НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
НАВЧАЛЬНО-НАУКОВИЙ КОМПЛЕКС
«ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ»
НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

МІЖНАРОДНИЙ НАУКОВО–ТЕХНІЧНИЙ ЖУРНАЛ

№ 3 2023

ЗАСНОВАНО У ЛИПНІ 2001 р.

АДРЕСА РЕДАКЦІЇ:

03056, м. Київ,
просп. Перемоги, 37, корп. 35,
ННК «ІПСА» КПІ ім. Ігоря Сікорського
Тел.: **204-81-44**; факс: **204-81-44**
E-mail: journal.iasa@gmail.com
http://journal.iasa.kpi.ua

У номері:

NATIONAL ACADEMY OF SCIENCES OF UKRAINE
EDUCATIONAL AND SCIENTIFIC COMPLEX
«INSTITUTE FOR APPLIED SYSTEM ANALYSIS»
OF THE NATIONAL TECHNICAL UNIVERSITY OF UKRAINE
«IGOR SIKORSKY KYIV POLYTECHNIC INSTITUTE»

SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

INTERNATIONAL SCIENTIFIC AND TECHNICAL JOURNAL

№ 3 2023

IT IS FOUNDED IN JULY 2001

THE EDITION ADDRESS:

03056, Kyiv,
av. Peremogy, 37, building 35,
Institute for Applied System Analysis
at the Igor Sikorsky Kyiv Polytechnic Institute
Phone: **204-81-44**; Fax: **204-81-44**
E-mail: journal.iasa@gmail.com
http://journal.iasa.kpi.ua

In the issue:

Шановні читачі!

Навчально-науковий комплекс «Інститут прикладного системного аналізу» Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського» видає міжнародний науково-технічний журнал

«СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ».

Журнал публікує праці теоретичного та прикладного характеру в широкому спектрі проблем, що стосуються системних досліджень та інформаційних технологій.

Провідні тематичні розділи журналу:

Теоретичні та прикладні проблеми і методи системного аналізу; теоретичні та прикладні проблеми інформатики; автоматизовані системи управління; прогресивні інформаційні технології, високопродуктивні комп'ютерні системи; проблеми прийняття рішень і управління в економічних, технічних, екологічних і соціальних системах; теоретичні та прикладні проблеми інтелектуальних систем підтримання прийняття рішень; проблемно і функціонально орієнтовані комп'ютерні системи та мережі; методи оптимізації, оптимальне управління і теорія ігор; математичні методи, моделі, проблеми і технології дослідження складних систем; методи аналізу та управління системами в умовах ризику і невизначеності; евристичні методи та алгоритми в системному аналізі та управлінні; нові методи в системному аналізі, інформатиці та теорії прийняття рішень; науково-методичні проблеми в освіті.

Головний редактор журналу — ректор Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», академік НАН України Михайло Захарович Згуровський.

Журнал «Системні дослідження та інформаційні технології» включено до переліку фахових видань ВАК України.

Журнал «Системні дослідження та інформаційні технології» входить до таких наукометричних баз даних: Scopus, EBSCO, Google Scholar, DOAJ, Index Copernicus, реферативна база даних «Україніка наукова», український реферативний журнал «Джерело», наукова періодика України.

Статті публікуються українською та англійською мовами.

Журнал рекомендовано передплатити. **Наш індекс 23918.** Якщо ви не встигли передплатити журнал, його можна придбати безпосередньо в редакції за адресою: 03056, м. Київ, просп. Перемоги, 37, корп. 35.

Завідувачка редакції **С.М. Шевченко**
Редакторка **Р.М. Шульженко**
Молодша редакторка **Л.О. Тарин**
Комп'ютерна верстка, дизайн **А.А. Патіохи**
Свідоцтво про реєстрацію КВ № 23234–13074 ПР від 22.03.2018 р.

Підписано до друку 29.09.2023. Формат 70×108 1/16. Папір офс. Гарнітура Times.
Спосіб друку – цифровий. Ум. друк. арк. 14,411. Обл.-вид. арк. 28,56. Наклад 100 пр. Зам. № 11/04

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Свідоцтво про державну реєстрацію: ДК № 5354 від 25.05.2017 р.
просп. Перемоги, 37, м. Київ, 03056.
ФОП Пилипенко Н.М., вул. Мічуріна, б. 2/7, м. Київ, 01014.
Виписка з Єдиного державного реєстру № 2 070 000 0000 0214697 від 17.05.2019 р.,
тел.. (044) 361 78 68.

Dear Readers!

Educational and Scientific Complex «Institute for Applied System Analysis» of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» is published of the international scientific and technical journal

«SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES».

The Journal is printing works of a theoretical and applied character on a wide spectrum of problems, connected with system researches and information technologies.

The main thematic sections of the Journal are the following:

Theoretical and applied problems and methods of system analysis; theoretical and applied problems of computer science; automated control systems; progressive information technologies, high-efficiency computer systems; decision making and control in economic, technical, ecological and social systems; theoretical and applied problems of intellectual systems for decision making support; problem- and function-oriented computer systems and networks; methods of optimization, optimum control and theory of games; mathematical methods, models, problems and technologies for complex systems research; methods of system analysis and control in conditions of risk and uncertainty; heuristic methods and algorithms in system analysis and control; new methods in system analysis, computer science and theory of decision making; scientific and methodical problems in education.

The editor-in-chief of the Journal is rector of the National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», academician of the NASU Michael Zaharovich Zgurovsky.

The articles to be published in the Journal in Ukrainian and English languages are accepted. Information printed in the Journal is included in the Catalogue of periodicals of Ukraine.

СИСТЕМНІ ДОСЛІДЖЕННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

3 • 2023

ЗМІСТ

SYSTEM RESEARCH AND INFORMATION TECHNOLOGIES

3 • 2023

CONTENT

ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ ПРОБЛЕМИ І МЕТОДИ СИСТЕМНОГО АНАЛІЗУ

UDC 331.101.262
DOI: 10.20535/SRIT.2308-8893.2023.3.01

ACADEMICIAN GLUSHKOV'S LEGACY: HUMAN CAPITAL IN THE FIELD OF CYBERNETICS, COMPUTING, AND INFORMATICS AT IGOR SIKORSKY KYIV POLYTECHNIC INSTITUTE

M.Z. ZGUROVSKY

Abstract. The role of academician Viktor Glushkov in the creation of scientific schools in the field of cybernetics, computing, and informatics at the Igor Sikorsky Kyiv Polytechnic Institute, which became a powerful national center for training specialists in this field, is considered. The significant influence of academician Hlushkov's ideas on the formation of generations of scientists, who to this day continue to build a digital society in Ukraine and far beyond its borders, is shown.

Keywords: academician Viktor Glushkov, cybernetics, computing, informatics, digital society.

