification. Emotions are intertwined with mood, temperament, personality, sentiment and motivation [17]. Emotions can be understood as a complex feeling of the mind that results in physiological and physiological changes. Human thoughts and behavior is influenced by emotions and there are instances of changes in body when it encounters different emotional states [18]. Literature [19; 20] proves that there is a considerable influence of emotional syndromes on human actions and reactions. Several applications created by researchers rely on emotion detection as an integral component for identification of behavioral patterns [21; 22]. Speech features can be extracted from various sources to accomplish predictive analytics. The list of sources includes vocal tract, excitation source and prosodic extraction.

Emotions can be broadly studied using both discrete and continuous approaches. Different classes represent different kinds of emotions and the continuous approach of studying of emotions is a derivation based on combination of several psychological measurements on different axes [24]. Speech Emotion Recognition in the main identifies emotions on the basis of categorical approach, which depends on usage of common words. Researchers derive emotions from expressions of face, speech and various physiological signals. The analysis of facial expressions is a great way to find emotions [25–29] since human face displays emotions very aptly even without a single word uttered. Voice recordings are potentially important for expressing the speakers' mental state and their intentions. Speech features can be studied as vectors for detection of emotion from a data set. [30–32]. The autonomous nervous system allows to assess an emotion and thus utilize physiological signals like ECG, RSP BP to recognize people's emotions and possibly help cure mental illnesses [17].

We took into consideration the line of temperament. We also integrate the state in which a representative of this type of temperament may be. There are 8 axes of Miller are applied by us. Also it is optimal to take 8 axes.

SER [23] – Speech Emotion Recognition System (Component and tasks)	
Feature Extractor	Emotion Classifier
Takes signal input and generates emotion feature	Map the speech with one or more emotions

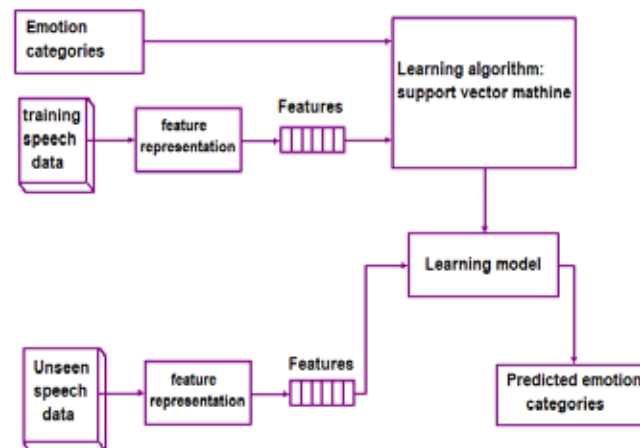


Fig. 1. Framework for supervised emotion classification [23]

Consciously controlling the volume level, for example, to emphasise a secret, even an angry person can make the voice calmer and quieter in order to show that it was a secret, a question or the essence of the issue. On the contrary, in order to highlight the characterisation of some hero of his story, the speaker can consciously increase the amplitude of the voice.

Physiological features of voice and hearing. It takes into account the average amplitude (frequency) and other characteristics of the person's voice, obtained on the basis of data from the database.

Given the above assumptions, if we wish to approach the study of temperament, we first need an understanding of emotion. Such understanding is not easy to come by a glance at a typical textbook of psychology will show

that «emotion» is used to refer to a ragbag of apparently disconnected facts and is never itself clearly defined at all. Yet, within one branch of psychology, namely, animal learning theory, there has long been a reasonably clear consensus that emotions consist of states elicited by stimuli or events which have the capacity to serve as reinforcers for instrumental behavior. This, for example, is the framework within which Miller analyzed the concept of «fear» and its role in avoidance learning [69]. The term temperament has considerable overlap with dimensions of personality and with emotional, cognitive, and behavioral functions.

Emotional Speech databases. There is need of suitable databases to train the emotion recognition systems. Researchers suggest several existing databases aligned with the task of detecting and classifying the emotions. These data bases can be categorized into three broad domains that cover acted emotions, natural emotions and felicitated emotions [33]. Out of the three mentioned domains enacted emotions are frequently supported by research since they are strong and reliable. EMODB is one of such highly used database for emotional classification. SAVEE is yet another enacted emotion database used for studying emotions. EMODB is a Berlin database for emotions while SAVEE is an English database specifying various emotions.

Statistical Data Corpus. Three databases have been utilized for training and testing:

1. BERLIN DATABASE
2. SENTIMENT DATABASE English
3. SAVEE DATASET

Berlin Database was created in 1999 and consists of utterances spoken by various actors. EMODB has different number of spoken utterances for seven emotions [34]. Emotions included in the database are anger, boredom, disgust, fear, happiness, indifference, sadness. The dataset contains more than 500 utterances spoken by 51 male and 60 female actors from the age 21 to 35 years. The emotions labelled in EMODB are listed in Table 1.

Surrey Audio-Visual Expressed Emotion (SAVEE) [35], [36]. The increasing demand of research in speech analysis led to the development of SAVEE database recordings to help study automatic emotion recognition system. The database contains recordings from 4 male actors in 7 different emotions, 480 British English utterances in total.

TIMIT corpus was used for sentence selection and contains phonetically-balanced emotions. The data were recorded in a visual media lab with efficient audio-visual equipment. The recordings were then processed and labelled. 10 subjects under audio, visual and audio-visual conditions evaluated quality of performance, of the recordings. The actors of the database utterances were four male speakers annotated as DC, JK, JE, KL. The speakers who contributed for recordings were postgraduate students and researchers at the University of Surrey. The age of speakers lies between 27 to 31 years. Seven discrete categories of emotions described as anger, disgust, fear, happiness, sadness and surprise were recorded [37]. The focused research was carried out on recognizing the discrete emotions [38]. Table 2 compares the features of both the datasets used in experimental analysis.

Table 1

EMODB Labels

Letter	Emotion(German)	Emotion (English)
W	Ärger (Wut)	Anger
L	Langeweile	Boredom
E	Ekel	Disgust
A	Angst	Fear/Anxiety
F	Freude	Happiness
T	Trauer	Sadness
N	Neutral	

Table 2

Comparison of EMODB and SAVEE

Attributes	EMODB	SAVEE
No. of speakers	111	4
Age of Speakers	21 to 35 years	27 to 31 years.
No. of utterances	500+	480
Language	German	British English
Emotions	angry, happy, anxious, fearful, bored, disgusted, neutral	anger, disgust, fear, sadness, happiness, surprise, neutral

Feature Extraction [39]–[41]. The core step for recognizing speech or emotions from speech is extracting the features of speech. The process of feature extraction refers to identifying the components of the vocals from the audio signal. The audio signal is good source of linguistic information if the noise is discarded in the signal. There is another interpretation of feature extraction which says that it is characterization and recognition of information specifically related to the actor's (speaker) mood, age, gender. The general process of feature extraction involves transformation of raw signal into feature vectors, which suppress the redundancies and emphasize on the speaker specific properties. The properties are like pitch, amplitude, frequency. The speaker dependencies such as health, voice tone, speech rate and acoustical noise variations may vary the speech signal during the testing and training sessions due to [31], [42], [43]. The shape of the vocal tract filters the sounds generated by human beings which if determined efficiently can be used to derive phoneme representation of the speech sample with high accuracy.

The features to be extracted from speech can be studied under three categories named as High level Features which may include phones, lexicon, accent, pronunciation; Prosodic and Spectra-temporal Features that can be studied as pitch energy duration, rhythm, temporal features) and Short term spectral and prosodic Features pertaining to spectrum glottal pulse [44]–[46]. Short-term spectral features aid in better prediction with higher accuracies for various applications. The spectrogram analysis can be used for information extraction from the short term spectral features. Linear Predictive Cepstral Coefficients (LPCC), Mel-Frequency Discrete Wavelet Coefficient (MFDWC), Mel-Frequency Cepstral Coefficients (MFCC) are most commonly used short term spectral features for speech analysis [31], [42], [47].

MFCC. MFCC are considered as the commonly used acoustic features for the task of identifying the speaker and the properties of the speech. MFCC takes into account human perception sensitivity with respect to frequencies. The combination of both is best for speech identification and differentiation. The importance of MFCC is inspired by the fact that the shape of vocal tract that includes tongue, teeth, throat etc. filters the sound generated by human speakers. The accuracy in determining the shape enables easy analysis of the sound that comes out through the vocal tract. The accurate in determination of the shape of sound can help in finding the phonetic information. The task of MFCC is to accurately represent envelop of the short time power spectrum of the sound when it traverse through the vocal tract [41], [47], [48].

Mel Frequency Cepstral Coefficients (MFCCs) were identified as a feature and is widely applicable to automatic speech recognition and identification of speaker. The correlation among the actual and heard signal frequencies can be derived efficiently by incorporating Mel scale. Davis and Mermelstein were pioneer in identifying MFCC as sound feature in the 1980's. MFCC ever since its discovery has been considered important feature for analysis of speech signals. There are few other features along with MFCC, like Linear Prediction Coefficients (LPCs) and Linear Prediction Cepstral Coefficients (LPCCs) that were coined before MFCC and remained the main features for automatic speech recognition (ASR), especially with classification algorithm such as HMM [49]. In practice 8 to 12 or 13 MFCC are considered for representing the shape of spectrum and hence are used for speech analysis [50]. MFCC are highly preferred choices in automatic speech recognition systems [51]. Authors found that MFCC is effective for end to end acoustic modelling using CNN [52], [53]. MFCC is widely used feature while considering speech modelling [54], [55]. MFCC based comparative study of speech recognition techniques was conducted by authors who found that MFCC with HMM gave recognition accuracy of 85 percent and with deep neural networks the score was 82.2 percent [56]. Computation of MFCCs includes a conversion of the Fourier coefficients to Mel-scale [57]. Mel-scale are the most popular variant used today, even if there is no theoretical reason that the Mel-scale is superior to the other scales [58].

Decision Tree Classifiers. There are several machine learning algorithms that can be applied for recognizing speech emotions. The algorithms can be used independently or in hybrid mode for classifying emotions. Decision tree are one of the machine learning algorithms that can be used for classification task [59], [60]. The Decision tree uses the supervised learning approach that works on labelled data. The data is split into train and test subsets for carrying out the classification task. The current work uses Random forest, KNN and XGBoost algorithms for classifying emotions. All the mentioned algorithms are the variations of decision tree classifiers and a brief description of each classifier is given below [14].

Random Forest. One of the most flexible and easily implementable learning algorithm in machine learning is Random Forest. The algorithm provides better solutions over basic decision trees. The random forest depends on few parameters which if tuned can provide good results. The algorithm is widely used due to its simple and flexible aspect of implementation. Random Forest supports both regression and classification task while modelling a solution. It is supervised learning technique that creates random forests. The ensemble decision trees are referred to forest and mostly use bagging for training [51, 52]. The importance of regressive bagging lies in the fact that it increases the overall results. Multiple decision trees are build and together to increase efficiency in random forest algorithm.

Random forest generation uses same hyper parameters that are used for decision tree or a bagging classifier. The class of classifier does not require combining the decision trees to bagging classification algorithm. The algorithm proceeds by searching for the best feature from available features subset. The selected feature will then be used for splitting the node. The node split diversifies and enhances the results. The relative importance of each feature is measured while prediction. SK-learn tool can be used to measure a feature importance. The tool reduces impurity at the tree nodes that use the feature, across all trees in forest. Score for each feature is automatically computed after training. Features and observations are randomly selected by random forest and averaged for building several decision trees. The decision uses rules and facts for decision making and trees from over fitting. Random forest prevents it by creating subsets and combining them to subtrees. The only limitation of random forest is slow computation which is affected by number of trees build by random forest [61].

XGBOOST. XGBoost [61] uses gradient boosting technique to ensemble decision trees. XGBoost is stands for “eXtreme Gradient Boosting”. Small, medium structured and tabular data uses XGBoost for classification. XGBoost is studied as improvisation upon the base GBM framework. Optimization and algorithmic techniques are used to improve the base framework of GBM. Regularization is used to enhance the performance of algorithm by preventing data overfitting. The algorithm automatically learns best missing values depending on the training loss and handle variety of patterns of sparsity more efficiently. It also has built-in cross-validation method at each iteration.

XGboost is sequential tree building algorithm implemented by parallelization. The interchangeable nature of the loops determine the base of building algorithm. The external loop is responsible for maintaining the tree count, and features are calculated by the internal loop. Loops are interchangeable and thus enhance run time performance. All the instances are globally scanned, initialized and sorting is done using parallel threads. This switch of loops increase the algorithmic performance. The parallelization overheads in computation are offset. The tree splitting within GBM framework for stopping the split is greedy in nature. Splitting of tree at node depends on the negative loss criterion at the point of split. XGBoost uses ‘max_depth’ parameter as specified instead of criterion first, and pruning of the trees is done backward. The computational performance is improved significantly by using this ‘depth-first’ approach improves [61].

Sklern KNeighbors classifier and KNN. The early description of KNN was found in 1950. KNN is labour intensive approach for large datasets. It was used for pattern recognition initially. The learning of KNN is based on the comparison test data with train data such that both have similarities. A set of N attributes describe the tuple data. An n- dimensional space is used to store all the training tuples where each of them corresponds to a point in space. The pattern space for k training tuples that are closest to unknown tuple is identified by the K-nearest neighbor classifier. The closest found points are referred to nearest neighbors and euclidian distance defines the nearness of the neighboring clones [62].

The Euclidean distance between two Co tuples represented by $\mathcal{A}_1 := \{a_{11}, a_{12}, \dots, a_{1n}\}$, $\mathcal{A}_2 := \{a_{21}, a_{22}, \dots, a_{2n}\}$ is obtained using following calculation,

$$d(A_1, A_2) = \sqrt{\sum_{i=1}^n (a_{1i} - a_{2i})^2}, \quad (1)$$

And in case of parametrized KNN we use in particular case the following generalization of formula (1):

$$d(A_1, A_2(y)) = \sqrt{\sum_{i=1}^n w_i (a_{1i} - a_{2i})^2},$$

After we choose a class y maximizing this distance. Weights depend on the neighbor’s number $w(x_i) = w(i)$. For brevity, we denote these quantities by w_i . In general case we utilize some realizations of classification by formula

$$a(x) = \arg \max_{y \in Y} [x(i) = y] w_i, \quad (2)$$

where $x(i)$ are points near testing point y and Y is assumed by us class of object y .

Thus the difference of values of attribute in A_1 and A_2 is obtained. The difference is then squared to accumulate total distance count. Attributes with large ranges can outweighs attributes within small ranges (binary attributes).

To normalize data, we will use Z-scaling based on the mean value and standard deviation; dividing the difference between the variable and the mean by the standard deviation. In practice, minimax scaler and Z-scaling have similar applicability and are often interchangeable. However, when calculating the distances between points or vectors, Z-scaling is used in most cases. And the minimax is useful for visualization, for example, to transfer the features encoding the color of the pixel into a range of [0... 255] [2].

Recall that Z-scaling based on the mean and standard deviation is dividing the difference between variable and mean by standard deviation;

$$z = \frac{x - \mu}{\sigma}, \quad (3)$$

where μ is an expected value and σ is the standard deviation of the value.

Because the K-Nearest Neighbours Algorithm (hereinafter the KNN algorithm) is about the distance from a point to a class, Z-scaling is usually used for its application, as it is known that in calculating the distances between points or vectors in most cases the result of Z- scaling is much more accurate.

The main advantage of Z-scaling is that it preserves the normal distribution of a random value.

Z-normalization keeps a distribution normal if it was, and a non-normal distribution converts to a non-normal distribution too. As a result of such a transformation, we get a value with 0-th mean (mean) and 1 dispersion.

Normalization is applied to each attribute value to resolve the issue.

Most common class is assigned to unknown tuple among its k-nearest neighbours. If the value of K equates 1, the unknown tuple is assigned the class of the training tuple that is closest to it in pattern space. Real value prediction are returned by KNN for unknown value tuples. The unknown values the classifier of KNN returns the average of the real valued labels associated with K-nearest neighbours of unknown tuple [62]–[64].

Min-max scaler keeps outliers so we have to use robust scaler of statistics that are robust to outliers. As an alternative normalization we propose to use Z-normalization (Z-scaler). This normalization holds a normal distribution.

Problem Statement. Authoritative literary sources mention MFCC as an important feature for analyzing and classifying various aspects of speech. Some of them state that only 13 MFCC features are sufficient enough to be considered for experimentation. There is currently no experimental validation for this statement. Moreover, there is no sufficient research on the identification of these 12 MFCC from the extracted 20 base features of Mel scale. MFCC also have derivatives of base features named as delta and double delta. The aim of the work is to establish experimental proof of considering only 8 to 13 MFCC from extracted 39 features of MFCC [23], [50]. The current work conducts experimental analysis on MFCC obtained from EMODB, a Berlin database that consists of more than 500 utterances, which were recorded from 111 both male and female speakers from various age groups.

Experimental confirmation of results. To reach more effectiveness we utilize parametric KNN method. To present a number row, you have to look at the dynamics of the signal change. In large sentences, emotions are placed either at the end of the sentence or at the beginning. Therefore, when parameterizing the distance vector, it is important to set weighting factors in such a way as to distinguish the significance of the start of the sentence distance and the distance to the end of the vector coordinate.