INTRODUCTION

Cybernetics as a distinct science was formed in the late 1940s. Its formation and development is closely related to the creation of electronic computing machines (ECM) and their expansion in the USA, Great Britain, and the USSR in the late 1940s and early 1950s. In the Soviet Union, a significant impetus for the further development of cybernetics was the creation of a small electronic computing machine (SECM) in Kyiv in 1947–1948 under the supervision of Academician S.O. Lebediev [1].

The government's decision of 1955 to create a new academic institute in 1957 – the Computational Center of the Academy of Sciences of the Ukrainian SSR – was the consequence of this achievement. The activity of this Center, with its subsequent transformation into the Institute of Cybernetics in 1962, became the initial stage of the development of cybernetics, computer science and informatics based on the combination of fundamental and applied research.

Being a leader of a new scientific field, Academician V.M. Glushkov perceived cybernetics as a science of management using computer technology, communication means, and mathematical methods of real-time processing of large arrays of information of arbitrary nature and content. According to Academician V.M. Glushkov, young scientists and specialists of the senior generation had to actively participate in these studies.

That is why, on the initiative of Academician Glushkov, in 1960, the Kyiv Polytechnic Institute (KPI) started training personnel in computing. And in 1969 at Taras Shevchenko Kyiv State University on the initiative of Viktor Glushkov and the dean of the Faculty of Mechanics and Mathematics, prof. I.I. Lyashko, the faculty of cybernetics was opened, which began training specialists in computational mathematics. Later, the training of personnel in computer technology and computational mathematics was started in a few other universities of Ukraine.

KPI'S CONTRIBUTION TO THE TRAINING OF SPECIALISTS IN THE FIELD OF CYBERNETICS, COMPUTER SCIENCE AND INFORMATICS

On March 16, 1960, the first Department of Computer Science (CS) at the KPI was detached from the Department of Automation and Telemechanics (Head of the Department Prof. Y.I. Greben, Dr.Sci. in Engineering). The first Head of the Department of Computer Science was O.G. Ivakhnenko, then the Associate Member of the Academy of Sciences of the Ukrainian SSR.

O.G. Ivakhnenko can truly be called one of the pioneers of Ukrainian cybernetics. In 1959, he published the first monograph on cybernetics in the Soviet Union [2]. Several generations of future specialists in cybernetics and informatics studied the materials of this monograph. O.G. Ivakhnenko was always a generator of ideas, which he generously shared with his students. His scientific school includes 20 doctors and more than 100 candidates of sciences, whom he trained personally. Ivakhnenko's first students were Acad. V.M. Kuntsevich, professors V.I. Ivanenko, V.I. Kostyuk, O.A. Pavlov, Yu.P. Zaichenko and other outstanding scientists who later started their own scientific schools and made a significant contribution to the development of cybernetics.

Some of Ivakhnenko's ideas of the 1960s and 1970s were several decades ahead of the state of science at that time. Thus, the well-known Group method of data handling (GMDH) [3] became the first method of deep learning, which was ahead of classical developments in the field of intelligent data analysis, forecasting, modeling of complex systems, optimization, pattern recognition and artificial intelligence by several decades.

In 1960, the International Congress of the International Federation of Automatic Control (IFAC) was held in Ukraine [4]. O.G. Ivakhnenko, who was one of the organizers of the event, invited the founder of cybernetics, the outstanding American scientist Norbert Wiener to come to Kyiv and take part in the Congress. Norbert Wiener gratefully accepted this invitation, spoke at the IFAC Congress and discussed current problems of cybernetics with O.G. Ivakhnenko in an informal and positive environment.

Professor K.G. Samophalov, Associate Member of the Academy of Sciences of Ukraine, headed the Department of Computer Science at the KPI from 1961 to 1990. Later, Prof. G.M. Lutsky, Dr.Sci. in Engineering, continued his successful activity. Over 50 years of their fruitful work, almost 50 Dr.Sci. and more than 350 PhDs, more than 8.000 specialists have been trained; they have formed several generations of specialists in the field of computer science, cybernetics, and informatics.

A new Department of Engineering Cybernetics of KPI, established in 1969 by Prof. V.I. Kostyuk, Dr.Sci. in Engineering, played a significant role in the further training of specialists in this field. Specialists in engineering cybernetics and robotics were trained at this Department. In total, 170 Dr. Sci., 110 PhDs, 6.200 specialists in engineering cybernetics, ACS and robotics were trained at the EC Department.

During the heyday of research in the field of cybernetics, the abovementioned Department of Automation and Telemechanics, headed by Prof. A.A. Krasnoproshyna, Dr.Sci. in Engineering, continued to work actively. Somewhat later, it was renamed the Department of Automation and Control in Engineering Systems (ACES). A scientific school was formed at the department, which conducted research in the field of industrial automation and data transmission; professors Yu.P. Zhurakovsky, G.B. Serdyuk, L.M. Kompanets, Dr.Sci. in Engineering, and other researchers were its members. Later, the ACES Department focused on research in the field of informatics, paying the main attention to the development and operation of information systems and technologies. During the entire period of the Department's existence, three Dr.Sci. and twenty PhDs, and over 5 thousand specialists in industrial automation and information systems and technologies were trained.

On Order of the Ministry of Higher and Secondary Special Education of the Ukrainian SSR No. 278 of October 29, 1985 "On Changes in the Structure of Higher Educational Institutions of the Ministry of Higher Education of the Ukrainian SSR in 1985", the above-mentioned departments were merged into the Faculty of Informatics and Computer Science (FICS). Prof. A.A. Krasnoproshyn, Dr.Sci in Engineering, became the first dean of the Faculty.

Since the early 1970s, the era of practical development of automated control systems (ACS) began, whose ideological foundations were formulated by A.I. Kitov and V.M. Glushkov in 1959–1962 [5]. As a result, in 1973, for the first time in Ukraine, the training of specialists in the specialty of ACS was initiated at the Department of Engineering Cybernetics.

Leading professors of the Department of Engineering Cybernetics: Acad. O.G. Ivakhnenko, Acad. V.I. Skurykhin, Acad. V.M. Kuntsevych, Prof. V.I. Kostyuk, Prof. V.V. Azhogin, Prof. Yu.P. Zaichenko, Prof. O.A. Pavlov, Assoc. Prof. M.K. Pechurin, Assoc. Prof. V.M. Tomashevsky and others participated in the formation of this new and future-oriented specialty.

It is worth noting the significant role of Associate Professor P.I. Akinin in the organization and formation of a new Department of Automated Production Control Systems (APCS), by detaching it from the Department of Engineering Cybernetics in 1978. Subsequently, Professor O.A. Pavlov headed the Department of APCS, later the Department of Automated Data Processing and Control Systems (ADPCS), who made a significant contribution to the development of the ACS specialty over the years. The Associate Professor V.O. Tykhonov was the Deputy Head of the Department; he assured the activity of the Department and development of methodological support of specialty at a high level. At that time, Associate Professors S.F. Telenyk and S.M. Grysha worked fruitfully at the Department and made a significant contribution to the development of the ACS specialty. The department trained 8 Dr.Sci., 50 PhDs, over 6,000 specialists in the ACS sphere.

The establishment of the Branch of the Department of ADPCS at the Cybernetic Center of National Acasdemy of Sciences of Ukraine, where Academicians of the NAS of Ukraine I.V. Sergienko (Head of the Branch), V.K. Zadiraka, A.O. Chykriy worked, as well as work of Academicians of the NAS of Ukraine N.Z. Shor, A.A. Letichevsky, V.I. Skurikhin at the Department of ADPCS, contributed to highly creative and innovative level of learning [6]. The transfer of the experience of scientific and practical work by prominent scientists of Ukraine, followers of V.M. Glushkov's work, to students, remained in their memory forever and significantly contributed to their professional growth.

The graduation of experts in the ACS sphere lasted for almost 20 years. Over the years, the foundation of the scientific school in the field of engineering cybernetics, ACS and informatics has been created; the leading scientists of the university, in particular, Professors V.I. Kostyuk, K.G. Samofalov, V.V. Azhogin, M.Z. Zgurovsky, O.A. Pavlov, G.M. Lutsky, Yu.P. Zaichenko made a significant contribution to it.

Scientific research in the field of information technologies in KPI became the basis for the development of the latest systems of automation, promising project and technological solutions. Scientists of the Cybernetic Center of the NAS of Ukraine: I.V. Sergienko, Director of the Institute of Cybernetics and Head of the Branch of the Department of Automated Data Processing and Control Systems, Academicians of the NAS of Ukraine I.M. Kovalenko, V.M. Kuntsevych, V.S. Mykhalevych, O.V. Palagin, Associate Member of the NAS of Ukraine B.M. Malynovsky, Prof. V.P. Derkach, Dr.Sci. in Engineering, and others made a significant contribution to the formation of the scientific school in KPI.

In 1990, the Faculty of Applied Mathematics (FAM) was formed at the KPI on the basis of the departments of Applied Mathematics, Mathematical Methods of Systems Analysis and Specialized Computing Systems, by its detaching from the Faculty of Informatics and Computer Science (FICS). Famous scientists M.Z. Zgurovsky, I.M. Kovalenko, K.G. Samofalov, Yu.L. Daletsky, Ye.M. Vavilov, O.A. Molchanov, O.A. Pavlov, V.P. Tarasenko participated in the creation of the FAM and its formation in different years.

The first Dean, the founder of the faculty was I.M. Kovalenko, Academician of the NAS of Ukraine, a well-known specialist in the field of probability and random processes theory. The next deans of the faculty – Prof. O.A. Molchanov and Prof. I.A. Dychka – continued the work of I.M. Kovalenko.

Currently, the FAM includes three Departments: Applied Mathematics; System Programming and Specialized Computer Systems; Computer Systems Software.

The Department of Applied Mathematics (AM) was established in 1973 by its detaching from the Department of Computer Science of the FICS. Prof. Ye.M. Vavilov was the first Head of the Department. Further, for 36 years, Professor O.A. Molchanov, Dr.Sci. in Engineering, was the Head of the Department; nowadays, Professor O.R. Chertov, Dr.Sci. in Engineering, heads the Department.

The Department of AM trains specialists in the specialty 113 Applied Mathematics, Educational Program "Data Science and Mathematical Modeling", which covers the following areas: development of mathematical methods for Data Mining and Big Data; Machine Learning methods; Computer information technologies for knowledge acquisition using Computational Intelligence methods.

The Department of System Programming and Specialized Computer Systems (SPSCS) was established in 1990 by its detaching from the Department of

Computer Science (FICS). Professor V.P. Tarasenko, Dr.Sci. in Engineering, became the first Head of the newly created Department. Nowadays, Professor V.O. Romankevych, Dr.Sci. in Engineering, heads the Department of SPSCS.

The Department of SPSCS traines specialists in the specialty 123 "Computer Engineering", Educational Program: System Programming and Specialized Computer Systems, which covers the areas of designing, development, testing, implementation and operation of computer systems, networks and their components; designing and modernization of specialized computer systems, their object and functional orientation; designing, development, testing, implementation and support of system software.

The Department of Computer Systems Software Support (CSSS) was established in 2009 by its detaching from the Department of System Programming and Specialized Computer Systems. Professor I.A. Dychka, Dr.Sci. in Engineering, became the first Head of the Department. Nowadays, Associate Professor Ye.S. Sulema, Dr.Sci. in Engineering, heads the Department.

The Department of CSSS trains specialists in the specialty 121 "Software Engineering", Educational Program "Software for multimedia and information retrieval systems", which covers the following areas: development and use of object-oriented engineering methods of projecting and designing complex software systems; development, testing, implementation and support of software for intelligent search systems (Information Retrieval, Data Mining, Text Mining); designing software methods for efficient processing, storage and protection of large data, including multimedia data (Multimedia and Mulsemedia).

Since 2010, the FAM is a member of the European Consortium of Mathematics in Industry, ECMI. The faculty trained 11 Dr.Sci., 46 PhDs and over 4,000 specialists in computer engineering, software engineering, data analysis and mathematical modeling.

The FAM graduates hold leading positions in national and foreign IT companies, including scientific subdivisions of Global Logic, Epam, Data Art, Ajax Systems, Samsung, Microsoft, etc.

COOPERATION OF KPI AND NAS OF UKRAINE IN THE SPHERE OF TRAINING SPECIALISTS IN THE FIELD OF CYBERNETIC SECURITY AND SYSTEMS ANALYSIS

One of the bright pages in the history of the KPI scientific school of cybernetics, computer science and informatics is the development of applied mathematics, information and cyber security at the Physical and Technical Institute (PhTI) of KPI [7]. The Institute was founded in 1995 on the initiative of the President of the National Academy of Sciences of Ukraine, Academician B.E. Paton and then Minister of Education of Ukraine, Academician M.Z. Zgurovsky in order to train academic personnel in the field of applied physics and applied mathematics for the National Academy of Sciences of Ukraine. O. M. Novikov, Associate Member of the NAS of Ukraine, is the PhTI Director during all these years.

It was the educational focus on applied mathematics that served as the basis for the establishment of two new departments at the PhTI in 1999–2000, namely, Information Security (O.M. Novikov, Associate Member of the National Academy of Sciences of Ukraine was the first Head of the Department) and Mathematical Methods of Information Protection (Prof. M.M. Savchuk, Dr.Sci. in Physics and Mathematics was the first Head of the Department), the new specialty "Security of information and communication systems" and the specialization "Mathematical methods of cryptanalysis" within the specialty "Applied Mathematics".