When applying discriminant recognition techniques to the object, the feature vector is displayed in five-dimensional space. The training matrix will be defined as the data matrix:

$$W = \begin{pmatrix} w_{11} & w_{12} & w_{13} & w_{14} & w_{15} & w_{16} \\ w_{21} & w_{22} & w_{23} & w_{24} & w_{25} & w_{26} \\ w_{31} & w_{32} & w_{33} & w_{34} & w_{35} & w_{36} \\ w_{41} & w_{42} & w_{43} & w_{44} & w_{45} & w_{46} \\ w_{51} & w_{52} & w_{53} & w_{54} & w_{55} & w_{56} \end{pmatrix}.$$

We highlight five characteristics: in addition to the three mentioned in the beginning, we will highlight the beginning of the sentence and its end, as they are emotionally loaded. They carry not only an information load, but also an emotional one. When analyzing the beginning and the end of a sentence, we will take into account how much each feature has changed relative to the average characteristics of the speaker. Therefore, forming a data-vector with 5 coordinate we substitute in the i -th coordinate characterizing the amplitude the ratio $\frac{a_i}{M(a)} = v_i$ of the amplitude of the current phrase a_i to the average value $M(a)$ of the amplitude of the person's voice.

The columns of the matrix W contain the weighting factors [66]–[67, 71] that most characterize this class of emotion, and the rows contain the features extracted from the phrase. For example, high amplitude is most characteristic for the emotion of aggression. The corresponding weight coefficient in the aggression column will be larger. Knowing the average values of the features of speech, we construct the matrix based on the changes in the parameters.

The following simulations and experiments were performed. The flow graph in Fig 2 shows the steps carried out for the experiment.

The experiment was done on MFCC extracted from the EMODB and SAVEE data set. Four subsets of MFCC features comprising 20 cepstral constants were analyzed for feature importance. This was done to identify which 13 MFCC should be used for speech emotion analysis. The result of each subset using different classifiers areas shown in Table 3. Each subset was used to classify emotions using supervised learning algorithm (variants of decision trees). It was observed that the results obtained using 20 MFCC over set of first 13 was very near to each other. There was no effective and substantial difference in the accuracy scores for classification while using 13 and 20 MFCC subsets. Increased number of features often increase the complexity of the system and so if 13 MFCC are used instead of 19 MFCC the results will not suffer much loss. The validation of the experiment was also done by extracting important features from PCA Analysis. It was seen that most of the important features corresponded to initial 13 MFCC extracted from dataset. The accuracy score of classification on EMODB using M0-M12 was 52%, 50%, 47%, 41% using subsets M0-M19, M10-M12, M15-M17 and M16-M19 respectively using Random Forest classifier. KNN shows 56%, 44%, 45% and 42% accuracy score using subsets M0-M19, M10-M12, M15-M17 and M16-M19 respectively. XGB showed poor performance on the original extracted dataset for classification task. SAVEE results as shown in Table 4 depicts accuracy scores of RF using subsets M0-M19, M10-M12, M15-M17 and M16-M19 are 57%, 55%, 50%, 50% respectively. For KNN the performance of four subsets is 67%, 54%, 55%, 50%. XGB showed poor performance on all the subsets. The results obtained in Table 4 and Table 5 clearly indicate that selecting M0-M12 would be a better choice for features from MFCC data set. It can be seen from the results that first 13 Mel coefficients can successfully be used for playing with speech over using 20 features. This selection shall only optimize the results but also reduce the complexity of the model thereby reducing the computation time.

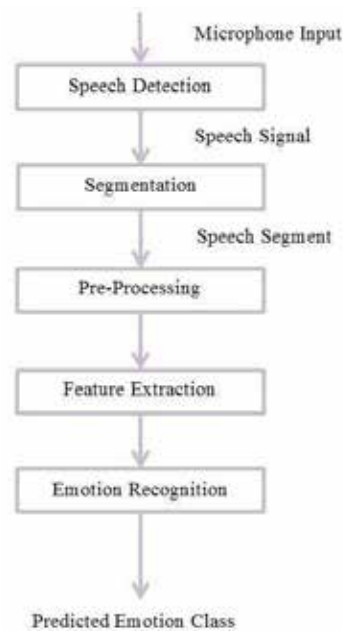


Fig 2. Flowgraph of the emotion classification using decision tree classifiers

Table 3

Label encoded emotions for EMODB

Emotion EMODB	Emotion SAVEE	Encoded label
Fear/Anxiety	Anger	0
Disgust	Disgust	1
Happiness	Fear	2
Boredom	Happiness	3
Neutral	Sadness	4
Sadness	Surprise	5

There after the dataset was minimized and re-experimented for feature importance and classification. The accuracy of results for classification increased effectively but the set of important features still contained features

M0 to M12 that initial 13 features. The reason behind this is that as the sound signal passes through the vocal tract and comes out as the utterance there is a subsequent addition of noise to the originally generated signal. Addition of noise disturbs the energy whose log is computed as the base of MFCC extraction. The induction of noise imputes the signal at later levels so the original signal remains intact for usage in analysis [65]. The features present here can be successfully used for better results as compared those extracted towards the end of the sample of each speech signal.

Table 4

Results of EMODB with subsets of MFCC

MFCC	M0-M19	M0-M12	M15-M17	M6-M19
RF	52%	50%	47%	41%
KNN	56%	44%	45%	42%
XGB	40%	38%	28%	27%

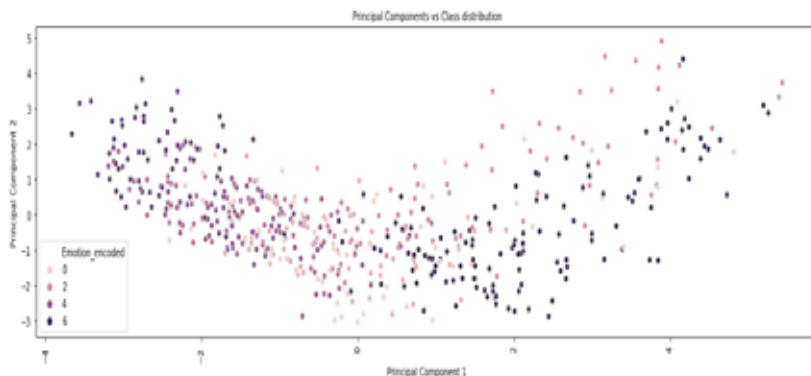
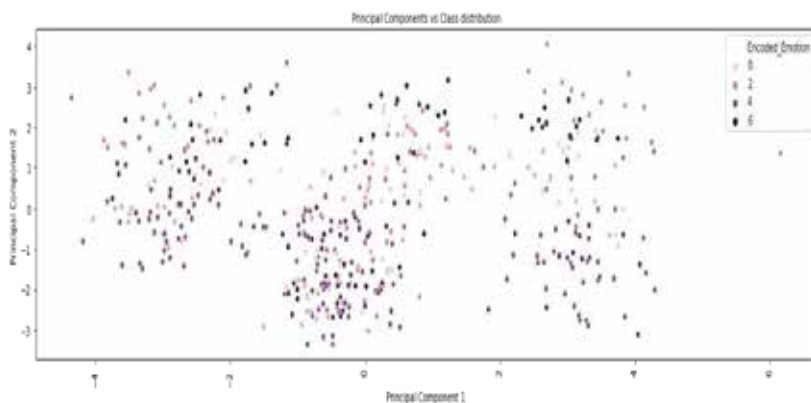
Table 5

Results of SAVEE with subsets of MFCC

MFCC	M0-M19	M0-M12	M15-M17	M6-M19
RF	57%	55%	50%	50%
KNN	67%	64%	55%	50%
XGB	30%	38%	30%	30%

The results in table 4 and Table 5 shows that M0-M19 and M0-M13 has nearly similar results for accuracy on classifiers. The datasets in Table 3 and Table 4 used the non-manipulated MFCC extracted from the speech utterances in in EMODB and SAVEE.

For further analysis the datasets were minimized and preprocessed using min max scaling. The important features identified for the minimized data using principle component analysis are in Fig 5 and Fig6. The x-axis of the plot shows various classes of emotion and y axis plot shows the MFCC features.

**Fig. 5. Principle component Vs Class Distribution for EMODB****Fig. 6. Principle component Vs Class Distribution for SAVEE**

The results of using 13 MFCC from minimized datasets EMODB and SAVEE are shown in Fig. 6(a), 6(b), 6(c). SAVEE results can be seen in Fig. 8(a), 8(b), 8(c). The plots in the mentioned figures clearly displays the precision. Recall and F1- scores for emotions in both the datasets using variants of decision trees.

Results showed that for EMODB all the three classifiers defined fairly variable results. Boredom (91%), anger (74%) and sadness (100%) had highest precision for random forest XGB and KNN respectively in EMODB. For SAVEE a higher precision rate for Disgust (84%) was identified using random forest. KNN and XGBoost identified sadness more precisely over remaining six emotions where the scores for sadness were 84%, 70% and 74% with KNN, random forest and XGBoost respectively. A common conclusion was obtained from the results of both the datasets that sadness was commonly identified with highest precision using KNN. So KNN can be effective for studying the emotion sadness in emotion analysis.

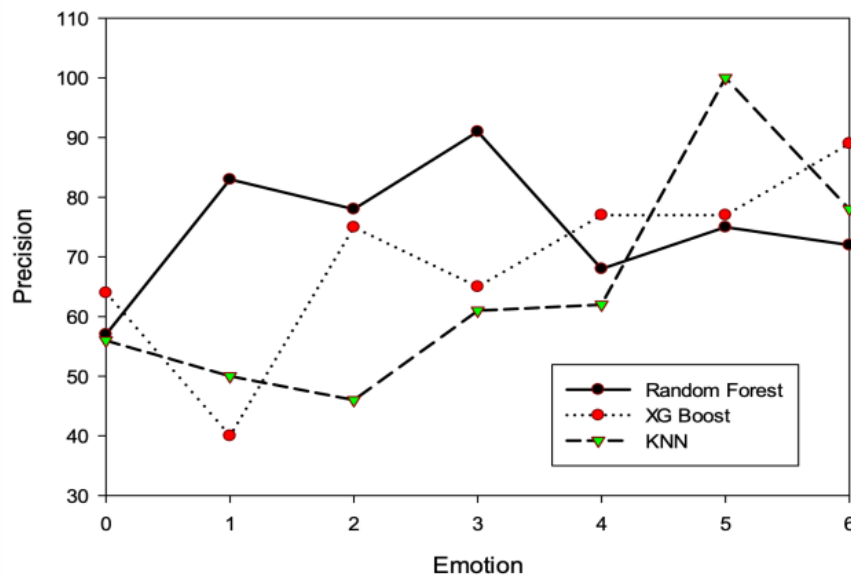


Fig. 6 (a). Emotion Vs Precision score for first 13 MFCC using KNN, Random Forest and XGB classifiers on EMODB

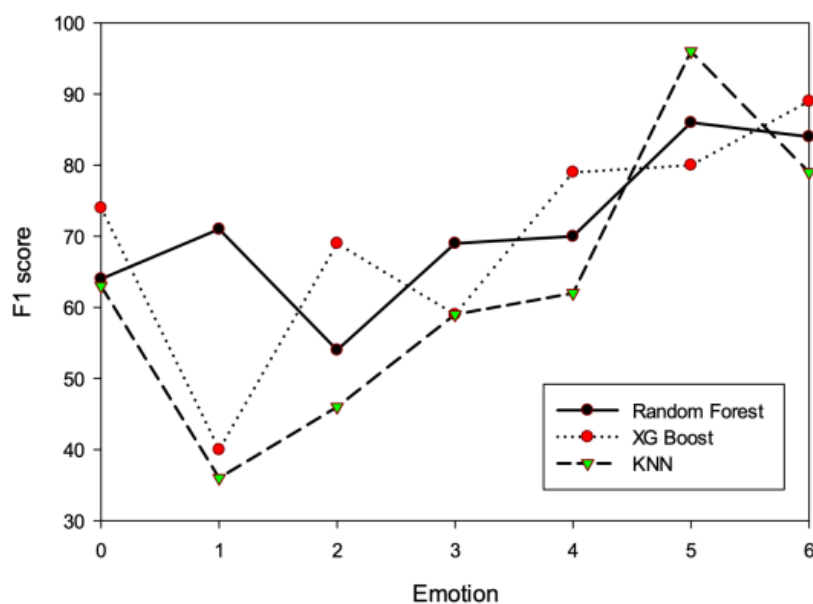


Fig. 6 (b). Emotion Vs F1-score for first 13 MFCC using KNN, Random Forest and XGB classifiers on EMODB

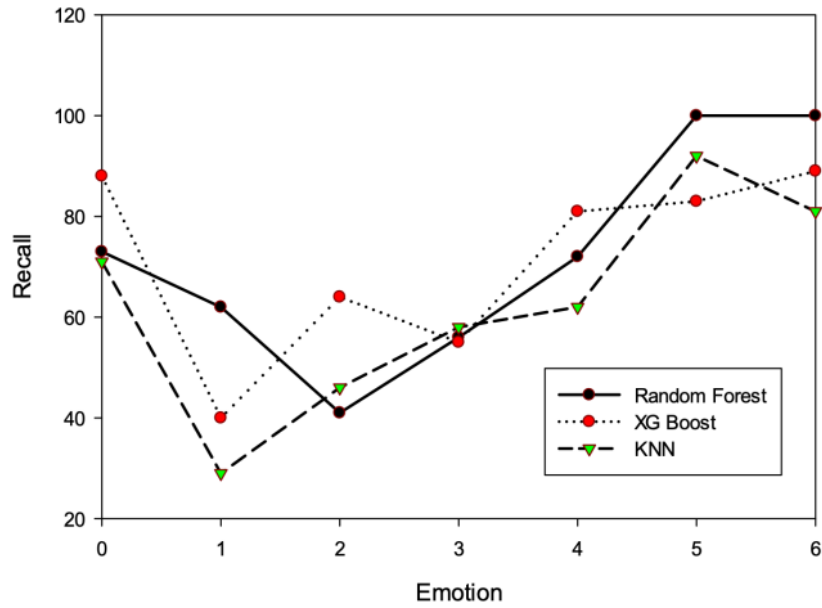


Fig. 6 (c). Emotion Vs Recall score for first 13 MFCC KNN, Random Forest and XGB classifiers on EMODB

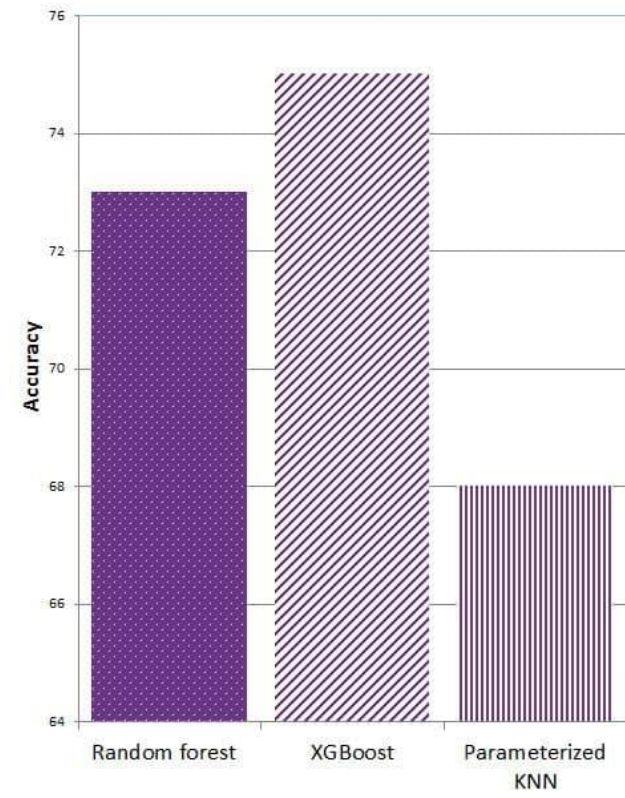


Fig. 7 (a). Accuracy score for Emotion classification using parameterized KNN, Random Forest and XGB classifiers on EMODB

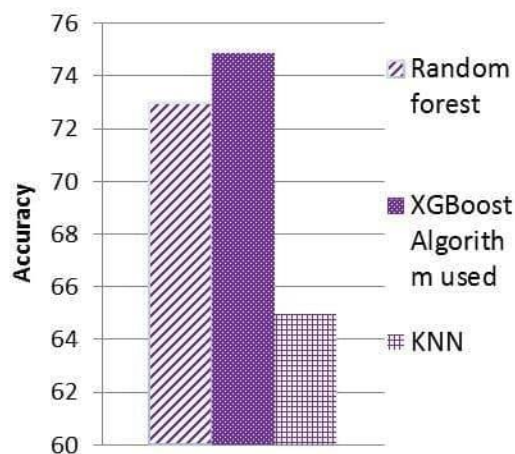


Fig. 7 (b). Accuracy score for Emotion classification using parameterized KNN, Random Forest and XGB classifiers on EMODB