The group of scientists headed by O.M. Novikov, Dr.Sci. in Engineering, trained in the scientific school of Academician M.Z. Zgurovsky, became the basis of the Department of Information Security, and the scientists from the scientific school of cryptography of the Institute of Cybernetics of the National Academy of Sciences of Ukraine, headed by Academician I.M. Kovalenko, formed the basis of the Department of Information Protection. Prof. Ye.A. Machusky, Dr.Sci. in Engineering, Prof. O.Ye. Arkhipov, Dr.Sci. in Engineering, Prof. D.V. Lande, Dr.Sci. in Engineering, prof. M.M. Savchuk, Associate Member of the NAS of Ukraine, Prof. M. Yu. Kuznetsov, Associate Member of the NAS of Ukraine, Prof. A.B. Kachynsky, Dr.Sci. in Engineering, Prof. A.M. Kudin, Dr.Sci. in Engineering, and others made a significant contribution to the development and strengthening of the sphere of information and cyber security in KPI.

The specialty "Informatics" within the framework of the "Applied Mathematics" educational sphere was established in the first years of the PhTI formation, which later gained an impetus within the educational programs "Mathematical methods of cryptographic protection of information", "Mathematical methods of modeling, pattern recognition and computer vision".

In 2021, the Department of Mathematical Modeling and Data Analysis was formed within the PhTI structure (Prof. N.M. Kussul, Dr.Sci. in Engineering is the Head of the Department). The scientists of the Institute of Space Research of the Cybernetic Center of the National Academy of Sciences of Ukraine, scientists and professors of the PhTI participated in the foundation of the Department.

Employers consistently give a high assessment of the qualifications of graduates in applied mathematics, information and cyber security of scientific schools of PhTI. In particular, according to the rating of the "Forbes Ukraine" magazine, PhTI of KPI was recognized as the best faculty of Ukraine in the nomination "STEM disciplines" in 2021.

For more than 20 years of activity, the teams of the Department of Information Security and the Department of Mathematical Methods of Information Protection, and later the Department of Mathematical Modeling and Data Analysis of PhTI, trained more than 2,000 specialists in applied mathematics, information and cybernetic security, 10 PhD and doctoral theses were defended. Teams of scientists received two State Prizes of Ukraine in the field of science and technology and the Prize of the President of Ukraine for young scientists; three PhTI scientists, Prof. M.M. Savchuk, Prof. M.Yu. Kuznetsov and Prof. O.M. Novikov were elected Associate Members of the National Academy of Sciences of Ukraine.

In September 1988, on the initiative of Academician V.S. Mykhalevych, Professors M.Z. Zgurovsky and Yu.L. Daletsky, and with the support of the Minister of Higher and Secondary Special Education V.D. Parkhomenko, the training of system analysts was started in Ukraine. For this purpose, the Department of Mathematical Methods of Systems Analysis (MMSA) was detached from the Department of Applied Mathematics of the Faculty of Control Systems (FCS) of KPI as a department of targeted personnel training for institutes of the Academy of Sciences of the Ukrainian SSR. The Department trained students in the specialty "Applied Mathematics" with the specialization "Systems Analysis and Control". Prof. M.Z. Zgurovsky, Dr.Sci in Engineering, headed the Department.

The group of Professors of the Department of Applied Mathematics of KPI: Yu.L. Daletsky, Yu.V. Bohdansky, V.G. Bondarenko, S.M. Paramonova, N.D. Tsvyntarna, and scientific group of Professor M.V. Zgurovsky whose members were V.D. Romanenko, O.M. Novikov, O.S. Makarenko, O.M. Selin, P.I. Bidyuk, Yu.M. Selin formed the nucleus of the MMSA Department.

The MMSA Department became the first and basic department in Ukraine for training specialists in the field of systems analysis and made a significant contribution to the development of the new "Systems Analysis" specialty.

For the first time in Ukraine, new specialties "Intelligent decision-making systems", "Decision-making systems and methods", "Social informatics", "Systems analysis", "Computer science" were introduced into the educational process of the MMSA Department, and new educational and professional programs "Systems Analysis and Control", "Systems Analysis of the Financial Market", "Systems and Methods of Artificial Intelligence" were developed and accredited.

Professors P.I. Bidyuk, Yu.L. Daletsky, V.Ya. Danilov, Yu.P. Zaichenko, M.Z. Zgurovsky, P.O. Kasyanov, V.Ya. Melnyk, N.D. Pankratova, V.D. Romanenko actively participated in the organization of the educational process, in the preparation of textbooks and teaching aids.

A series of new textbooks and monographs on systems analysis, data mining, and artificial intelligence by M.Z. Zgurovsky, N.D. Pankratova, and Yu.P. Zaichenko, were prepared and printed, they were used for more than 20 years to train system analysts in Ukraine and abroad.

On the initiative of A.I. Kukhtenko, Academician of the National Academy of Sciences of Ukraine, the Research Institute of Interdisciplinary Studies (RIIDS) was established at KPI in 1990 based on the research sector of the MMSA Department. The main focus of the RIIDS activity was to conduct large-scale exploratory and fundamental research of an interdisciplinary nature in such scientific areas as mathematical, physical, chemical and engineering methods of environmental protection; artificial intelligence systems; theory of fractals and its application in physics, chemistry, materials science; catastrophe theory and bifurcation, synergy and "chaos", and others.

36 Dr.Sci. and 105 PhDs, 4,300 specialists in the field of systems analysis and information technologies were trained at the MMSA Department during 35 years of its activity.

Researchers of the Department P.I. Bidyuk, Yu.P. Zaichenko, M.Z. Zgurovsky, P.O. Kasyanov, N.D. Pankratova, V.D. Romanenko became laureates of the State Prizes of Ukraine in the field of science and technology. M.Z. Zgurovsky and Yu.L. Daletsky were elected Academicians of the National Academy of Sciences of Ukraine for their fundamental developments in the field of science and technology, and V.S. Melnyk and N.D. Pankratova were elected Associate Members of the National Academy of Sciences of Ukraine.

In 1997, by Resolution of the Cabinet of Ministers of Ukraine No. 1351, dated December 2, 1997, the Institute of Applied Systems Analysis (IASA) under the dual subordination of the Ministry of Education and Science of Ukraine and the National Academy of Aciences of Ukraine, was established on the basis of the MMSA Department of KPI and the departments of “Numerical Optimization Methods” (Head of the department B.M. Pshenychny, Academician of the National Academy of Sciences of Ukraine) and “Applied Nonlinear Analysis” (Head of the Department Prof. V.I. Ivanenko, Dr.Sci. in Engineering) of the Institute of Cybernetics of the National Academy of Sciences of Ukraine [8].

The foundation of the Institute of dual subordination resulted from the implementation of the concept of science and education integration. The integration of educational and scientific areas in IASA provided an opportunity to carry out unique research and implement the methodology of training system analysts of a practical orientation. M.Z. Zgurovsky, Academician of the National Academy of Sciences of Ukraine, became the Director of the Institute, and later, its academic adviser. Since 2018, Prof. P.O. Kasyanov, Dr.Sci. in Physics and Mathematics, is the IASA Director.

The main goal of IASA foundation is to carry out innovative educational activities at various levels of higher education and areas of scientific research, to develop the methodology of systems analysis, mathematical methods and software tools for conducting comprehensive analysis, forecasting and solving the state’s social, economic, ecological, and technological problems, meeting social, economic and cultural needs of society, achievement of innovative development of the country.

Currently, the IASA educational and scientific structure includes three Departments: Mathematical Methods of Systems Analysis; System Design; Artificial Intelligence, which train highly qualified specialists in the specialties: 124 Systems Analysis and 122 Computer Sciences, and three scientific departments of the National Academy of Sciences of Ukraine: System Mathematics; Mathematical Methods of Systems Analysis; Applied Nonlinear Analysis; they perform fundamental and applied scientific research and participate in the educational process by delivering lectures, conducting practical classes, giving academic advises to graduate students, supervising bachelor’s and master’s theses. The Institute has subdivisions of pre-university and course training as well.

At the current stage, the Department of Mathematical Methods of Systems Analysis traines system analysts capable to design, create and operate complex man-machine aggregates for analysis, forecasting, control and design of dynamic processes in macroeconomic, technical, technological, security, environmental and financial systems. The MMSA graduates work in the analytical subdivisions of state administration bodies, special services, large companies, data centers, other institutions and enterprises, whose activities are based on modern system analytics.

The Department of System Design trains specialists in the modern theory and practice of building and operating intelligent computing environments (application of service-oriented computing and architectures (SOA and SOC), distributed grid, cloud, fog, and serverless computing in computer processing of big data and automated design; construction of multi-agent systems and infrastructures; use of SOC and SOA in the Internet of Things). Graduates are in demand by scientific centers, IT companies, enterprises involved in the support and maintenance of modern information systems and technological complexes.

The Department of Artificial Intelligence trains specialists in the field of artificial intelligence and intelligent data analysis, computer vision, text and voice information processing, intelligent medical diagnostic systems, and intelligent defense systems.

The Institute of Applied Systems Analysis publishes the international magazine "System Research and Information Technologies" [9], which is included in the register of specialized scientific publications of Ukraine under category "A", the published articles are indexed in the Scopus scientometric database. Staff members of the Institute are members of a number of editorial boards of periodical professional Ukrainian and foreign editions.

The UNESCO International Chair "Higher Technical Education, Applied Systems Analysis and Informatics" [10] operates on the IASA basis, whose main purpose is the implementation of a comprehensive research program in the field of higher technical education, methods of applied systems analysis, advanced information technologies and their application in the educational system, analysis of complex systems of various nature, in particular, socio-economic, ecological, technical and others.

The Institute conducts active international activities. According to the Resolution of the Presidium of the National Academy of Sciences of Ukraine dated 29/04/2015, No. 118 "About the distribution of responsibilities among the members of the Presidium of the National Academy of Sciences of Ukraine", the Scientific Director of IASA M.Z. Zgurovsky, Academician of the National Academy of Sciences, is responsible for the implementation of relations with the Committee on Data for Science and Technology (CODATA), United Nations Industrial Development Organization (UNIDO). He maintains relations with the "International Scientific Council" (ISC) together with the President of the National Academy of Sciences of Ukraine.

For more than a quarter of a century, IASA trained more than 6000 highly qualified system analysts who successfully work in Ukraine and 50 countries around the world. IASA scientific schools trained 2 academicians, 2 associate members of the National Academy of Sciences of Ukraine, more than 50 Dr.Sci. and 130 PhDs.

A powerful scientific base and personnel potential made it possible for KPI in cooperation with the National Academy of Sciences of Ukraine to create the Faculty of Physics and Mathematics (FMF) in 1996 [11]. The faculty was headed by an outstanding scientist, academician of NASU, Hero of Ukraine V.G. Baryakhtar. Since 2007, his case has been continued by Doctor of Technical Sciences, Prof. V.V. Vanin.

At different times, scientists with world-famous names worked here and headed the departments of fundamental training: academicians I.V. Skrypnyk, A.M. Samoilenko, V.I. Loktev, corresponding member of the National Academy of Sciences of Ukraine V.Ya. Valakh, professors A.V. Pavlov, F.P. Yaremchuk, V.V. Buldygin, Yu.I. Horobets, N.O. Virchenko, O.I. Klesov, V.M. Gorshkov, M.M. Kukharchuk, I.V. Beiko and others.

The faculty trains specialists in fundamental and applied problems in various branches of modern mathematics and physics — from the study of aerodynamics and hydrodynamics to the physics of the atomic nucleus and elementary particles using methods of mathematical and computer modeling, analysis and forecasting of complex processes and natural phenomena to practical applications of modern mathematical methods and tools in the field of economics, finance business and national security.

There are six departments within the FMF: general physics and modeling of physical processes; mathematical analysis and probability theory; general physics; mathematical physics and differential equations; descriptive geometry; engineering and computer graphics.

Over the 27 years of its work, FMF has trained 21 doctors, 53 candidates of science, 2285 specialists in modern areas of mathematics, physics, mathematical and computer modeling and their applications in various fields of science, national economy and national security of Ukraine. Scientists of the faculty have been awarded two state prizes in the field of science and technology.

In 2002, the educational and scientific Institute of Telecommunication Systems (ITS) was established in KPI [12] as a logical result of the development of the Department of Telecommunications, which, since 1993, trained specialists and designed telecommunication networks for large computing complexes and control facilities in ACS systems.

Nowadays, ITS includes three departments: Department of Electronic Communications and Internet of Things (ECIT); Department of Information Technologies in Telecommunications (ITT); and Department of Telecommunications (TC).

Educational and scientific work of the ITS is carried out in cooperation with scientific institutions and companies from Ukraine, USA, Great Britain, France, Germany, China, Poland, etc. Currently, the main ITS partners are the following large companies: Lifecell, ALCATEL-LUCENT, HUAWEI TECHNOLOGIES (China), Kyivstar, Ukrtelecom, Vodafone, Bankomzvyazok, Data Group, NOKIA, etc.

Over the years of its activity, the Institute of Telecommunication Systems has trained 12 Dr.Sci. and 42 PhDs, more than 2.500 specialists in the field of telecommunications; 15 ITS scientists were awarded the title of laureates of state prizes of Ukraine in the field of science and technology; Prof. M. Yu. Ilchenko, Dr.Sci. in Engineering, was elected academician of the National Academy of Sciences of Ukraine.