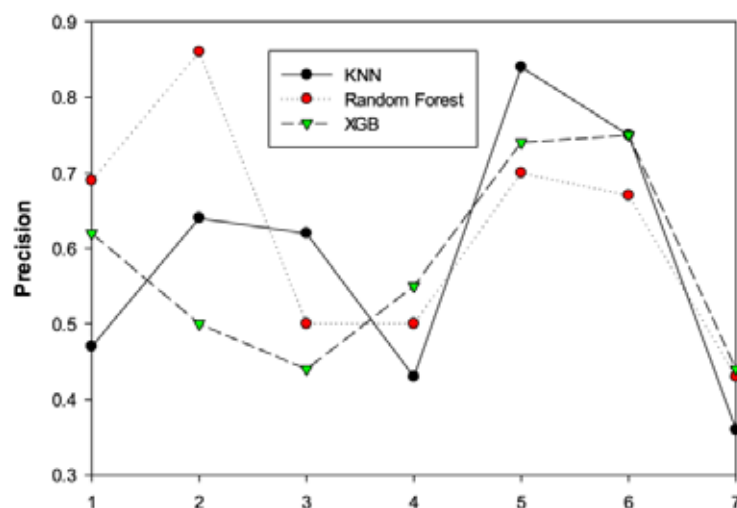


Fig. 8 (a). Emotion Vs Precision score for first 13 MFCC using KNN, Random Forest and XGB on SAVEE

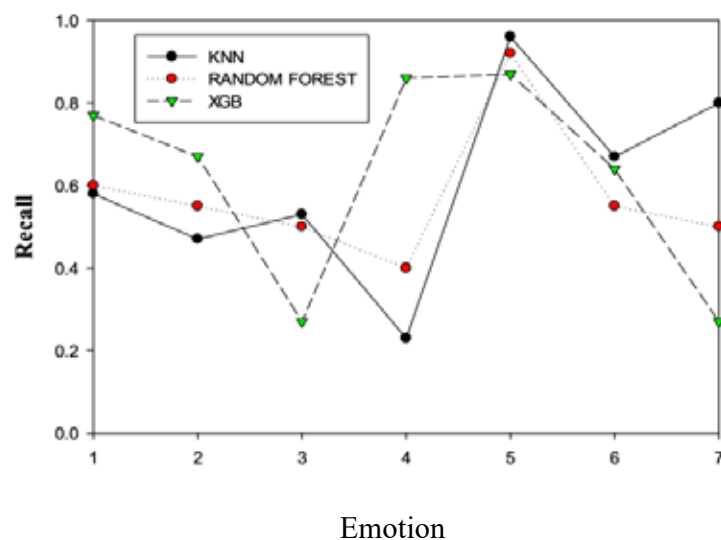


Fig. 8 (b). Emotion Vs Recall score for first 13 MFCC using KNN, Random Forest and XGB on SAVEE

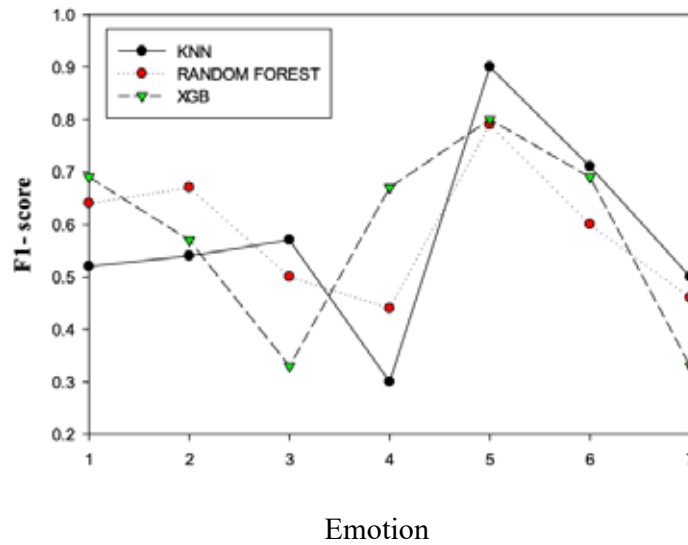


Fig.8 (c). Emotion Vs F1 score for first 13 MFCC KNN, Random Forest and XGB on SAVEE

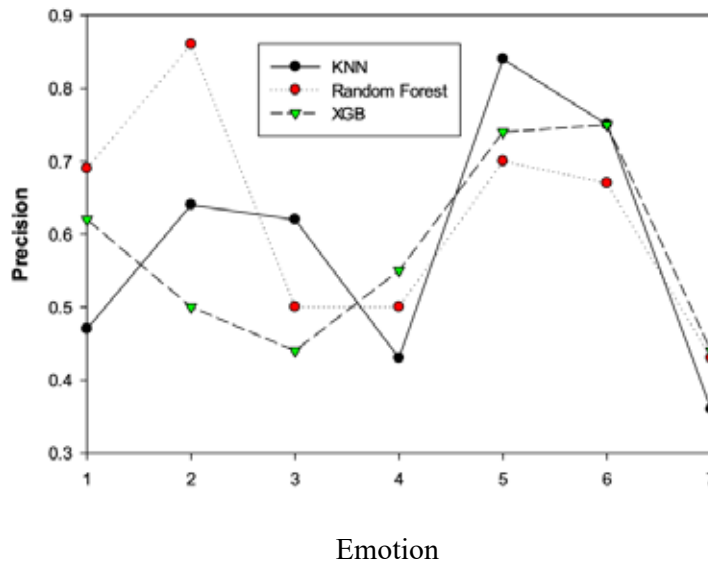


Fig. 8 (d). Emotion Vs Recall score for 13 MFCC using three decision tree classifiers on EMODB

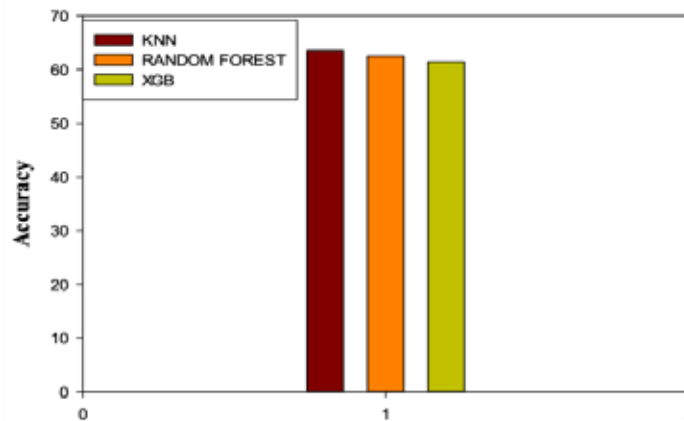


Fig. 9. Accuracy score obtained with first 13 MFCC using KNN, Random Forest and XGB

Conclusion. Speech features have always remained the one of the regressively studied topic in research. Speech or utterances contain vital information regarding the intention, emotion and psychology of the speaker [70]. The paper studied the use of one such speech feature called MFCC and utilized it to classify emotions using two datasets. The work also tried to establish the importance of using first 13 MFCC when we have a set of 20 Mel constants that can be extracted for speech based on the vocal physiology of human mouth. Accuracy scores for emotion classification using variants of decision tree approach has been obtained for EMODB and SAVEE for two datasets. KNN was identified as the common classification algorithm for both datasets. The score of sadness as obtained from KNN were highest for both the datasets. The results of the experiments can be utilized for predicting emotions and personality of the speaker. The results can be integrated with various application pertaining to human psychology and medical treatments.

References

1. Koolagudi, S.G., Rao K.S. (2012). Emotion recognition from speech: a review. *International Journal of Speech Technology*, no. 15, pp. 99–117.
2. Marechal, C. et al. (2019). *Survey on AI-based multimodal methods for emotion detection*. High-Performance Modelling and Simulation for Big Data Applications, Springer.
3. Rao, K.S., Koolagudi, S.G., Vempada, R.R. (2013). Emotion recognition from speech using global and local prosodic features. *International Journal of Speech Technology*, vol. 16, no. 2, pp. 143–160.
4. Koolagudi, S.G., Barthwal, A., Devliyal, S., Rao, K.S. (2012). Real life emotion classification from speech using gaussian mixture models. *Communications in Computer and Information Science*, no. 38, pp. 3892–3899.
5. S. Latif, R. Rana, S. Younis, Qadir, J., Epps, J. (2018). Transfer learning for improving speech emotion classification accuracy. *Interspeech: Proceedings of the Annual Conference of the International Speech Communication Association*, pp. 257–261.
6. Lee, C.M., Narayanan, S.S. (2005). Toward detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 2, pp. 293–303.
7. Banse, R. Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, vol. 70, no. 3, pp. 614–636.
8. Hozjan, V., Kačič, Z. (2003). Context-independent multilingual emotion recognition from speech signals. *International Journal of Speech Technology*, no. 6, pp. 311–320.
9. Ramakrishnan, S. (2012). Recognition of Emotion from Speech: A Review. *Speech Enhancement, Modeling and Recognition-Algorithms and Applications*, pp. 121–138.
10. Sebe, N. Cohen, I., Huang, T.S (2005). Multimodal emotion recognition. *Handbook of Pattern Recognition and Computer Vision*, 3rd Edition.
11. Zhang, Q., Wang, Y., Wang, L., Wang, G. (2007). Research on speech emotion recognition in E-learning by using neural networks method. *IEEE International Conference on Control and Automation, ICCA*, vol. 167, pp. 114–177.
12. Jing, S., Mao, X., Chen, L. (2018). Prominence features: Effective emotional features for speech emotion recognition. *Digital Signal Processing: A Review Journal*, vol. 72, no. pp. 216–231.
13. Albornoz, E.M., Milone, D.H., Rufiner, H.L. (2011). *Spoken emotion recognition using hierarchical classifiers*. Computer Speech and Language.
14. A. Özseven, T. Düğenci, M., Durmuşoğlu (2018). A Content Analysis of The Research Approaches in Speech Emotion. *International Journal of Engineering Sciences & Research Technology*.
15. Krishna Kishore K.V, Krishna Satish P. (2013). Emotion recognition in speech using MFCC and wavelet features. *Proceedings of the 2013 3rd IEEE International Advance Computing Conference, IACC*.
16. Yousefpour, A., Ibrahim, R., Hamed, H.N.A. (2017). Ordinal-based and frequency-based integration of feature selection methods for sentiment analysis. *Expert Systems with Applications*, no. 75, pp. 80–93.
17. Shu L. et al. (2018). A review of emotion recognition using physiological signals. *Sensors (Switzerland)*.
18. Oosterwijk, S., Lindquist, K.A., Anderson, E., Dautoff, R., Moriguchi, Y., Barrett, L.F. (2012). States of mind: Emotions, body feelings, and thoughts share distributed neural networks. *NeuroImage*.
19. Pessoa, L. (2010). Emotion and cognition and the amygdala: From “what is it?” to “what’s to be done?”. *Reprinted from Neuropsychologia*, vol.48.
20. Koolagudi, S.G., Rao, K.S. (2012). Emotion recognition from speech: A review. *International Journal of Speech Technology*, no. 15, pp. 99–117.
21. Winkelman, P., Niedenthal, P., Wielgosz, J., Eelen, J., Kavanagh, L.C. (2014). Embodiment of cognition and emotion. *APA handbook of personality and social psychology*, vol. 1, pp.151–175.
22. Fernández-Caballero, A. et al. (2016). Smart environment architecture for emotion detection and regulation. *Journal of Biomedical Informatics*.
23. Guan, H., Liu, Z., Wang, L., Dang, J., Ruiguo, Yu. (2018). Speech Emotion Recognition Considering Local Dynamic Features. *Computer Science*.
24. Cen, L., Zhu, L.Y., Wu, F., Yu, Z.L. Hu F.A. (2016). Real-Time Speech Emotion Recognition System and its Application in Online Learning. *Emotions, Technology, Design, and Learning*.
25. Shuman, V., Scherer, K.R. (2015). Emotions, Psychological Structure of. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*.
26. Ekman, P. (2005). Basic Emotions. *Handbook of Cognition and Emotion*.

27. Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, no. 24(8), pp. 1377–1388.
28. Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, no. 48(4), pp. 384–392.
29. Bourke, C., Douglas, K., Porter, R. (2010). Processing of facial emotion expression in major depression: A review. *Australian and New Zealand Journal of Psychiatry*. 2010, no. 44(8), pp. 681–696.
30. Van den Stock, J., Righart, R., B. de Gelder, (2007). Body Expressions Influence Recognition of Emotions in the Face and Voice. *Emotion*.
31. Banse, R., Scherer, K.R. (1996). Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology*.
32. Gulzar, T., Singh, A., Sharma, S. (2014). Comparative Analysis of LPCC, MFCC and BFCC for the Recognition of Hindi Words using Artificial Neural Networks. *International Journal of Computer Applications*.
33. Shrawankar, U., Thakare, V.M. (2013). Techniques for Feature Extraction In Speech Recognition System : A Comparative Study. *International Journal Of Computer Applications In Engineering, Technology and Sciences*.
34. Haamer, R.E., Rusadze, E., Lüsi, I., Ahmed, T., Escalera, S., Anbarjafari, G. (2018). Review on Emotion Recognition Databases. *Human-Robot Interaction – Theory and Application*.
35. Lalitha, S., Geyasruti, D., Narayanan, R., Shravani, M. (2015). Emotion Detection Using MFCC and Cepstrum Features. *Procedia Computer Science*, vol. 70, pp. 29–35.
36. Jackson, P., Haq, S. (2014). *Surrey audio-visual expressed emotion (savee) database*. University of Surrey: Guildford.
37. Liu, Z.T., Xie, Q., Wu, M., Cao, W.H., Mei, Y., Mao, J. W. (2018). Speech emotion recognition based on an improved brain emotion learning model. *Neurocomputing*.
38. Ekman, P. et al. (1987). Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion. *Journal of Personality and Social Psychology*.
39. Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S. (2009). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
40. Koduru, A., Valiveti, H.B., Budati, A.K. (2020). Feature extraction algorithms to improve the speech emotion recognition rate. *International Journal of Speech Technology*.
41. Kumar, K., Kim, C., Stern, R.M. (2011). Delta-spectral cepstral coefficients for robust speech recognition. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings*.
42. Tiwari, V. (2010). MFCC and its applications in speaker recognition. *International Journal on Emerging Technologies*.
43. Dave, N. (2013). Feature Extraction Methods LPC, PLP and MFCC In Speech Recognition. *International Journal for Advance Research in Engineering and Technology*.
44. Yankayi, M. (2016). Feature Extraction Mel Frequency Cepstral Coefficients (Mfcc), pp. 1–6.
45. Ananthakrishnan, S., Narayanan, S.S. (2008). Automatic prosodic event detection using acoustic, lexical, and syntactic evidence. *IEEE Transactions on Audio, Speech and Language Processing*.
46. Kinnunen, T., Li, H. (2010). An overview of text-independent speaker recognition: From features to supervectors. *Speech Communication*.
47. Wang, W.Y., Biadys, F., Rosenberg, A., Hirschberg, J. (2013). Automatic detection of speaker state: Lexical, prosodic, and phonetic approaches to level-of-interest and intoxication classification. *Computer Speech and Language*.
48. Lyons, J. (2014). Mel Frequency Cepstral Coefficient. *Practical Cryptography*.
49. Palo, H.K., Chandra, M., Mohanty, M.N. (2018). Recognition of Human Speech Emotion Using Variants of Mel-Frequency Cepstral Coefficients. *Lecture Notes in Electrical Engineering*, vol. 442, pp. 491–498.
50. Yazici, M., Basurra, S., Gaber, M. (2018). Edge Machine Learning: Enabling Smart Internet of Things Applications. *Big Data and Cognitive Computing*.
51. Wang, Xia, Dong, Yuan, Hakkinen, J., Viikki, O. (2002). Noise robust Chinese speech recognition using feature vector normalization and higher-order cepstral coefficients.
52. Davis, S.B., Mermelstein, P. (1990). Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. *Readings in Speech Recognition*.
53. Palaz, D., Magimai-Doss, M., Collobert, R. (2019). End-to-end acoustic modeling using convolutional neural networks for HMM-based automatic speech recognition. *Speech Communication*.
54. Passricha, V., Aggarwal, R.K. (2020). A comparative analysis of pooling strategies for convolutional neural network based Hindi ASR. *Journal of Ambient Intelligence and Humanized Computing*.
55. Vimala, C., Radha, V. (2014). Suitable Feature Extraction and Speech Recognition Technique for Isolated Tamil Spoken Words. *International Journal of Computer Science and Information Technologies*.
56. Dalmiya, C.P., Dharun, V.S., Rajesh, K.P. (2013). An efficient method for Tamil speech recognition using MFCC and DTW for mobile applications. *2013 IEEE Conference on Information and Communication Technologies, ICT*.
57. NithyaKalyani, A., Jothilakshmi, S. (2019). Speech summarization for tamil language. *Intelligent Speech Signal Processing*.
58. Stevens, S.S., Volkman, J., Newman, E.B. (1937). A Scale for the Measurement of the Psychological Magnitude Pitch. *Journal of the Acoustical Society of America*.
59. Mitrović, D., Zeppelzauer, M., Breiteneder, C. (2010). *Features for Content-Based Audio Retrieval*.
60. Caruana, R., Niculescu-Mizil, A. (2006). An empirical comparison of supervised learning algorithms. *ACM International Conference Proceeding Series*.
61. Kotsiantis, S.B. (2007). Supervised machine learning: A review of classification techniques. *Informatica (Ljubljana)*.
62. Luckner, M., Topolski, B., Mazurek, M. (2017). Application of XGboost algorithm in fingerprinting localisation task. *Computer Science*.

63. Sutton, O. (2012). Introduction to k Nearest Neighbour Classification and Condensed Nearest Neighbour Data Reduction. *Introduction to k Nearest Neighbour Classification*
64. Deng, Z., Zhu, X., Cheng, D., Zong, M., Zhang, S. (2016). Efficient kNN classification algorithm for big data. *Neurocomputing*.
65. Okfalisa, Gazalba, I., Mustakim, Reza, N.G.I. (2018). Comparative analysis of k-nearest neighbor and modified k-nearest neighbor algorithm for data classification. *2nd International Conferences on Information Technology, Information Systems and Electrical Engineering, ICITISEE*.
66. Skuratovskii, R.V. (2020). The timer compression of data and information Proceedings. *IEEE 3rd International Conference on Data Stream Mining and Processing (DSMP)*, pp. 455–459.
67. Skuratovskii, R.V. (2019). Employment of Minimal Generating Sets and Structure of Sylow 2-Subgroups Alternating Groups in Block Ciphers. *Advances in Computer Communication and Computational Sciences, Springer*, pp. 351–364.
68. Gnatyuk, V.A. (2001). Mechanism of laser damage of transparent semiconductors. *Physica B: Condensed Matter*, pp. 308–310.
69. Zgurovsky, M.Z., Pankratova, N.D. (2007). *System Analysis: Theory and Applications*. Springer Verlag. Berlin.
70. Romanenko, Y.O. (2016). Place and role of communication in public policy. *Actual Problems of Economics*, vol. 176, no. 2, pp. 25–26.
71. Skuratovskii, R.V. (2021). On commutator subgroups of Sylow 2-subgroups of the alternating group, and the commutator width in wreath products. *European Journal of Mathematics*, vol. 7, pp. 353–373.

UDC 619.12

DOI <https://doi.org/10.32689/maup.it.2021.2.8>

Ruslan SKURATOVSKIY

Lecturer at the Department of Computational Mathematics and Computer Modeling, Interregional Academy of personnel management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039

ORCID: <https://orcid.org/0000-0002-5692-6123>

Руслан СКУРАТОВСЬКИЙ

викладач кафедри обчислювальної математики і комп'ютерного моделювання, ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом», вул. Фрометівська 2, Київ, Україна, індекс 03039

Bibliographic description of the article: Skuratovskiy, R. (2021). Analitichnyi ohliad ta systemnyi analiz filtratsii fosfororhanichnykh kyslot v ahrarnomu sektori Yevropeiskoho Soiuzu [Analytical overview and system analysis of organophosphates filtration in the agricultural sector of the European Union]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 75–84. DOI: <https://doi.org/10.32689/maup.it.2021.2.8>

Бібліографічний опис статті: Скуратовський Р. Аналітичний огляд та системний аналіз фільтрації фосфорорганічних кислот в аграрному секторі Європейського Союзу. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 75–84. DOI: <https://doi.org/10.32689/maup.it.2021.2.8>

ANALYTICAL OVERVIEW AND SYSTEM ANALYSIS OF ORGANOPHOSPHATES FILTRATION IN THE AGRICULTURAL SECTOR OF THE EUROPEAN UNION

Abstract. This research focused on reviewing organophosphate monitoring in European countries that was carried out by various researchers independently to National authorities or official European monitoring bodies, during the last ten years. A modified QuEChERS method (a solid phase extraction technique for detection of pesticide residues in food) seems to be the most preferable method used lately in private or educational laboratories, while ultra high pressure liquid chromatography (UHPLC) systems, are gradually gaining ground in trace analysis. Considering the collected published data, organophosphates seem not to be so intriguing to monitor and research, besides being still one of the most frequently applied pesticides used for cultivation purposes today. The question is, will public health be exposed to adverse effects of produce sold, especially in local markets, without the necessary “free from pesticide residue” certification? Traceability and food labeling techniques can provide evidence for certain production methods (eg organic food) or specific food product origin, while facilitate traceability in case of MRLs violation.

Key words: pesticides, organophosphates, residual analysis, agricultural sector.

АНАЛІТИЧНИЙ ОГЛЯД ТА СИСТЕМНИЙ АНАЛІЗ ФІЛЬТРАЦІЇ ФОСФОРОРГАНІЧНИХ КИСЛОТ В АГРАРНОМУ СЕКТОРІ ЄВРОПЕЙСЬКОГО СОЮЗУ

Анотація. Це дослідження було зосереджено на перегляді моніторингу фосфорорганічних кислот у європейських країнах, який проводився різними дослідниками незалежно від національних органів влади або офіційних європейських моніторингових органів протягом останніх десяти років. Модифікований метод QuEChERS (метод твердофазної екстракції для виявлення залишків пестицидів у харчових продуктах), здається, є найбільш переважним методом, який останнім часом використовується у приватних чи освітніх лабораторіях, тоді як системи рідкої хроматографії надвисокого тиску (UHPLC) поступово завойовують свою популярність. аналіз. Беручи до уваги зібрані опубліковані дані, органофосфати, здається, не настільки інтригуючі для моніторингу та досліджень, крім того, що вони досі є одним із найчастіше застосовуваних пестицидів, які використовуються для вирощування на сьогоднішній день. Питання полягає в тому, чи буде громадське здоров'я піддане негативно впливу реалізованої продукції, особливо на місцевих ринках, без необхідної сертифікації «без залишків пестицидів»? Методи відстеження та маркування харчових продуктів можуть надати докази для певних методів виробництва (наприклад, органічних харчових продуктів) або специфічного походження харчових продуктів, одночасно полегшуючи простежуваність у разі порушення МДГОВ.

Ключові слова: пестициди, органофосфати, залишковий аналіз, аграрний сектор.

Graphical Abstract



Overview. Agriculture. Agricultural products are produced throughout the world, providing the basis of human nutrition with vitamins, minerals and fibers [1], while many national economies depend on agricultural practices. Yesterday's agriculture was based on conventional cultivation methods, and focused only on mass production, without quality control schemes and insufficient control measures. These cultivation methods act like a chain, starting with producers experiencing continuous fluctuations and instability in their production and hence, their income, next the consumers not able to distinguish if a product falls within the acceptable safety region, its quality or its origin, and finally the increased risk to human health and the environment as shown from the following years of such practices [2; 3; 4].

Ways of agriculture however, have changed considerably during the last decades. The industrial takeover produced mono-cultured agricultural production in developed countries, with local economies evolving into global economies and focused more on qualitative production than quantitative. The introduction of greenhouse cultivation, mainly in the Mediterranean region, offered a "controlled environment" in order to produce specific products otherwise not possible.

The introduction of a wide variety of pesticides for all types of treatment, fertilizers for all types of deficiencies, agricultural tools and instruments and new irrigation systems, gave a new impulse in the agriculture sector, increasing gradually the agricultural income, but at the same time, introducing new or recurring problems (contaminated soil and ground water, polluted air, food-borne illness, toxic chemicals in foods, animal feed and fiber) so much in the rural exploitation, as well as in the wider environment and consequently, in the social structure [5]. One of the main reasons this occurs is due to the thoughtless and reckless use of pesticides from non-trained farm producers.

In Europe, certain legislations, regulations and directives have been established by the European Parliament focusing on placing of plant protection products on the market (Regulation (EC) No 1107/2009), maximum residue levels of pesticides in or on food and feed of plant and animal origin (Regulation (EC) No 396/2005), laying down rules for making available on the market of EU fertilizing products (Regulation (EU) 2019/1009) and bio-cidal products (Regulation (EU) No 528/2012) and establishing a framework for Community action to achieve the sustainable use of pesticides (Directive 2009/128/EC) [6; 7; 8; 9], while pesticide monitoring reports are published for all European countries every year [10]. Additionally, European Community's Rapid Alert System for Food and Feed [11], publishes weekly overviews of alerts and information notifications on its website, with chemical hazards being one the reported hazard categories. Additionally, the European Commission publishes EU statistics on pesticide sales in regards to geographical location, year, unit of measure and pesticide groups and an updated EU pesticide database that categorizes products/active substances as "approved" or "not approved".

Pesticide Usage. The basic agricultural pesticide categories, based on plant protection activity, are insecticides (insects), fungicides (fungi) and herbicides (weeds). The most common non organic insecticides applied are organochlorines, organophosphates, carbamates, neonicotinoids and synthetic pyrethroids. Although specific categories have been banned in EU countries (eg organochlorines).

Even though pesticides are designed to target specific organisms, this is not guaranteed. Chemical structures differ in both within and between these categories and their induced modes of action cause collateral damage to unwanted species. This can lead to an abrupt change in the ecosystem, un-stabilizing the natural processes of the environment and increasing the health risk to humans as the induced toxicity varies widely within each type of group [12; 13; 14; 15; 16; 17; 18].

The importance of pesticide residue exposure plays a pivotal health role for producers and their families that are exposed directly, as well as the general population that is exposed indirectly via marketing and application of pesticides, domestic vegetable cultivation, close proximity to rural areas of agriculture fields and by the consumption of exposed food and water [19; 20; 21; 22; 23; 24; 25; 26]. This has been noted especially for intense cultivated European countries (i.e. greenhouse in Spain, Portugal, Italy and Greece), where pesticides have been used at least for three decades. The lack of proper information (what pesticide for what crop, adverse health effects, safety measures etc.), false estimations (pesticide dosage per crop) and training (correct clothing and storage, application tools, how and when to apply the pesticide mixture, safe harvesting post period, disposal of the empty packages etc.) result in a silent chronic exposure detrimental to human health and in some cases of acute exposure, fatal to human health.

Organophosphates. Organophosphate pesticides was one of the most important preferred categories due to its increased activity against a wide spectrum of insects at relatively low application doses. Organophosphates and carbamates had replaced organochlorines world-wide as they were regarded safer for their toxic impact, although today, most organophosphates have started to be banned as they are now being replaced by other insecticidal products (especially synthetic pyrethroids) that have an even less toxic impact.

Organophosphate insecticides have been widely used for more than three decades in the U.S. accounting for almost 30% of the global insecticide market sectors [EPA, 2017].

Sales of insecticides based on organophosphates, in regards to the 27 geopolitical entities of the EU, of the year 2018, show the immense level of this preferred insecticide group as shown in Fig. 1.

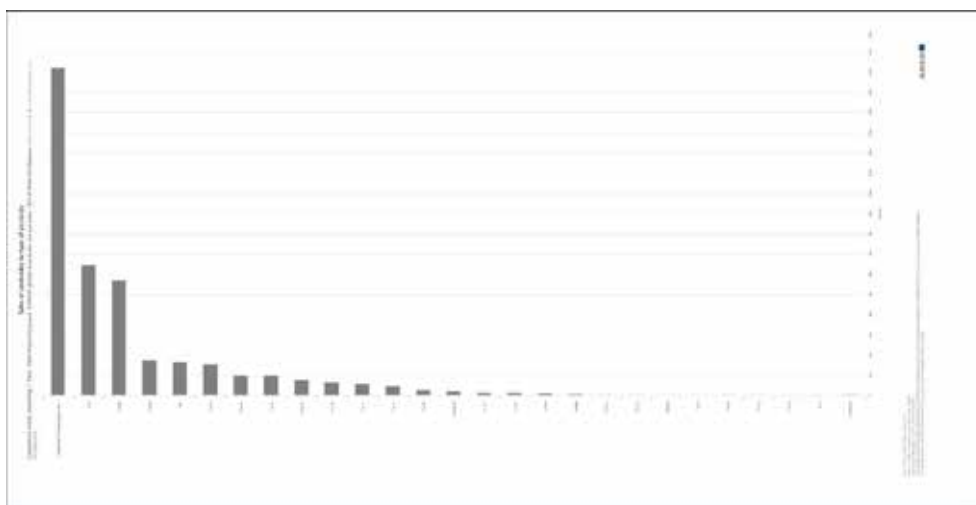


Fig. 1. Volume of pesticides sold by type of pesticide (based on organophosphates) within the current 27 EU member states for the year 2018. [Eurostat (2020). Sales of pesticides by type of pesticide. Retrieved from: <https://ec.europa.eu/eurostat/databrowser/view/tai02/default/bar?lang=en>]

A few of the most used organophosphates that are not approved based on the EU pesticide database, are Parathion, Phoxim, Chlorpyrifos, Azinphos-Methyl, Fenitrothion, Dichlorvos, Diazinon amongst others, although some are approved, but as for all active substances, there are specific restrictions of use (Phosmet, Malathion i.a).

They were the most widely pesticides sold in USA and European countries, while some of them are even registered for public health applications [27; 28; 29; 30].

Organophosphorus pesticides are esters of phosphoric acid and their derivatives have a main mode of action via inhibition processes of acetylcholine esterase (AChE). AChE is a known key enzyme of the cholinergic system responsible for the breakdown of acetylcholine that function as a neurotransmitter. Inhibition of this enzyme results in accumulation of the neurotransmitter acetylcholine (ACh) at the synapses and overstimulation of nerves and muscles.

Organophosphorus pesticides exhibit a high activity/low persistence time, are relatively unstable in the environment and degrad quickly to toxic bi-products. Nearly 75% of organophosphate pesticides are metabolized to dialkyl phosphate metabolites (DMP, DMTP, DMDTP, DEP, DETP, DEDTP), which do not inhibit acetylcholinesterase enzymes. These metabolites are not considered toxic, but they rather provide an exposure marker of the parent pesticide, suggesting a possible recent exposure. Since dialkyl phosphate metabolites may also exist in the environment as an organophosphate degradation product, the detection of these metabolites in urine may also indicate the person's exposure to the metabolite itself.

Additionally, since each of these six urinary metabolites can be produced from more than one parent pesticide, the metabolite detection alone is not enough exposure evidence to a specific pesticide and further investigation is needed. In contrary to dialkyl phosphate metabolites, other specific metabolites can also be produced by organophosphate pesticides, from one or only a few parent pesticides (eg malathion dicarboxylic acid can be produced from malathion). The detection of the specific metabolites suggests both exposure to parent pesticide and presence of metabolite in person's food or environment, [31]. Health concerns linked to organophosphate exposure are either acute or chronic. All levels of exposures have a particular toxic impact that is amplified or reduced based on age, underlying diseases, demographics etc.

Organophosphate residual analysis. Pesticide residual analysis is a methodology, used for the investigation and quantitation of banned or registered pesticides that may occur in an agricultural product. A residue is a trace (mg/Kg, µg/kg, ng/kg) of a substance, present in a matrix. The European Union has established a pesticide maximum residue levels (MRLs) for each legal pesticide, in order to avoid high limits of residual quantities.

The analytical techniques for the identification and quantification of pesticide residues are being developed and improved continuously. Gas chromatography is good for detecting volatile and thermally stable compounds, high-performance liquid chromatography is for non-volatile and high molecular weight samples and capillary electrophoresis (CE), is best for less consumptions of organic solvents with faster compound separa-

tion processes [75, 76]. Although official analytical methods exist (eg AOAC International: Association of Analytical Communities, SANCO documents from European Commission Directorate General Health and Consumer Protection, Ministry of Public Health in Netherlands etc), multi-residue in house methods are developed and followed by the vast majority of laboratories throughout Europe, mainly due to the great variety in sample preparation and detection apparatus that can be used. Considering organophosphate analysis in vegetables, a number of sample preparation steps are followed before the analysis takes place. Sub-sampling, homogenization, extraction and clean-up are considered of major importance steps since they can introduce a significant uncertainty into the final result.

Vegetable samples are reaching laboratory in certain quantities (usually 1 Kg or more), consisted from a number of individual products (eg 1-2 cucumbers, 3-5 tomatoes, 5-10 peppers). In order to acquire a representative result all individual items are chopped, homogenized in a blender and further divided in smaller quantities (eg 50 or 10g). These sub-samples can be stored at -20°C or immediately utilized.

After sample preparation the extraction step is followed in order to move (extract) the organic contaminants, such as the organophosphate pesticides, from the solid phase of the vegetable to a liquid one. Analytes are isolated from the primary matrix and their concentration is increased above the detection limit, for the specific analytical technique to be followed [32]. Many extraction techniques have been reported in the literature, with the most frequent to be: liquid-liquid extraction (LLE), solid-phase extraction (SPE), accelerated solvent extraction (ASE), gel permeation chromatography (GPC), microwave-assisted extraction (MAE), matrix solid-phase dispersion (MSPD), supercritical fluid extraction (SFE). During the extraction of vegetable samples a number of interfering compounds are co-extracted with the analytes (eg fats, carbohydrates, water, chlorophyll) [32], hence the need of an additional analytical step in order to purify the extract (clean up step). Solid phase extraction, gel permeation, sulphuric acid treatment or saponification are some of the techniques used for purification purposes.