CONCLUSIONS

1. Academician Viktor Glushkov left an outstanding legacy in the field of training high-quality human capital in the field of cybernetics, computing and information science. Over the past 63 years, the legacy of the KPI has been replenished with three educational and scientific institutes, four faculties, and thirty-one departments. In this area, the KPI has trained more than 61.300 computer engineers, 195 Doctors of Sciense, 1.123 PhD's, four Academicians and seven Corresponding Members of the National Academy of Sciences of Ukraine. 59 KPI scientists were awarded by the title of Laureate of State Prizes of Ukraine in the field of science and technology.

2. Diligent and talented youth who studies and works in many universities, scientific institutions and leading companies of Ukraine and the world, is a follower of Academician Glushkov's achievments. KPI students made a significant contribution to strengthening the image of academician Glushkov's

scientific school at the national and international levels. They won more than 170 awards at Olympiads in mathematics, cyber security, informatics in Ukraine and abroad.

3. Among the significant achievements of students there are 41 awards (including 8 Grand Prix) at the World Mathematical Olympiads (IMC), the World Team Programming Championship (ICPC) according to the ACM version in London, Tokyo, Prague, Warsaw, Bucharest, Munich, Skopje (Macedonia), Blagoevgrad (Bulgaria), etc. Students O. Rybak, A. Mellit, S. Torba, M. Vlasenko, A. Gogoliev, K. Vedensky, B. Nagirnyak, B. Borysyuk, B. Baydenko, O. Slyusarenko, K. Marovetska, K. Matviyiv, V. Mykhaylovsky, Ye. Polishchuk and others were multiple winners of mathematics and informatics Olympiads in different years.

4. Since 2012, a series of victorious performances at the world championships among "white" hackers in the CTF version (capture the flag) of DCUA team supervised by Associate Professor M.I. Ilyin, Ph.D., was added to the bright victories of the students [13]. In 2016, the team won the title of World Champions among "white" hackers among more than 12 thousand participating teams.

5. Currently, students and followers of Academician Glushkov actively build a digital country and its economy, making scientific breakthroughs and training new generations of talented youth, increasing the defense capability of our state.

REFERENCES

1. *Encyclopedia of cybernetics in 2 volumes*, (in Ukrainian); editor: V.M. Glushkov (chief editor); Academy of Sciences of the Ukrainian SSR; Vol. 2. Kyiv: URE ed., 1973, 570 p.
2. A.G. Ivakhnenko, *Engineering cybernetics*, (in Russian). Kyiv: Technika Publ. house, 1959.
3. A.G. Ivakhnenko, "Group method of data handling – a competitor of the stochastic approximation method," *Automatics*, no. 3, pp. 58–72, 1968.
4. *IFAC – International Federation of Automatic Control*, (in Ukrainian). Available: https://www.ifac-control.org/about/history
5. O. Gavva, *Integrated automated control systems: history and development potential*, (in Ukrainian). Available: http://ogas.kiev.ua/perspective/integrovani-avtomatyzovani-systemy-upravlinnya-istoriya-ta-perspektyvy-rozvytku-504
6. *Department of Information Systems and Technologies*, (in Ukrainian). Available: https://ist.kpi.ua/uk/history/
7. *Physical and Technical Institute of Igor Sikorsky KPI*, (in Ukrainian). Available: https://kpi.ua/ipt
8. *Institute of Applied Systems Analysis of the Ministry of Education and Science of Ukraine and the National Academy of Sciences of Ukraine at Igor Sikorsky KPI*, (in Ukrainian). Available: https://kpi.ua/institute/ipsa
9. *Systems Research and Information Technologies (academic journal)*, in Ukrainian. Available: http://iasa.kpi.ua/iasa-journal
10. *UNESCO International Chair "Higher Engineering Education, Applied Systems Analysis and Informatics*", (in Ukrainian). Available: https://kpi.ua/web_unesco
11. *Physics and mathematics faculty of KPI named after Igor Sikorsky*. Available: https://fmf.kpi.ua/

12. *Institute of Telecommunication Systems of NTUU "Ihor Sikorsky KPI"*. Available: https://kpi.ua/en/faculty/its.html/
13. "Ukrainian "white" hackers beat 12,000 teams in the world ranking," (in Ukrainian), *Ukrinform*. Available: https://www.ukrinform.ua/rubric-technology/2164581-ukrainski-bili-hakeri-obijsli-12-tisac-komand-u-svitovomu-rejtingu.html/

Received 01.09.2023

INFORMATION ON THE ARTICLE

Michael Z. Zgurovsky, ORCID: 0000-0001-5896-7466, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: zgurovsm@hotmail.com

СПАДЩИНА АКАДЕМІКА ГЛУШКОВА: ЛЮДСЬКИЙ КАПІТАЛ У СФЕРІ КІБЕРНЕТИКИ, ОБЧИСЛЮВАЛЬНОЇ ТЕХНІКИ ТА ІНФОРМАТИКИ В КИЇВСЬКОМУ ПОЛІТЕХНІЧНОМУ ІНСТИТУТІ ІМ. ІГОРЯ СІКОРСЬКОГО / М.З. Згуровський

Анотація. Розглянуто роль академіка Віктора Глушкова у створенні наукових шкіл в галузі кібернетики, обчислювальної техніки та інформатики в Київському політехнічному інституті ім. Ігоря Сікорського, який став потужним національним центром підготовки фахівців у цій сфері. Показано значний вплив ідей академіка Глушкова на формування поколінь учених, які дотепер продовжують розбудовувати цифрове суспільство в Україні та далеко за її межами.

Ключові слова: академік Віктор Глушков, кібернетика, обчислювальна техніка, інформатика, цифрове суспільство.

UDC 62-50
DOI: 10.20535/SRIT.2308-8893.2023.3.02

AUGMENTED SECURITY SCHEME FOR SHARED DYNAMIC DATA WITH EFFICIENT LIGHTWEIGHT ELLIPTIC CURVE CRYPTOGRAPHY

DIPA D. DHARMADHIKARI, SHARVARI C. TAMANE

Abstract. Technology for Cloud Computing (CC) has advanced, so Cloud Computing creates a variety of cloud services. Users may receive storage space from the provider as Cloud storage services are quite practical; many users and businesses save their data in cloud storage. Data confidentiality becomes a larger risk for service providers when more information is outsourced to Cloud storage. Hence in this work, a Ciphertext and Elliptic Curve Cryptography (ECC) with Identity-based encryption (CP-IBE) approaches are used in the cloud environment to ensure data security for a healthcare environment. The revocation problem becomes complicated since characteristics are used to create cipher texts and secret keys; therefore, a User revocation algorithm is introduced for which a secret token key is uniquely produced for each level ensuring security. The initial operation, including signature, public audits, and dynamic data, are sensible to Sybil attacks; hence, to overcome that, a Sybil Attack Check Algorithm is introduced, effectively securing the system. Moreover**,** the conditions for public auditing using shared data and providing typical strategies, including the analytical function, security, and performance conditions, are analyzed in terms of accuracy, sensitivity, and similarity.