Liquid-liquid extraction is a widely used method used for the isolation of pesticides from fruit and vegetable matrices, initially based on two basic approaches: acetone followed by partitioning with dichloromethane and petroleum ether and extraction with ethyl acetate at the presence of sodium sulfate. These two approaches have been modified by many laboratories in order to take into consideration health and environmental aspects [33]. Even though acetone is completely mixed with water there is a need of a non-polar solvent in order to enhance a distinct separation from the water phase, a fact that may lead to lower recoveries. Ethyl acetate is not completely miscible with water; hence water can be removed by the excess of sodium sulfate [34]. A typical example of liquid-liquid extraction is the addition of ethyl acetate, anhydrous Na_2SO_4 and NaCO_3 to the weighted sample, homogenization with Ultra-Turrax, filtration and concentration prior to analysis [35]. In recent years acetonitrile extraction has an increased use, particularly after the development of the QuEChERS method, where anhydrous NaCl with MgSO_4 are used, leading to increased recoveries of polar compounds [36].

Liquid-liquid extraction can be used with out further purification steps, or it can be combined by various clean-up steps, in order to minimize matrix effect, improve ruggedness, decrease LOQ and prohibit contamination in detection apparatus [35; 37]. Supercritical fluid extraction (SFE), solid-phase microextraction (SPME), stir bar sorption extraction (SBSE) and matrix solid-phase dispersion (MSPD) are additional purification steps after the first extraction of the vegetable matrix. SFE can lead to a pure extract but with high instrumentation cost and a large number of parameters needed to be optimized [38; 39; 40], while SPME can provide a fast and cheap method more oriented to qualitative analysis [41]. SBSE has been reported as an efficient method for less polar pesticides [42] and MSPD can utilize reversed phase material (C_{18} and C_8) and Florisil sorbent for more polar pesticides [43; 44; 45].

Solid-phase extraction (SPE) initially used Florisil and Silica sorbents, due to strong adsorption of organophosphates, while later graphitized carbon black (GCB) sorbents were used, in order to increase pigment removal from fruit and vegetable extracts [46; 47; 48; 49]. Since carbon sorbents did not eliminate matrix problems, the use of reverse phase C_{18} was introduced in order to remove non-polar co-extractives from vegetable extracts [50; 51; 52]. Chemically bonded stationary phases (aminopropyl, $-\text{NH}_2$), primary-secondary amine (PSA) and strong anion exchanger (-SAX) were also used in multi-residue pesticide methods. The bonded normal phase SPE columns (with PSA and $-\text{NH}_2$) has been reported as an effective step in order to reduce fatty acids, while C_{18} and -SAX remove less matrix co-extractives from the vegetable extract. Researchers [50; 51] have also reported the combination of two or three SPE columns (GCB, C_{18} , -SAX, PSA, $-\text{NH}_2$) as an effective clean-up step of sample extracts. SPE sorbents were also used by Anastassiades et al [36] during the development of the QuEChERS method with improved removing of matrix co-extractants and reduced laboratory handling. Letohay and others [53; 54] further modified the QuEChERS method, in order to improve recoveries for a wide range of pesticides from various non-fatty matrices, with pH from 2 to 7. Dispersive solid-phase extraction (DSPE)

was called the method, where after liquid-liquid extraction the SPE material was mixed with a portion of the vegetable crude extract. DSPE with a combination of PSA and GCB can be used for vegetable samples high in carotinoides (eg red pepper) or chlorophyll (eg lettuce), while QuEChERS method with a combination of C₁₈ and PSA has been used for a large number of pesticides from various food matrices, [33]. Gel permeation chromatography (GPC) is usually used for fatty matrices with high molecular weight co-extractants, but with an increase in solvents amounts and laboratory time.

Organophosphate pesticide detection is usually performed with gas chromatography (GC) coupled with nitrogen-phosphorous detectors (NPD), following specific temperature programs. The use of auto-samplers facilitates analysis and eliminates errors. The use of programmed temperature vaporizer (PTV) can facilitate large volume sample injections. Confirmation of positive samples can be performed with mass spectrometers-detectors (MS or MS/MS), usually either with electron impact ionization (EI) in full scan mode (as a means to evaluate clean-up efficiency) or selective ion monitoring (SIM) (evaluation of recovery efficiency) or atmospheric pressure chemical ionization mode (APCI). Liquid chromatography (LC) can also be used coupled with mass selective detectors (MS, MS/MS) with atmospheric pressure chemical ionization (Pizzutti, 2007). Ultra high pressure liquid chromatography (UHPLC) systems, working with higher operating pressures than in normal LC are gradually gaining ground in pesticide analysis, [56].

Review of Results – Discussion. This paper focused on reviewing organophosphate pesticide monitoring in European countries that was carried out by various researchers, independently to National authorities or official European monitoring bodies, during the last five years. Results showed a limited amount of information available for organophosphate pesticides. Even though hundreds of papers are published about pesticide monitoring in agricultural products and the vast majority of work deals with the development of new multi-residue in house methods, new sampling preparation methods, modern analysis apparatus, validation of new methods or evaluation of various pesticide analysis parameters, organophosphates are not preferred by private or educational laboratories in their analysis schedule. Organophosphates seem to be forgotten by researchers these days. Nevertheless, the review found a limited number of scientific publications and selected representative information of these data is presented in Table 1, so as to highlight the various sample preparation techniques, the analysis apparatus used, as well as any validation that is performed in order to prove method suitability, as far as organophosphate analysis is considered.

Data are presented in Table 1, initially categorized by country, vegetable commodity, number of samples monitored and number of various organophosphates analyzed. The majority of countries found during this review to deal with organophosphates are: Spain, Greece, Netherlands, Portugal and Italy. This is well explained since these countries are the main source of vegetable production through out Europe, due to excellent climate conditions (most of them being Mediterranean countries). Nevertheless other countries are entering vegetable market such as Bulgaria or Czech Republic.

Considering sample preparation for organophosphate analysis, all laboratories sample a significant amount of vegetable (1 or 2 Kg), consisted of various vegetable individual items, washing is avoided, vegetables are chopped and homogenized in mixers. Sub-sampling is usually used, in order to end with a representative quantity of the initial sample and not jeopardize the trace analysis. Efforts are taken, in order sample preparation and the following extraction or purification steps to be identical and with small variations in the final procedure to end with different extracts for different pesticide categories. Extraction is followed by all researchers, with liquid extraction being the most common method, using acetone, ethyl acetate or acetonitrile. There are cases where supercritical fluid extraction (SFE), single drop microextraction (SDME), accelerated solvent extraction (ASE), matrix solid phase dispersion (MSD) or stir bar sorptive extraction (SBSE) are used. Some researchers did not use any further purification steps, while those who use clean-up procedures prefer solid phase extraction with GCB or SAX and PSA. Modified QuEChERS method with acetonitrile seems to be the most preferable method used lately in private or educational laboratories, since it improves extraction [53; 57] and provides acceptable recovery rates, resulting to a simple, cheap and reliable sample preparation technique (extraction and purification) in order to meet the low levels required by European MRLs standards [58].

The sample preparation techniques, used for extraction and purification of vegetable samples before the sample analysis, exhibit a number of variations in initial matrix amount required, solvent usage, toxic or flammable chemicals disposal, time and laboratory work required. Liquid – liquid extraction techniques are reported as time consuming with large amount of solvents usage. The use of Soxhlet apparatus reduces laboratory cost but it still requires large amount of solvents, while microwave assisted Soxhlet extraction reduces sample preparation time. Ultrasonification extraction has been reported to facilitate the solvent leaching process. In order to further reduce extraction time and solvent usage microwave assisted extraction (MAE), accelerated solvent extraction (ASE) and supercritical fluid extraction (SFE) can be used, providing a better analyte desorption and diffusion from the solid matrix. These techniques can lead to lower temperature requirements, higher

extraction rates, less solvent usage, while in certain cases (eg SFE) can produce extracts with no further clean up step required [32]. Additionally to liquid extraction, solid adsorbent materials can be used for extracting analytes from the vegetable matrix. Solid-phase extraction (SFE), solid-phase microextraction (SPME), matrix solid-phase dispersion (MSPD) and stir-bar sorptive extraction (SBSE) are reported to provide more advantages comparing to liquid – liquid extraction techniques (simultaneous concentration of analytes and removal of interfering compounds, less solvent usage, more samples analyzed in shorter times), [32]. Solid phase extraction techniques involve purification of analyte extracts while gel permeation is mainly used to purify extracts with an increased fat content.

Organophosphate detection is carried out with GC-NPD, while there are cases where GC-ECD methods have been developed. Confirmation is performed by the majority of researchers, usually using GC-MS, GC-MS/MS or LC-MS/MS. Currently, the use of LC-MS or triple quadrupole MS for pesticide trace analysis is increasing, despite the high purchase cost for private or education laboratories. Additionally, all laboratories have developed some sort of validation procedure, evaluating recovery (Re %), relative standard deviation (RSD %), limit of detection (LOD in mg/Kg or µg/Kg) and limit of quantitation (LOQ mg/Kg or µg/Kg), (Table 1), along with other parameters not shown in Table 1 (eg precision, linearity, uncertainty etc). These values were found mostly within SANCO requirements [59]. There is not enough information about the accreditation status of these laboratories (ISO 17025) or the participation in official proficiency tests.

Table 1

Selected published data on pesticide monitoring with focus to organophosphates

Food description			Sample preparation (Extraction, and/or Clean-up steps)	Detection Apparatus (Characteristics)	Method Validation Characteristics (range)				References
Matrix Country	Sno	OPPs			Re (%)	RSD (%)	LOD (µg/Kg)	LOQ (µg/Kg)	
Spain Vegetables	23	2	Liquid extraction (acetonitrile) Solid phase extraction (multi – layer Superclean Envi Carbil / PSA SPE cartridge, acetonitrile : toluene)	PTV – GC- ITMS	80 – 96	1 – 7	1 – 2	1- 4	[60]
Greece Vegetables	13	8	Single – drop microextraction (SDME)	GC – NPD	29.9 – 51	4.9 – 8.7	0.6-10.0	-	[61]
Spain Vegetables	-	5	Liquid extraction (acetonitrile, acetone, ethyl- acetate)	GC – NPD GC – MS (SIM)	71.3-115	2.1 – 7	0.1-3.8	0.5-12.6	[62]
Austria Vegetables	-	22	Modified QuEChERS method (acetonitrile)	GC – MS	70-132	2.9-17.3	0.4-48.2	1.2-161	[63]
Spain Vegetables	-	1	Liquid extraction (dichloromethane)	GC / PFPD	89.2 – 92	8.2-14.9	3	10	[64]
Spain Vegetables	-	5	Supercritical fluid extraction (SFE)	RPLC – GC LVI-GC-NPD	-	-	1.49-7.66	-	[65]
Spain Fruits	429	11	Liquid extraction (ethyl – acetate)	GC-NPD	60-92	5-18		0.01-100	[66]
Spain Fruits	116	1	Liquid extraction (ethyl – acetate)	LC-MS	72	19	-	20	[45]
Bulgaria Fruits Vegetables	-	5	Liquid extraction (acetone) Solid-phase extraction (GCB, SAX-PSA)	GC-ECD GC-MS	74-114	2-16	1- 5	5 – 10	[67]
Slovak Republic Fruits	-	7	QuEChERS method (acetonitrile)	GC-NCI-MS GC-EI-MS	-	1.7-8.1 6-7.5	0.0068-0.113 0.14-0.5	0.0234-0.160 0.36-1.66	[68]
Czech Republic	-	37	QuEChERS method (acetonitrile) (clean up- PSA)	UHPLC-TOF MS	-	2.25-13.63	-	<10	[69]
Portugal Fruits	28	11	Dispersive liquid-liquid microextraction (DLLME)	MD – GC/MS	66-106	3-20	0.2-4.1	0.06-1.24	[70]
Portugal Fruits	-	7	QuEChERS method (acetonitrile)	LP-GC-MS	74-90	4-12	1.2-4	-	[71]
Czech Republic Fruits	-	2	Liquid extraction (acetonitrile)	LC-MS/MS GC-MS	93-94	7-8	-	9-11	[72]
Slovak Republic Fruits Vegetables	-	4	QuEChERS method (acetonitrile)	GC-EI-MS GC-NCI-MS	-	3.4-5.3 7.5-12	0.033-0.055 0.08-0.34	0.11-0.18 0.27-1.13	[73]
Spain Vegetables	-	2	QuEChERS method (acetonitrile)	GC-MS (SIM) GC-MS/MS	84.9-98.2 -	2.0-4.6 6.0-7.5	0.2-0.3 0.1-0.3	-	[74]

Conclusions – Recommendations. Even though European legislation exists and applies strict guidelines and regulations in order to reduce the use of persistent pesticides and encourage the development of target specific pesticides, this review shows a tendency in pesticide research towards more new and modern chemical molecules that are used in today's agriculture. Organophosphates seemed to be old fashioned among various laboratories in European countries, even though they are still used for cultivation purposes. Certainly, we should take into consideration the existence of National laboratories in each European country, which are monitoring pesticides and informing European Union in order to produce specific reports. Considering private laboratories in each country, we should acknowledge that routine analysis is taking place and certainly organophosphates are monitored among the other pesticide categories. In this case, we should underline that private laboratories operate with agricultural samples originating from independent farmers, farmer groups, super-markets or factories that export vegetables abroad, to other European countries and the majority of samples are analyzed as a prerequisite for certification purposes. So, assuming all agricultural products exported to other European countries are certified and the requirements of each certification scheme are strictly followed, we could conclude that public health is not questioned. But, what happens inside each European country, where certification is not necessary for selling agricultural products (as a marketing requirement) and products reach consumers, sold in local markets or supermarkets, with only National authorities been responsible for inspection and monitoring? Analysis cost, insufficient laboratory methods (eg pesticide metabolites that are not covered by in-house methods), new pesticides or foreign imported pesticides, are some of the parameters that may limit the successful pesticide monitoring from National authorities. Not only organophosphates but all pesticides are in question and there is not substantial evidence (published data) to prove that farmers have been using legal-registered pesticides or they have followed the necessary pre-harvest days after pesticide usage (strictly defined for each pesticide).

Finally, we should underline the importance of accreditation to all major pesticide categories from National laboratories and the usefulness of specific agricultural sampling protocols, which would take into consideration cultivation process, product origins, marketing strategies, as well as weather fluctuation and consumer demands. Certification schemes used for vegetable cultivation in European countries (such as GLOBALGAP, formerly known as EUREPGAP) should become a prerequisite, not only for produce intended to be sold abroad but also for produce that is sold inside each country. Traceability and food labeling techniques can provide evidence for certain production methods (eg organic food) or specific food product origin, while facilitate traceability in case of MRLs violation.

Abbreviations:

Sno : number of samples analyzed, OPPs no : number of organophosphate analytes analyzed, Re (%) : recovery %, RSD (%) : relative standard deviation %

LOD (mg/Kg): limit of detection, LOQ (mg/Kg): limit of quantitation,

PTV – GC- ITMS : programmable temperature vaporization injector – gas chromatography with ion trap mass spectrometry

GC-NPD : gas chromatography – nitrogen phosphorous detector

GC – MS (SIM) : gas chromatography – mass spectrometry with selected ion monitoring

GC / PFPD : gas chromatography – pulse flame photometric detection

RPLC – GC : reversed phase liquid chromatography – gas chromatography

LVI-GC-NPD : large volume injection – gas chromatography – nitrogen phosphorous detector

LC-MS : liquid chromatography – mass spectrometry

GC-ECD : gas chromatography – electron capture detector

GC-NCI-MS : gas chromatography -negative chemical ionization mass spectrometry

GC-EI-MS : gas chromatography – electron impact ionization – mass spectrometry

UHPLC-TOF MS : ultra high performance liquid chromatography – time of flight mass spectrometry

MD – GC/MS : multidimensional gas chromatography – mass spectrometry

LP-GC-MS : low pressure gas chromatography – mass spectrometry

References:

1. FAO/WHO Report of a Joint Workshop (2004) Fruit and Vegetables for Health. Joint FAO/WHO Workshop on Fruit and Vegetables for Health. Fruit and vegetables for health: Report of a Joint FAO/WHO Workshop, Kobe, Japan. Pp 1-39, ISBN: 92-4-159281-8
2. Pomadera C (2006) Word Development Report 2008. Agriculture for Development. Contract Agriculture. Lessons from experiences in Costa Rica.
3. Prudent P, Loko S, Deybe D, Vaissayre M (2007) Factors limiting the adoption of IPM practices by cotton farmers in begin : a participatory approach. *Experimental Agriculture* 43:113-124
4. Kokkinakis E, Boskou G, Fragkiadakis GA, Kokkinaki A, Papadaki M, Lapidakis N (2007) Microbiological quality of tomatoes and peppers produced under the Good Agricultural Practices protocol AGRO 2-1 & 2-2 in Crete, Greece, *Food Control* 18(12):1538-1546