Keywords: ciphertext, user revocation, data sharing, CC, ECC, security issues.

INTRODUCTION

Innovative developments have made it possible to implement progressive solutions to improve the nature of human existence. In order to gather information and address challenges relating to wellbeing, analysts who are thinking about the growth of innovation have collected and evaluated wellbeing data from these sources. Accordingly, the development of integrated medical care innovation has the potential to improve efficiency and results comprehension at every level of the medical care framework [1]. By utilizing robust patient well-being controls, pervasive information access, remote patient checking, quick clinical intervention, and decentralized electronic-medical care records, new electronic health (e-Health) application frameworks are being developed that can address specific issues related to traditional medical care frameworks [2]. These systems can manage patient and well-being data, boost individual satisfaction, foster teamwork, enhance outcomes, cut costs, and generally improve the effectiveness of e-medical care administrations [3].

IoT usage and the advancement of wireless communication technologies enable real-time streaming of patients' health conditions to caregivers [4]. Additionally, a number of readily available sensors and portable devices may measure particular human physiological parameters with a single touch, including blood pressure (BP), respiration rate (RR), and heart rate (HR) [5]. Although it is still in the early stages of development, businesses and industries have quickly incorporated the power of IoT into their current systems and seen gains in both user experiences and production [6]. However, the use of IoT technology in the healthcare sector poses a number of difficulties, including those related to data management, storage, and transmission between devices as well as issues with security and privacy. Among the potential answers to these is Cloud Computing technology [7].

Cloud Computing is incredibly beneficial since it has many distinctive characteristics. Cloud computing makes a variety of services easily available [8]. Some advantages of this technology include suppleness, cost consumption, pay-as-you-go, increased effectiveness, and nimbleness. Through dedicated CS, cloud computing services offer client-specific applications and data storage [9]. Through the use of the cloud, businesses are able to avoid upfront framework costs and investments, allowing them to launch their applications more quickly, more intelligently, and with less maintenance. To provide the best services to mobile clients, portable distributed computing combines cloud computing and mobile devices [10].

Cloud Computing becoming increasingly important for the rapid growth of technology, particularly in the industry of the health care system [11]. Cloud storage is many times less costly than server storage, storage, equipment materials, and HR training to strengthen required operations [12]. Because in the cloud all the information of patients is stored, and retrieve the prescription from the cloud at any time and from any location. Because the data is stored in the cloud, healthcare professionals and identified patients will be able to access it via mobile devices such as smartphones and PDAs [13].

Cloud Computing enables internet-connected devices to access healthcare information from anywhere in the world. Furthermore, medical practitioners may exchange their resources and medical knowledge with other famous researchers in the same area from across the world [14]. Healthcare organizations, medical medication producers, pharmacists, medical insurance providers, researchers, and patients must all exchange EH records [15]. This presents a significant challenge in terms of keeping sensitive patient data secure. While there are numerous benefits, there are also some risks, particularly with regard to data security in the cloud, which is the most difficult issue at all times. It becomes more difficult in cloud computing because the actual data is stored in another location. As a result, providing security for data in the cloud is a time-consuming task for cloud computing organizations [16].The current healthcare sector is beset with computational and processing challenges. Inherent issues in the traditional healthcare sector [17; 18]. Patients' records include sensitive information that must be kept secure at all times. The current method has several irregularities in securing patient data. Medical data takes up more memory space, which is inefficient [19]. As a consequence, the purpose of this study is to enhance the performance and strengthen the operations of the existing cloud system in the healthcare industry. The proposed solution's key contributions include security and privacy, low-cost access policies for SHRs (Smart Health Records), a lightweight IoT detection system, rapid detection of Sybil attacks to reduce their impact on the network, and Sybil node identification using node properties. The proposed system in the given paper includes:

- A reversible and efficient no-pairing data-sharing technique for cloud storage systems based on Elliptic Curve Cryptography (ECC) with Identity-based encryption (IBE) approach.
- Cipher text Based Encryption (CBE), which presents an encryption access control (EAC) technique to satisfy User revocation that includes both user revocation and attribute, is seen here as a means of ensuring data security.
- A threat detection model to identify resource depletion attacks and Sybil attacks detection in the cloud system ensuring the safe storage of data.

LITERATURE SURVEY

In their study, Yan et al. [20] created the retrieval and storage-based indexing framework (RSIF) to enhance concurrent user and service provider access to healthcare data stored in the cloud. Concurrent access to stored data was made possible through continuous, replication-free indexing and time-constrained retrieval. Deep learning is used for all storage instances to categorize the restrictions for data augmentation and update. The learning process determines the approximate indexing and ordering for storage and retrieval, respectively, through conditional assessment. As long as the processes are independent, this helps to shorten the time for access and retrieval occurring at the same time. This data analysis should be carried out using various indexing techniques and storage and processing techniques.

PRCL, a privacy-aware, and resource-saving collaborative learning protocol was proposed by Hao et al. [21]. They created a unique model splitting technique that divides the network into three pieces, which offloads the intensive center half for CS to decrease the overhead. The original data, labels, and model parameters are all kept private by PRCL through the use of a mild perturbation of data and filled partly homomorphic encryption. Additionally, they examined the suggested protocol's security and showed how PRCL performed better with correctness and effectiveness. Future work will need to focus on developing adaptive data perturbation techniques, lowering the system's overhead of communication.

For a system of exchanging personal health records, Zhang et al. [22] suggested an effective identity-based distributed decryption technique. They may easily share their data with different parties without having to reassemble the decryption private key. They demonstrated the scheme's resistance to chosen-ciphertext attacks (CCA). Additionally, they used an Android phone and a laptop to implement the plan using the Java pairing-based cryptography (JPBC) package. The outcomes of the trial demonstrated the system's viability in an electronic personal health record system. In the future, it will be necessary to look into some more effective strategies, such as removing the zero-knowledge proof from the scheme and dispersing the secret without the need for a secret channel.

Using JavaScript-based smart contracts, Singh et al. [23] proposed a patient-centric architecture for a decentralized healthcare management system with a blockchain-based EHR. Additionally, a functional prototype based on the composer and Hyper ledger fabric technology has been put into place, ensuring the security of the suggested paradigm. Performance metrics including latency, throughput, resource usage, and others are measured in experiments using the hyper ledger calliper benchmarking tool under various situations and control parameters. The outcomes support the effectiveness of the suggested strategy. The authors want to expand their work based on fault tolerance in the future.