5. Sarig Y (2003) Traceability of Food Products. *Agricultural Engineering International : the CIGR Journal of Scientific Research and Development*. Invited Overview Paper. Vol V December 2003. Presented at the Club of Bologna meeting. November 16, 2003. Bologna Italy
6. Regulation (EC) No 1107/2009 of the European Parliament and of the Council of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC.
7. Regulation (EC) No 396/2005 of the European Parliament and of the Council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EEC Text with EEA relevance.
8. Regulation (EU) 2019/1009 of the European Parliament and of the Council of 5 June 2019 laying down rules on the making available on the market of EU fertilising products and amending Regulations (EC) No 1069/2009 and (EC) No 1107/2009 and repealing Regulation (EC) No 2003/2003 (Text with EEA relevance)
9. Directive 2009/128/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for Community action to achieve the sustainable use of pesticides (Text with EEA relevance) Text with EEA relevance.
10. European Commission (2010) Food and Veterinary Office – Special Reports: Pesticide Monitoring Reports. (http://ec.europa.eu/food/fvo/specialreports/pesticides_index_en.htm) Accessed 5.2.2010.
11. RASFF (2010) European Community's Rapid Alert System for Food and Feed (http://ec.europa.eu/food/food/rapidalert/index_en.htm) Accessed 5.2.2010.
12. Conacher HBS, Mess J (1993) Assessment of human exposure to chemical contaminants in foods. *Food Additives and Contaminant* 10:5-15.
13. Casarett and Doull's Toxicology (2001) *The Basic Science of Poisons* (6th Edition). Editor Klaassen, Curtis D. 2001 McGraw-Hill.
14. Toft G, Hagmar L, Giwercman A, Bonde JP (2004) Epidemiological evidence on reproductive effects of persistent organochlorines in humans. *Reproductive Toxicology* 19(1):5-26.
15. Lee JW, Baccarelli A, Tretiakova M, Gorbanev S, Lomtev A, Klimkina I, Tchibissov V, Averikna O, Dosemeci M (2006) Pesticide exposure and lung cancer mortality in Leningrad province in Russia. *Environment International* 32:412-416.
16. Lo S, King I, Allera A, Klingmuller D (2007) Effects of various pesticides on human 5 α -reductase activity in prostate and LNCaP cells. *Toxicology in vitro* 21(3):502-508.
17. Irigaray P, Newby JA, Clapp R, Hardell L, Sasco AJ, Howard V, Montagnier L, Epstein S, Belpomme D (2007) Risk factors and environmental agents causing cancer: an overview. *Biomedicine & Pharmacotherapy* 61(10):640-658.
18. Nasterlack M (2007) Pesticides and childhood cancer : an update. *International Journal of Hygiene and Environmental Health* 210:645-657.
19. Tutudaki M, Tsakalof KA, Tsatsakis MA (2003) Hair analysis used to assess chronic exposure to the organophosphate diazinon : a model study with rabbits. *Human & Experimental Toxicology* 22(3):159-164.
20. Tsatsakis A, Tutudaki M (2004) Progress in pesticide and POPs hair analysis for the assessment of exposure. *Forensic Science International* 145:195-199.
21. Zhou R, Zhou L, Yang K, Chen Y (2006) Distribution of organochlorine pesticides in surface water and sediments from Qiantang River, East China. *Journal of Hazardous Materials* 137:68-75.
22. Claver A, Ormad P, Rodriguez L, Ovelheiro JL (2006) Study of the presence of pesticides in surface Waters in the Ebro River basin (Spain). *Chemosphere* 64:1437-1443.
23. Lauria L, Settini L, Spinelli A, Figa-Talamanca I (2006) Exposure to pesticides and time to pregnancy among female greenhouse workers. *Reproductive Toxicology* 22:425-430.
24. Maloschik E, Ernst A, Hegedus G, Darvas B, Szekacs A (2007) Monitoring water-polluting pesticides in Hungary. *Microchemical Journal* 85:88-97.
25. Carreno J, Rivas A, Granada A, Lopez-Espinosa JM, Mariscal M, Olea N, Serrano F (2007) Exposure of young men to organochlorine pesticides in Southern Spain. *Environmental Research* 103:55-61.
26. Bretveld WR, Hooiveld M, Zielhuis AG, Pellegrino A, van Rooij I, Roeleveld N (2008) Reproductive disorders among male and female greenhouse workers. *Reproductive Toxicology* 25(1):107-114.
27. Abbott JA (1999) Quality measurement of fruits and vegetables. *Postharvest Biology and Technology* 15:207-225.
28. Abou-Arab AAK (1999) Behavior of pesticides in tomatoes during commercial and home preparation. *Food Chemistry* 65:509-514.
29. Bai Y, Zhou L, Wang J (2006) Organophosphorus pesticide residues in market foods in Shaanxi area, China. *Food Chemistry* 98:240-242.
30. CDC (2009) Centers for Disease Control and Prevention. Department of Health and Human Services. Fourth National Report on Human Exposure to Environmental Chemicals.
31. Lu C, Bravo R, Caltabiano LM, Irish RM, Weerasekera G, Barr DB (2005) The presence of dialkylphosphates in fresh fruit juices: implication for organophosphorus pesticide exposure and risk assessments. *Journal of Toxicology and Environmental Health part A* 68(3):209-27.
32. Beyer A, Biziuk M (2008) Applications of sample preparation techniques in the analysis of pesticides and PCBs in food. *Food Chemistry* 108(2):669-680.
33. Hercegova A, Domotorova M, Matisova E (2007) Sample preparation in the analysis of pesticide residues in baby food with subsequent chromatographic determination. *Journal of Chromatography A* 1153:54-73.
34. Van der Hoff R, G, Van Zoonen P (1999) Trace analysis of pesticides by gas chromatography. *Journal of Chromatography A* 843:301-322.
35. Patel K, Fussell JR, Goodall MD, Keely JB (2004) Evaluation of large volume-difficult matrix introduction-gas chromatography-time of flight-mass spectrometry (LV-DMI-GC-TOF-MS) for the determination of pesticides in fruit-based baby foods. *Food Additives and Contaminants part A* 21:658-669.

36. Anastassiades M, Lehotay JS, Stajnbaher D, Schenck JF (2003) Fast and easy multiresidue method employing acetonitrile extraction/partitioning and «dispersive solid-phase extraction» for the determination of pesticide residues in produce. *Journal of AOAC International* 86(2):412-431.
37. Cajka T, Mastovska K, Lehotay JS, Hajslova J (2005) Use of automated direct sample introduction with analyte protectants in the GC-MS analysis of pesticide residues. *Journal of Separation Science* 28(9-10):1048-1060.
38. Lehotay JS (1997) Supercritical fluid extraction of pesticides in foods. *Journal of Chromatography A* 785(1-2):289-312.
39. Eller IK, Lehotay JS (1997) Evaluation of hydromatrix and magnesium sulfate drying agents for supercritical fluid extraction of multiple pesticides in produce. *Analyst* 122(5):429-435.
40. Valverde-Garcia A, Fernandez-Alba RA, Agüera A, Contreras M (1995) Extraction of methamidophos residues from vegetables with supercritical fluid carbon dioxide. *Journal of AOAC International* 78(3):867-873.
41. Lehotay JS, Hajslova J (2002) Application of gas chromatography in food analysis. *TrAC – Trends in Analytical Chemistry* 21(9-10):686-697.
42. Sandra P, Tienpont B, David F (2003) Multi-residue screening of pesticides in vegetables, fruits and baby food by stir bar sorptive extraction–thermal desorption–capillary gas chromatography–mass spectrometry. *Journal of Chromatography A* 1000:299-309.
43. Navarro M, Pico Y, Marín R, Manes J (2002) Application of matrix solid-phase dispersion to the determination of a new generation of fungicides in fruits and vegetables. *Journal of Chromatography A* 968(1-2):201-209.
44. Kristenson ME, Hoverkate EGJ, Slooten JC, Ramos L, Vreuls RJJ, Brinkman UATH (2001) Miniaturized automated matrix solid-phase dispersion extraction of pesticides in fruit followed by gas chromatographic-mass spectrometric analysis. *Journal of Chromatography A* 917(1-2):277-286.
45. Blasco C, Pico Y, Manes J, Font G (2002) Determination of fungicide residues in fruits and vegetables by liquid chromatography-atmospheric pressure chemical ionization mass spectrometry. *Journal of Chromatography A* 947(2):227-235.
46. Pang GF, Cao YZ, Fan CL, Zhang JJ, Li XM (1999) Multiresidue gas chromatographic method for determining synthetic pyrethroid pesticides in agricultural products: Collaborative study. *Journal of AOAC International* 82(1):186-196.
47. Cook J, Beckett MP, Reliford B, Hammock W, Engel M (1999) Multiresidue analysis of pesticides in fresh fruits and vegetables using procedures developed by the Florida Department of Agriculture and Consumer Services. *Journal of AOAC International* 82(6):1419-1435.
48. Dalluge J, Van Rijn M, Beens J, Vreuls RJJ, Brinkman UATH (2002) Comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection applied to the determination of pesticides in food extracts. *Journal of Chromatography A* 965(1-2):207-217.
49. Fernandez Moreno JL, Arrebola Liebanas FJ, Garrido Frenich A, Martinez Vidal JL (2006) Evaluation of different sample treatments for determining pesticide residues in fat vegetable matrices like avocado by low-pressure gas chromatography–tandem mass spectrometry. *Journal of Chromatography A* 1111(1):97-105.
50. Schenck JF, Lehotay JS (2000) Does further clean-up reduce the matrix enhancement effect in gas chromatographic analysis of pesticide residues in food? *Journal of Chromatography A* 868(1):51-61.
51. Schenck JF, Lehotay JS, Vega V (2002) Comparison of solid-phase extraction sorbents for cleanup in pesticide residue analysis of fresh fruits and vegetables. *Journal of Separation Science* 25(14):883-890.
52. Stajnbaher D, Zupancic-Kralj L (2003) Multiresidue method for determination of 90 pesticides in fresh fruits and vegetables using solid-phase extraction and gas chromatography-mass spectrometry. *Journal of Chromatography A* 1015:185-198.
53. Lehotay SJ, Mastovska K, Lightfield AR (2005a) Use of buffering and other means to improve results of problematic pesticides in a fast and easy method for residue analysis of fruits and vegetables. *Journal of AOAC International* 82(2):615-629.
54. Domotorova M, Hercegova A, Matisova E (2006) Monitoring of pesticide residues in apples from Slovakia for baby food production. *Czech Journal of Food Sciences* 24:84.
55. Pizzutti RI, de Kok A, Zanella R, Adaimea BM, Hiemstra M, Wickert C, Prestes DO (2007) Method validation for the analysis of 169 pesticides in soya grain, without clean up, by liquid chromatography–tandem mass spectrometry using positive and negative electrospray ionization. *Journal of Chromatography A* 1142:123-136.
56. Guilleme D, Schappler J, Rudaz S, Veuthey JL (2010) Coupling ultra-high-pressure liquid chromatography with mass spectrometry. *TrAC Trends in Analytical Chemistry* 29(1):15-27.
57. Lehotay SJ, De Kok A, Hiemstra M, Van Bodegraven P (2005b) Validation of a fast and easy method for the determination of residues from 229 pesticides in fruits and vegetables using gas and liquid chromatography and mass spectrometric detection. *Journal of AOAC International* 82(2):595-614.
58. Leandro CC, Hancock P, Fussell RJ, Keely BJ (2006) Comparison of ultra-performance liquid chromatography and high-performance liquid chromatography for the determination of priority pesticides in baby foods by tandem quadrupole mass spectrometry. *Journal of Chromatography A* 1103(1):94-101.
59. SANCO (2010) European Commission. Plant Health. Guidance Documents (Pesticide Residues). Method Validation and Quality Control Procedures for Pesticide Residues Analysis in Food and Feed. Document No. SANCO/10684/2009. (http://ec.europa.eu/food/plant/protection/resources/publications_en.htm#residues) Accessed 5.2.2010.
60. Gonzalez-Rodriguez RM, Rial-Otero R, Cancho-Grande B, Simal-Gandara J (2008) Determination of 23 pesticide residues in leafy vegetables using gas chromatography–ion trap mass spectrometry and analyte protectants. *Journal of Chromatography A* 1196-1197, 100-109.

61. Amvrazi GE, Tsiropoulos GN (2009) Application of single-drop microextraction coupled with gas chromatography for the determination of multiclass pesticides in vegetables with nitrogen phosphorus and electron capture detection. *Journal of Chromatography A* 1216:2789–2797.
62. Fenoll J, Helli'n P, Marti'nez MC, Miguel M, Flores P (2007) Multiresidue method for analysis of pesticides in pepper and tomato by gas chromatography with nitrogen–phosphorus detection. *Food Chemistry* 105:711–719.
63. Lesueur C, Knittl P, Gartner M, Mentler A, Fuerhacker M (2008) Analysis of 140 pesticides from conventional farming foodstuff samples after extraction with the modified QuEChERS method. *Food Control* 19:906–914.
64. Gonzalez FJE, Vidal JLM, Castro Cano LM, Galera MM (1998) Levels of metamidophos in air and vegetables after greenhouse applications by gas chromatography. *Journal of Chromatography A* 829:251–258.
65. Cortes MJ, Vázquez A, Santa-Marva G, Blanch PG, Jesús Villan J (2009) Pesticide residue analysis by RPLC–GC in lycopene and other carotenoids obtained from tomatoes by supercritical fluid extraction. *Food Chemistry* 113:280–284.
66. Berrada H, Fernández M, Ruiz JM, Moltø CJ, Mapes J, Font G (2010) Surveillance of pesticide residues in fruits from Valencia during twenty months (2004/05). *Food Control* 21:36–44.
67. Balinova A, Mladenova R, Shtereva D (2007) Solid-phase extraction on sorbents of different retention mechanisms followed by determination by gas chromatography–mass spectrometric and gas chromatography–electron capture detection of pesticide residues in crops. *Journal of Chromatography A* 1150:136–144.
68. Hušková R, Matisová E, Hrouzková S, Svorc L (2009a) Analysis of pesticide residues by fast gas chromatography in combination with negative chemical ionization mass spectrometry. *Journal of Chromatography A* 1216:6326–6334.
69. Lacina O, Urbanová J, Poustka J, Hajslova J (2010) Identification/quantification of multiple pesticide residues in food plants by ultra-high-performance liquid chromatography–time-of-flight mass spectrometry. *Journal of Chromatography A* 1217:648–659.
70. Cunha CS, Fernandes OJ, Oliveira MBPP (2009a) Fast analysis of multiple pesticide residues in apple juice using dispersive liquid–liquid microextraction and multidimensional gas chromatography– mass spectrometry. *Journal of Chromatography A* 1216:8835–8844.
71. Cunha CS, Fernandes OJ, Oliveira MBPP (2009b) Fast low-pressure gas chromatography–mass spectrometry method for the determination of multiple pesticides in grapes, musts and wines. *Journal of Chromatography A* 1216:119–126.
72. Ticha J, Hajslova J, Jech M, Honzicek J, Lacina O, Kohoutkova J, Kocourek V, Lansky M, Kloutvorova J, Falta V (2008) Changes of pesticide residues in apples during cold storage. *Food Control* 19:247–256
73. Hušková R, Matisová E, Svorc L, Mocaka J, Kirchner M (2009b) Comparison of negative chemical ionization and electron impact ionization in gas chromatography–mass spectrometry of endocrine disrupting pesticides. *Journal of Chromatography A* 1216:4927–4932.
74. Mezcua M, Ferrer C, Garcia-Reyes FJ, Bueno MJM, Sigrist M, Fernandez-Alba RA (2009) Analyses of selected non-authorized insecticides in peppers by gas chromatography/mass spectrometry and gas chromatography/tandem mass spectrometry. *Food Chemistry* 112:221–225.
75. Antonia Maria Rojano-Delgado, Maria D. Luque de Castro (2014). Capillary electrophoresis and herbicide analysis: Present and future perspectives. *Electrophoresis* 35, 2509–2519
76. Abdalla A. Elbashia and Hassan Y. Aboul-Enein (2015). Separation and analysis of triazine herbicide residues by capillary electrophoresis *Biomedical Chromatography* 29: 835–842

УДК 004.9

DOI <https://doi.org/10.32689/maup.it.2021.2.9>

Денис ШИБАЄВ

аспірант кафедри технічної кібернетики та інформаційних технологій, Одеський національний морський університет, вул. Мечникова 34, Одеса, Україна, індекс 65029 (denscreamer@gmail.com)

ORCID: <https://orcid.org/0000-0002-3260-5843>

Наталя ШИБАЄВА

кандидат технічних наук, доцент кафедри інформаційних технологій, Державний університет «Одеська Політехніка», просп. Шевченко 1, Одеса, Україна, індекс 65001 (nati.shibaeva@gmail.com)

ORCID: <https://orcid.org/0000-0002-7869-9953>

Тетяна ОТРАДСЬКА

кандидат технічних наук, доцент кафедри інформаційних систем, Державний університет «Одеська Політехніка», просп. Шевченка 1, Одеса, Україна, індекс 65001 (tv_61@ukr.net)

ORCID: <https://orcid.org/0000-0002-5808-5647>

Микита КУЦ

студент, Міжрегіональна Академія управління персоналом, вул. Фрометівська 2, Київ, Україна, індекс 03039 (nickkutz@gmail.com)

ORCID: <https://orcid.org/0000-0001-5264-9475>

Denis SHIBAYEV

Graduate student of the Department of Technical Cybernetics and Information Technologies, Odessa National Maritime University, 34 Mechnikova Street, Odessa, Ukraine, postal code 65029 (denscreamer@gmail.com)

Natalia SHIBAYEVA

Candidate of Technical Sciences, Associate Professor of Information Technology, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (nati.shibaeva@gmail.com)

Tatiana OTRADSKA

PhD in Technical Sciences, Associate Professor of Information Systems Department, Odessa Polytechnic State University, 1 Shevchenko Avenue, Odessa, Ukraine, postal code 65001 (tv_61@ukr.net)

Mykyta KUTS

Student of the Interregional Academy of Personnel Management, 2 Frometivska Street, Kyiv, Ukraine, postal code 03039 (nickkutz@gmail.com)

Бібліографічний опис статті: Шибаяєв Д., Шибаяєва Н., Отрадська Т., Куц М. Розробка інформаційно-аналітичної системи з прийняття рішень в кредитному скорингу. *Інформаційні технології та суспільство*. 2021. Вип. 2. С. 85–93. DOI: <https://doi.org/10.32689/maup.it.2021.2.9>

Bibliographic description of the article: Shibayev, D. Shibayeva, N., Otradska, T., Kuts, M. (2021). Rozrobka informatsiino-analitychnoi systemy z pryiniattia rishen v kredytnomu skorynhu [Development of information and analytical system for decision making in credit scoring]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 2, 85–93. DOI: <https://doi.org/10.32689/maup.it.2021.2.9>

**РОЗРОБКА ІНФОРМАЦІЙНО-АНАЛІТИЧНОЇ СИСТЕМИ
З ПРИЙНЯТТЯ РІШЕНЬ В КРЕДИТНОМУ СКОРИНГУ**

Анотація. Впровадження інформаційних систем до фінансової галузі є популярним рішенням вже понад 20 років. Це пов'язано зі стрімким зростанням фінансово-економічних розрахунків та клієнт-орієнтованості фінансових закладів. Більшість операцій, які стосуються розрахунків, фінансових обчислень, систем кредитування, тощо, виконуються завдяки роботі з інформаційно-програмним забезпеченням. Це дозволяє мінімізувати помилки операторів

системи, зберігати цифрову форму документації, впроваджувати різні Інтернет-сервіси для популяризації фінансових послуг.

Метою статті є опис процесу проектування та визначення функціональної логіки розробки програмного засобу з прийняття рішень при аналізі клієнтів в фінансових установах. Реалізація поставленої мети передбачає вирішення низки завдань: 1) формування логіки оцінки вхідних критеріїв від клієнтів; 2) проектування концепції системи з використанням функціональної мови проектування UML; 3) розробка демонстраційної версії інформаційної системи на основі сформованого проекту та визначеної функціональної логіки.

Наукова новизна. Спроектване та розроблене рішення використовує полегшені методи обробки вхідної інформації, що дозволяє збільшити швидкість обробки заяв від користувачів. В роботі застосовуються динамічні методи аналізу вхідних факторів із застосуванням додаткових зовнішніх перевірок клієнтів. Застосовані засоби розробки оптимізують роботу системи, за рахунок застосування web-засобів розробки.

Як висновок, у статті наголошується, що сучасний кредитний скоринг є актуальною сферою застосування в різноманітних фінансових установах, та має включати в себе функціональні заходи з обробки вхідної інформації при прийнятті рішень в кожному індивідуальному випадку. Застосований метод обробки вхідних даних дозволяє мінімізувати час на обробку даних та збільшити вихідну точність рекомендацій для кредитної справи. Наведені проектування дозволили визначити якісні засоби реалізації програмної частини, а також сформувати необхідні технологічні компоненти для реалізації інформаційної системи. Подальшим розвитком системи може бути застосування більш якісних засобів обробки вхідної інформації, розширення методів початкового діагностування клієнтів та впровадження матричного методу збільшення точності кінцевого прогнозу для кожної окремої фінансової справи.

Ключові слова: кредитний скоринг, фінансовий менеджмент, обробка запитів, прийняття рішень, web-системи.

DEVELOPMENT OF INFORMATION AND ANALYTICAL SYSTEM FOR DECISION MAKING IN CREDIT SCORING

Abstract. The introduction of information systems in the financial sector has been a popular solution for over 20 years. This is due to the rapid growth of financial and economic settlements and customer orientation of financial institutions. Most operations related to settlements, financial calculations, credit systems, etc., are performed by working with information software. This allows you to minimize the errors of system operators, save the digital form of documentation, implement various Internet services to promote financial services.

The aim. The purpose of the article is to describe the process of designing and determining the functional logic of software development for decision-making in the analysis of customers in financial institutions. Realization of the set purpose provides the decision of a number of tasks: 1) formation of logic of an estimation of entrance criteria from clients; 2) designing the concept of the system using the functional design language UML; 3) development of a demonstration version of the information system on the basis of the formed project and the defined functional logic.

Scientific novelty. The designed and developed solution uses simplified methods of processing input information, which allows to increase the speed of processing requests from users. The paper uses dynamic methods of analysis of input factors with the use of additional external customer reviews. Used development tools optimize the system through the use of web-development tools.

In conclusion, the article emphasizes that modern credit scoring is a relevant area of application in multilevel financial institutions, and should include functional measures for processing input information in decision-making in each individual case. The applied method of input data processing allows to minimize time for data processing and to increase initial accuracy of recommendations for credit business. The given designs allowed to define qualitative means of realization of a software part, and also to form necessary technological components for realization of information system. Further development of the system may be the use of better means of processing input information, expanding the methods of initial diagnosis of customers and the introduction of a matrix method to increase the accuracy of the final forecast for each individual financial case.

Key words: credit scoring, financial management, request processing, decision making, web-systems.

Актуальність проблеми. Завдання кредитного скорингу виникає в банках і інших кредитних організаціях при прийнятті рішень про видачу кредитів. Завдання полягає в тому, щоб на основі деякої інформації про заявника обґрунтовано прийняти рішення – чи варто йому видавати кредит, і якщо так, то на яких умовах. Якщо мова йде про фізичних осіб, то вихідною інформацією для класифікації позичальників на «добрих» і «поганих» є створена клієнтська анкета, інформація про минулі кредити позичальників (кредитна історія), в деяких випадках – інформація про рух коштів на рахунках позичальника. В анкеті вказується сума і призначення кредиту, вік, стать, соціальний стан, освіта, професія тощо. Співробітники банку можуть вносити додаткові відомості, наприклад, психологічний портрет позичальника або результати перевірки, проведеної службою безпеки банку. В результаті формується опис позичальника, що містить від 20 до 200 ознак, залежно від використовуваної банком методики оцінки кредитоспроможності [1–3].

Поряд з класифікацією банк вирішує завдання визначення ризику, який несе в собі видача запитуваного кредиту в кожному конкретному випадку і прийняття рішення про прийнятність такого ризику для кредитного портфеля. При цьому банк намагається мінімізувати ризик і максимізувати прибуток шляхом варіювання умов кредиту, що видається – ставки, суми, терміну тощо. Для вирішення цих завдань, алгоритм повинен видавати не тільки саму класифікацію, але і оцінку ймовірності того, що даний позичальник належить до класу «поганих». Крім того, повинна бути можлива більш тонка класифі-

кація ризиків, ніж «поганий»/«добрий»: кредит може повертатися не повністю, або з затримками (що тягне операційні витрати), або достроково (що тягне втрату частини прибутку) [4].

Одне з найбільш істотних обмежень полягає в тому, що алгоритм класифікації позичальників повинен бути простим, зрозумілим кредитному експерту і повинен допускати запис на природній мові в термінах предметної області.

Для попереднього контролю, банки і компанії-продавці розробляють правила оформлення кредитної документації, а також перевірки документів, представлених потенційним кредитним позичальником. У них співробітникам рекомендується звернути увагу на:

- поведінку позичальника;
- його зовнішній вигляд;
- збіг інформації про кредитну історію позичальника, відображену в різних документах;
- пропорційність прибутків клієнта з розмірами кредиту і вартістю потенційної покупки;
- дійсність паспорта та видимі ознаки підробки документів (наприклад, потертості в паспорті, сліди переклеювання фотографії, нечітка друк, закінчення строку дії паспорта, невідповідність підписів позичальника в різних документах) [5].

Крім того, скорингові карти складаються з урахуванням напрацьованого досвіду, удосконалюються рекомендації щодо виявлення недобросовісних осіб (відсівання на першому етапі – метод візуального профайлінга). Ряд авторів відзначають впровадження автоматизованих систем оцінки кредитного позичальника, як найбільш дієвий метод виявлення шахрайства.

Аналіз останніх досліджень і публікацій. Виходячи з аналізу актуальності означеної проблеми завдання полягає в розробці методу та інструментарію який його реалізує, та дозволить виявити, оцінити і спрогнозувати кредитну поведінку здобувача роздрібного кредиту, виключаючи дані про його фінансові спроможності, які не представляється можливим підтвердити однозначно на момент отримання кредиту, але при цьому не виключаючи можливість подальшого їх доповнення.

Скоринг як метод оцінки ризиків в роздрібному кредитуванні в даний час набув найширшого застосування. Ризик, в першу чергу, залежить від того, наскільки добре оцінена можливість повернення кредитних коштів [6-7].

Поряд з очевидними перевагами скорингових моделей визначають:

- швидкість і неупередженість прийняття рішень;
- можливість диверсифікації кредитного ризику між позичальниками, тобто можливість ефективного управління кредитним портфелем;
- відсутність тривалого навчання співробітників кредитного департаменту;
- можливість провести експрес-аналіз заявки на кредит в присутності клієнта.

Експерти відзначають і безліч недоліків, які в повній мірі проявляються на російському ринку роздрібного кредитування, а саме:

- скорингові карти складаються для конкретних кредитних продуктів і завдань аналізу;
- скорингові карти мають обмежений термін споживання: тривалість періоду актуальності залежить від характеру і масштабу змін в економіці і може варіюватися від року до декількох років, якщо період зміни тенденцій можна порівняти з періодом накопичення даних статистичного аналізу, то скорингова карта може стати застарілою вже до моменту її розрахунку;
- скорингові моделі будуються на основі вибірки з числа найбільш «ранніх» клієнтів, але на даний момент в процесі кредитування задіяні вже не тільки найбільш надійні (найзаможніші, найякісніші) позичальники, далі цей процес розвивається стрімко і якість позичальників знижується;
- скоринг не захищає від позичальників, які повідомляють про себе частково недостовірні дані;
- виявляє шахраїв лише формально (наприклад, використання одного і того ж паспорта при неодноразовому отриманні кредиту);
- розробка, впровадження, обслуговування скорингових систем, що містять централізовані бази даних, вимагає високих витрат, що значно знижує рентабельність бізнесу;
- визначення оціночних ознак проводиться тільки на базі інформації про тих клієнтів, яким банк вже надав кредит, і невідомо, скільки клієнтів упустив банк;
- масова недостовірність заявлених даних.

Для вирішення аналогічних завдань аналізу, інтерпретації та прогнозування, широке застосування знаходять інтелектуальні інформаційні системи, які дозволять мінімізувати помилки в роботі кредитних менеджерів, та збільшити якість перевірки клієнтів [8-10].

Метою статті є опис особливостей розробки функціональної частини для інформаційно-аналітичної системи з прийняття рішень в фінансовому скорингу з можливим запровадженням системи в кредитно-фінансових установах.

Виклад основного матеріалу. Центральна парадигма інтелектуальних технологій сьогодні – обробка знань, системним ядром яких є база знань (БЗ) або модель предметної області.

Інший підхід до визначення обсягу вибірки – на основі критерію потужності при завданні максимально-допустимої помилки оцінки співвідношення «поганих» і «добрих» позичальників у генеральній сукупності. Навчальна вибірка позичальників для побудови моделі повинна відображати генеральну сукупність всіх потенційних позичальників з невідомими долями «поганих» і «добрих» кредитів. Припустимо, ми хочемо бути впевненим на 95%, що співвідношення «поганих» і «добрих» позичальників в навчальній вибірці повинно відображати генеральну популяцію позичальників. Знаючи розподіл залежною змінною на тестових даних, можна за формулою розрахувати необхідний обсяг вибірки:

$$n = \frac{z_y^2 w * (1 - w)}{\Delta_w^2} \quad (1)$$

Де розподіл, визначається в залежності від обраного рівня надійності γ , w – частка «поганих» клієнтів за тестовою вибіркою, Δ_w – максимально допустима гранична помилка оцінки частки «поганих» позичальників.

При визначенні мінімального обсягу вибірки можна скористатися і формулою інтервальної оцінки генеральної частки:

$$w - z_y \sqrt{\frac{w(1-w)}{n}} \leq P \leq w + z_y \sqrt{\frac{w(1-w)}{n}} \quad (2)$$

де P – оцінювана частка «поганих» позичальників у генеральній сукупності, w – частка «поганих» позичальників за тестовою вибіркою, n – обсяг початкової історичної вибірки, z_y – значення стандартного нормального закону розподілу, визначається в залежності від обраного рівня надійності γ .

Для визначення етапів взаємодії менеджера з клієнтом та інформаційною системою використовується діаграма станів, на якій відображається послідовність дій з урахування логічної складової, рис. 1.

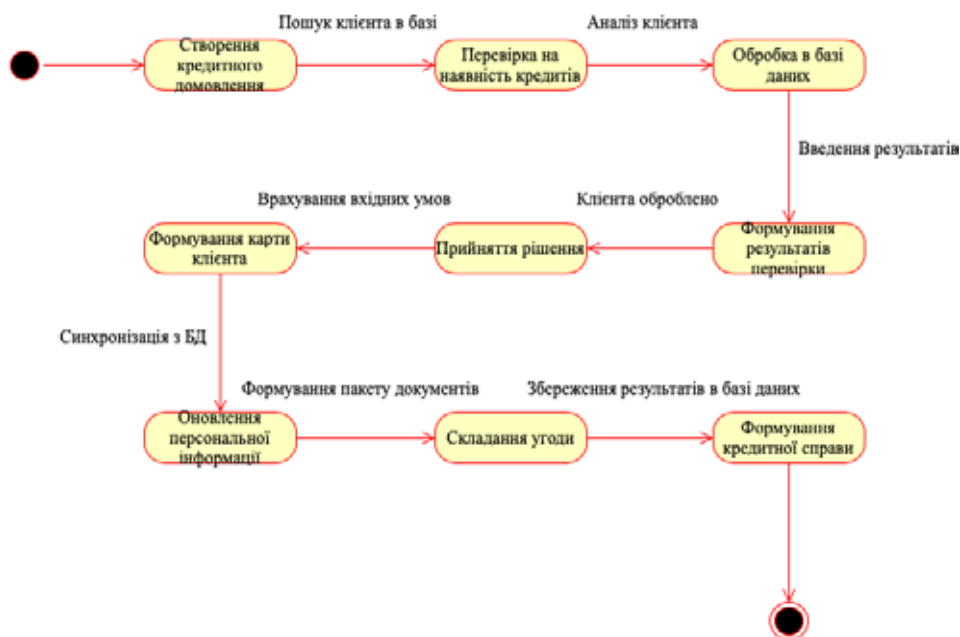


Рис. 1. Діаграма станів роботи менеджера з кредитування

Інформаційна система має клієнт-серверну архітектуру, що дозволяє оброблювати велику кількість аналітичної інформації в залежності від сформованого початкової умови. Важливим компонентом системи є визначення достовірності даних клієнта та поведінки системи в процесі блокування клієнта по поточній справі, рис. 2. Цей алгоритм має визначати повну процедуру дій менеджера в системі, для запобігання шахрайських схем.

Цільове визначення логіко-функціональної складової інформаційної системи дозволяє спроектувати функціональні особливості проекту. Для цього застосовують стандарти проектування на мові UML з розробкою діаграми варіантів використання, рис. 3.

Головним функціональним діючим актором системи є – менеджер з кредитування, який використовує інформаційну систему для перевірки доданих даних клієнтом, та створення кредитної справи згідно визначеної форми. До головних функціональних складових компонентів діаграми варіантів відносять:

– Міжнародне кредитування – функція програмної системи яка аналізує кредитні домовленості в міжнародній валюті. Є спеціалізованим програмним рішенням яке допомагає більш уважно організувати кредитне домовлення.



Рис. 2. Діаграма станів роботи менеджера з блокуванням клієнта

- Перевірка клієнта – функція пошуку клієнта та наявності в нього поточних, завершених кредитів, або чи є в нього борги.
- Чорний перелік користувачів – дозволяє сортувати клієнтів, які мають заборгованості та не сплачують кредитні домовлення.
- Умови кредитування – набір правил які дозволяють отримати кредит в залежності від обраних функціональних можливостей.
- Залогове майно – опис майна яке додається до кредитної умови та є гарантованою заставою фінансової позики. Може включати в себе набір вхідної інформації з детальним описом.
- Запит до інших кредитних установ – функція перевірки клієнтів фінансової установи на наявність кредитних контрактів в інших банках. Це дозволяє підвищити надійність формування кредитних домовлень та створення спеціалізованих скорингових моделей.
- Видати кредит – функція формування кредитної справи та договору згідно сформованої інформації. Виконується разом з модулем прийняття рішення.
- Вимоги кредитування – опція контролю різних варіантів кредитування в залежності від обраних пропозицій. Може суттєво впливати на вибір системи в залежності від обраного результату.
- Причина кредитування – дозволяє системі прийняття рішень більш детально аналізувати поточну інформацію щодо кожного окремого випадку кредитування та збільшувати точність формування кредитної справи.
- Поточні кредити – можливість перегляду усіх поточних кредитних справ та вартості наданих послуг різними менеджерами фінансової установи. Зберігає в собі гнучку систему фільтрації та перегляду інформації.
- Візуалізація платежів – дозволяє будувати графіки динаміки виплат та сформованих кредитних доручень згідно роботи фінансової установи.
- Боржники – виведення переліку з баз даних інформації стосовно клієнтів які не виконують фінансові доручення та не сплачують контракти.
- Динаміка виплат – можливість відстежувати та сортувати усі платежі від клієнтів установи.
- Додати запит до виконавчої служби – формування листа до виконавчої служби країни з доданням персональної інформації клієнта, який відмовляється виконувати рішення кредитного домовлення.

- Додати персональні дані – функція завантаження графічних матеріалів до персональної інформації користувачів.

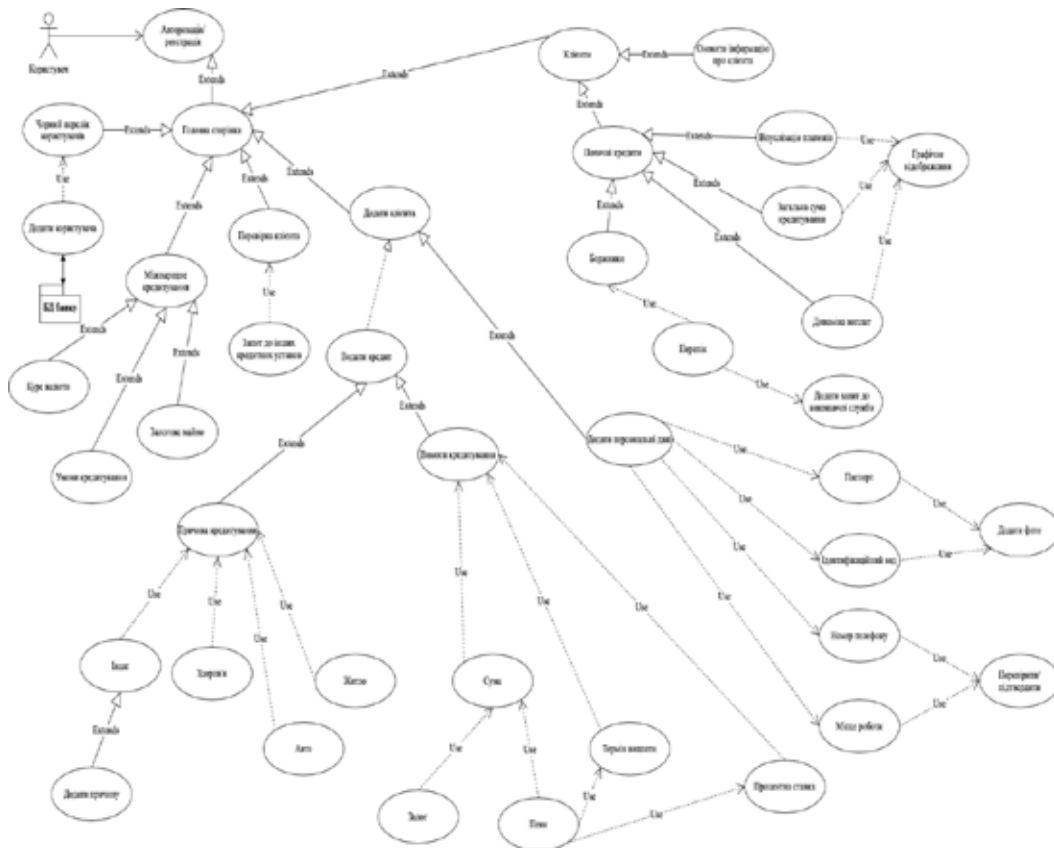


Рис. 3. Діаграма варіантів використання інформаційної системи

Проектування бази даних є важливою складовою будь-якого програмного проекту. Серед розглянутих баз даних, було обрано PostgreSQL як оптимальне поєднання для сучасних web-проектів. Таке рішення підтримує добру масштабованість та адаптацію для роботи з різними вимогами від клієнтів та середовища розробки. Для візуалізації спроектованих таблиць, застосовано нотацію проектування UML, рис. 4. До переліку таблиць, спроектованих в інформаційній системі відносяться наступні:

- Manager – таблиця зберігає інформацію кожного з користувачів програмного засобу та його облікові дані.
- Client – таблиця зберігає перелік інформації щодо кожного клієнта програмного засобу. До такої інформації відносяться персональна інформація, паспортні дані, кредитні домовлення тощо.
- Category – таблиця розподілу користувачів в системі, які дозволяють формувати спеціалізовані переліки та контролювати процес створення кредитних домовлень.
- Payment – таблиця кредитування та усієї інформацію, яка пов'язана з роботою фінансової організації.
- Loan – таблиця яка дозволяє зберігати інформацію щодо переліків клієнтів, кредитні справи яких не були розглянуті чи відмовлені в залежності від різних причин.
- Bank – таблиця яка поєднує роботу спеціалізованих таблиць електронного документообігу, які впроваджені в програмну систему.
- Graduator – таблиця яка зберігає робочі контракти користувачів та клієнтів програмної системи.

Проаналізувавши засоби розробки, було визначено що оптимальним рішенням для розробки сучасного програмного засобу, необхідно застосовувати мову розробки JavaScript та набору технологічних рішень для візуальної частини проекту: React, Apollo Client, styled-components. Для серверної частини використовуються: Postgres, Prisma2, GraphQL. На основі такого стеку технологічних рішень, формується діаграма класів до інформаційної системи, рис. 5.

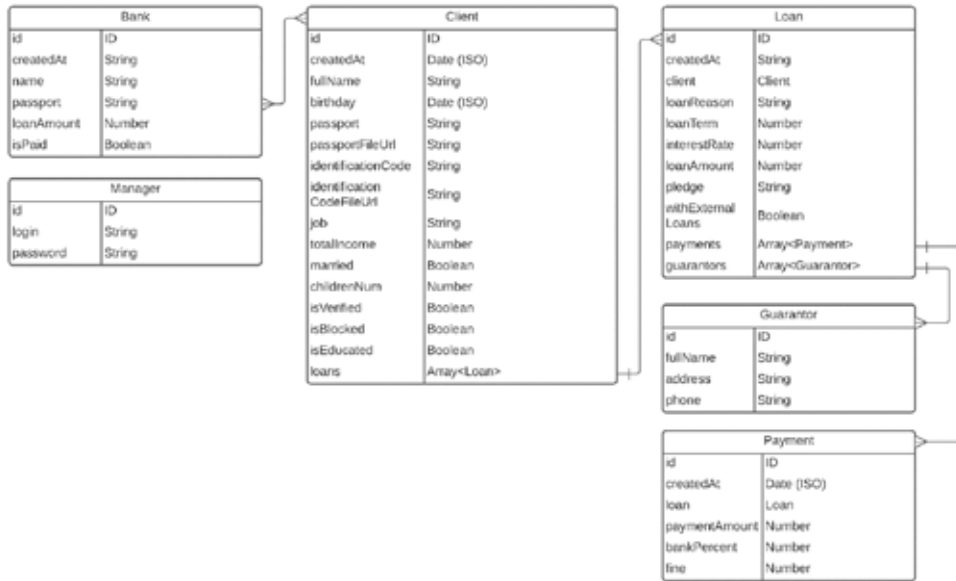


Рис. 4. Діаграма бази даних

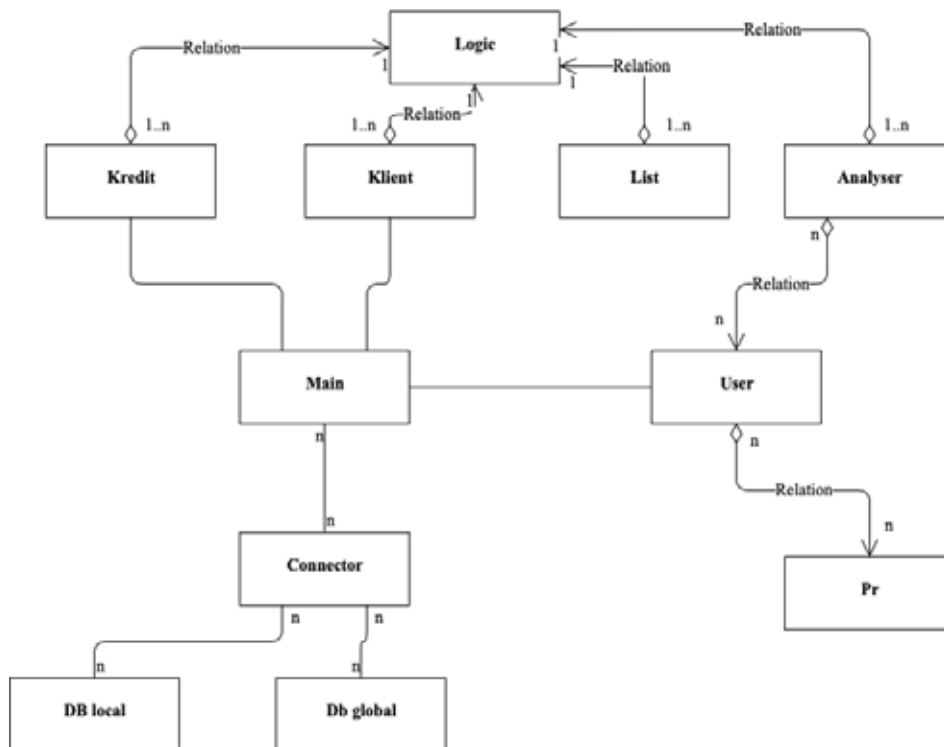


Рис. 5. Діаграма класів інформаційної системи

Кожен з класів має своє функціональне значення та використовується в якості складової інформаційної системи. При розробці було сформовано наступні класи:

- Logic – клас який поєднує в собі набір математичних та логічних операцій, необхідних для прийняття рішення.
- Kredit – клас, необхідний для формування клієнтських доручень та створень кредитних угод згідно сформованої документації.
- Klient – клас який поєднує в собі усі функціональні дії програмного засобу пов'язані з клієнтами фінансової установи.
- List – клас який функціонує з клієнтами та дозволяє формувати переліки клієнтів в залежності від типів обраної кредитної послуги.

- Analyser – математична система аналізу кредитних домовлень, яка дозволяє менеджерам приймати рішення в залежності від різних наборів вхідної інформації.
- Main – головний клас, який поєднує усі функціональні можливості інших класів та дозволяє їх ієрархічно пов'язувати між собою.
- User – клас який відповідає за роботу клієнта в програмній системі.
- Connector – клас який відповідає за синхронізацію інформації між програмним засобом та базами даних.
- DB local – клас який відповідає за роботу з локальною базою даних з використанням спеціалізованого програмного коннектора.
- DB global – клас який відповідає за роботу з глобальними базами даних та використовує в роботі спеціалізований програмний коннектор.
- Pr – клас який дозволяє заносити графічну інформацію до програмної системи та оновлювати персональні дані кожного клієнта системи.

Спроектована та розроблена інформаційна система має функціональну інтерфейсну частину яка оптимізована для комфортної роботи менеджера з кредитування. Головні об'єкти сформовані у вигляді глобальної таблиці з інформацією, яка включає в себе прив'язування інформації до дати створення картки клієнта та поточного статусу його картки.

Для кожної кредитної справи формується персоналізована діаграма виплат, на якій можливо переглянути динаміку повернення кредиту до установи. В залежності від доданої інформації щодо платежів, динаміка перераховується з урахуванням усіх особливостей кредитної справи та конкретного кредитного випадку. На рис. 6 відображено розроблену робочу сторінку менеджера з кредитування.

Создан	ФИО Клиента	Срок выплаты	Сумма кредита	Процентная ставка	Статус
2020-11-04	Зинovyа Дмитрий Германович	3 мес	8000 UAH	7%	Активный
2020-11-03	Алмyлов Альберт Юрьевич	3 мес	8000 UAH	7%	Завершенный
2020-11-02	Зинovyа Дмитрий Германович	12 мес	27000 UAH	7%	Активный
2020-11-01	Зинovyа Дмитрий Германович	12 мес	12000 UAH	7%	Должник

Рис. 6. Робоча сторінка менеджера з кредитування

Розроблена система оптимізована для роботи в фінансових установах з кредитування та дозволяє швидко перевіряти персональні дані клієнтів, за рахунок використання сучасних методів скорингового аналізу даних.

Висновки та перспективи подальших досліджень. Розроблена програмна система може використовуватися як один з заходів в прийнятті рішень при формуванні кредитних доручень, або як система оцінки ефективності роботи менеджерів кредитних спілок в фінансових установах. За рахунок використання гнучких алгоритмів розробки та динамічній можливості розширювати функціональні особливості системи, можлива адаптація розробленої системи в інші напрямки роботи фінансових установ.

Система побудована в якості активного помічника, який верифікує запити від менеджерів, аналізує фінансову заборгованість в інших фінансових установах, а також дозволяє спрощувати процес заповнення персональної інформації користувачів. Розроблена система може використовуватися в якості незалежного програмного рішення, або в якості окремого компонента в існуючій програмній системі.

Список використаних джерел:

1. Алешин В.А., Рудаева О.О. Кредитный скоринг как инструмент повышения качества банковского риск-менеджмента в современных условиях. *Terra economicus*. 2012. Т. 10. № 2. Ч. 3. С. 27–30.
2. Клейнер Г.Б., Коробов Д.С. История современного кредитного скоринга. *Проблемы региональной экономики*. 2012. № 17. С. 49–62.
3. Лукашевич Н.С. Об автоматизации кредитного процесса в банке. *Актуальные вопросы современной науки*. 2014. № 33. С. 132–140.
4. Снайдер Д., О'Брайен Т. Скоринг при кредитовании малого и среднего предпринимательства. *Деньги и кредит*. 2011. № 10. С. 59–63.
5. Хачатурян А.Г. Методы оценки кредитоспособности заемщика. *Фундаментальные и прикладные исследования кооперативного сектора экономики*. 2013. № 4. С. 102–105.
6. Алексейчук Н.Е. Экономическая теория. М. : Юрайт, 2012. 516 с.
7. Ачаповская М.З. Экономическая теория. Минск : ФУАинформ, 2010. 431 с.
8. Баликов В.З. Общая экономическая теория. М. : Омега, 2010. 685 с.
9. Земцов А.А., Осипова Т.Ю. Кредитный скоринг. Косвенный метод оценки богатства домашних хозяйств. *Вестник Томского Государственного Университета*. 2008. № 2(3). С. 17–38.
10. Глинкина Е.В. Кредитный скоринг как инструмент эффективной оценки кредитоспособности. *Финансы и кредит*. 2011. № 16(448). С. 43–47.

References:

1. Aleshin, V.A., Rudaeva, O.O. (2012). Credit scoring as a tool for improving the quality of banking risk management in modern conditions. *Terra economicus*, vol. 10, no. 2, p. 3, pp. 27–30.
2. Kleiner, G.B., Korobov, D.S. (2012). The history of modern credit scoring. *Regional economy problems*, no. 17, pp. 49–62.
3. Lukashevich, N.S. (2014). On the automation of the credit process in the bank. *Topical issues of modern science*, no. 33, pp. 132–140.
4. Snyder, D., O'Brien, T. (2011). Scoring in lending to small and medium-sized businesses. *Money and credit*, no. 10, pp. 59–63.
5. Khachaturian, A.G. (2013). Methods for assessing the creditworthiness of the borrower. *Fundamental and applied research of the cooperative sector of the economy*, no. 4, pp. 102–105.
6. Alekseychuk, N.E. (2012). *Economic theory*. Moscow:Yurayt.
7. Achapovskaya, M.Z. (2010). *Economic theory*. Minsk: FUAinform,.
8. Balikoev, V.Z. (2010). *General economic theory*. Moscow: Omega.
9. Zemtsov, A.A., Osipova, T.Yu. (2008). Credit scoring. An indirect method for assessing household wealth. *Bulletin of Tomsk State University*. 2008, no. 2(3), pp. 17–38.
10. Glinkina, E.V. (2011). Credit scoring as a tool for effective assessment of creditworthiness. *Finance and credit* 2008, no. 16(448). pp. 43–47.

НАУКОВЕ ВИДАННЯ

**ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ
ТА СУСПІЛЬСТВО**

**INFORMATION TECHNOLOGY
AND SOCIETY**

**ВИПУСК 2
ISSUE 2**

2021

Коректура
Ірина Чудеснова

Комп'ютерна верстка
Наталія Кузнєцова

Формат 60x84/8. Гарнітура Cambria.
Папір офсет. Цифровий друк. Ум. друк. арк. 10,93. Замов. № 1021/386. Наклад 300 прим.

Видавництво і друкарня – Видавничий дім «Гельветика»
65101, Україна, м. Одеса, вул. Інглєзі, 6/1
Телефон +38 (048) 709 38 69, +38 (095) 934 48 28, +38 (097) 723 06 08
E-mail: mailbox@helvetica.ua
Свідоцтво суб'єкта видавничої справи
ДК No 6424 від 04.10.2018 р.