For cloud-based WBANs, Yang et al. [24] suggested a brand-new effective and anonymous authentication technique. The security study demonstrated that the system could fix the flaws in earlier ones and satisfy all security criteria. They also demonstrated the benefits of the suggested plan via presentation analysis of functionalities, computing transparency, storage overhead, and communication overhead demonstrating that the plan is better suited for real-world applications in the healthcare industry. The authors intended to create a universal authentication system that may be used in a variety of application scenarios in the future.

Son et al. [25] use blockchain to ensure data integrity and cipher text-User attribute-based encryption (CP-ABE) to create access controls to store data on CS. They used automated validation of internet security protocols and applications to do informal analysis, Burrows-Adabi-Needham (BAN) logic analysis, and formal validation of the proposed protocol to verify its robustness (AVISPA). As a consequence, they proved that the proposed protocol is more secure and performs better than comparable protocols. Future work will include modelling the full network as well as the security protocol in order to develop a new, more practical solution.

From the survey it is observed that for [20] data analysis should be carried out using various indexing techniques and storage and processing techniques, [21] need to focus on developing adaptive data perturbation techniques and lowering the system's communication overhead, for [22] it is necessary to look into some more effective strategies, such as removing the zero-knowledge proof from the scheme and dispersing the secret without the need of a secret channel, [23] work has to be expanded based on fault tolerance, for [24] the authors intended to create a universal authentication system that may be used in a variety of application scenarios and [25] requires simulation of the entire network and the secure protocol in order to build a new and more workable. Hence order to achieve the abovementioned necessities it is essential to develop a model.

AUGMENTED SECURITY SCHEME FOR SHARED DYNAMIC DATA WITH EFFICIENT LIGHTWEIGHT ELLIPTIC CURVE CRYPTOGRAPHY

This paper takes into account a system for auditing cloud storage that consists of the cloud, users, a group, and proxies. The company can host and share data to the cloud. Users who produce data and exchange it with one another make up the group. Users in the group are able to govern the group collaboratively since they trust one another. Here initially a container is created to store Cloudlets in the CloudSim library. The Data centers and Brokers are created in which the VMs and Cloudlets are loaded. Hosts with specific IDs are created and the simulation is run (Fig. 1).

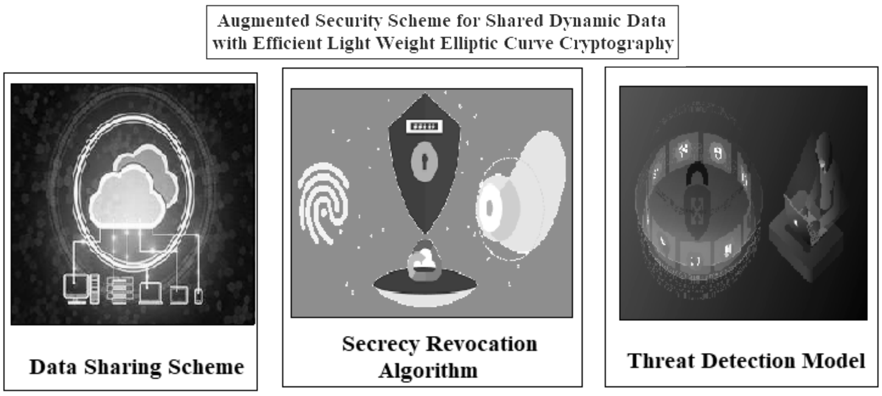


Fig. 1. Framework of the proposed model

Data sharing scheme

Traditional encryption techniques consume a lot of storage space since they require duplicates for all cipher text in every particular user through a unique key. To create a secure, reliable, and accurate model sharing data, the dispute must consider, importantly the key distribution among new users necessitates that data owners remain online constantly. Initially, the system must allow data owners to add or delete users. Then, the system must allow data owners to guarantee that data confidentiality is protected from CS, attribute authorities, and unauthorized users. Consequently, consumers should confirm the data correctness they have received. At last, users have to have mobile access to shared data. Hence, considering the above-mentioned requirements this research presented a reversible and efficient no-pairing data-sharing technique for cloud storage systems based on ECC withIdentity-based encryption (CP-IBE) approach. Cloud computing, smart grids, the Internet of Things, and other distributed systems may all benefit from the one-to-many encryption technique known as CP-IBE, which is based on public keys and enables flexible and fine-grained data access management. Unfortunately, because they rely on pricey bilinear pairing algorithms and have large encryption and decryption computation overhead costs, the majority of modern data-sharing approaches are ineffective for cloud systems with limited resources. The paper here proposed an effective data-sharing scheme for cloud storage systems based on the fact that the ECC algorithm has stronger bit security than exponential-based public key cryptographic algorithms like RSA and can achieve the same level of security with smaller key sizes and higher computational efficiency. The four entities in the proposed system model are a trusted expert, a cloud source, senders, and users which are depicted in Fig. 2.

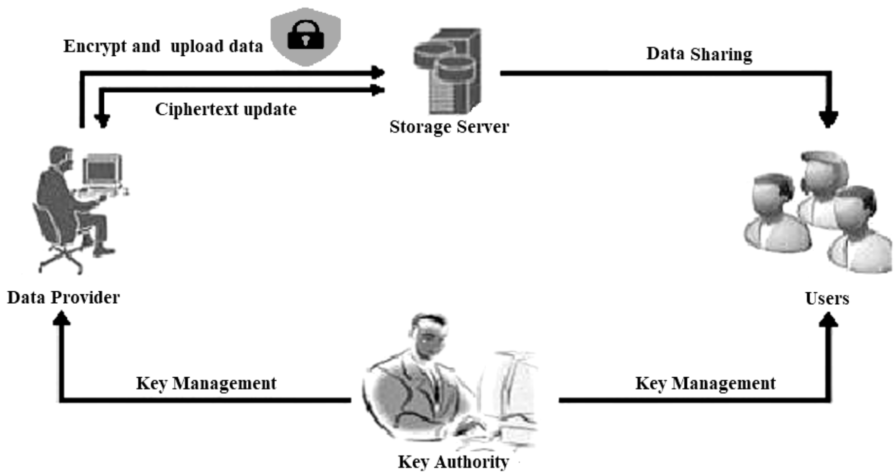


Fig. 2. Entities of the system model [31]

Trusted Expert (TE). The TE is generate both master secret keys and global public parameters, which produce and share out users' matching private keys. Additionally, it is in charge of blocking users. The TE is expected to be trustworthy but inquisitive, which means refuse service for authorized users provide adhere to established procedures appropriately, if interested in the data's substance, and want to learn as much as it can about its customers' private information.

Cloud Source (CS). Data from the owners are gathered and stored by the CS, a potent computing entity with limitless resources. Additionally, computing a sizable amount of decryption overhead aids users in decrypting the cipher text. Like the TE, the CS is expected to be sincere but curious.

Data owner. A data owner is a group that needs to outsource a data file to the CS. Data is encrypted initially under a set of characteristics before transmitting it.

User. It is an entity with unrestricted access to the ciphertext of the cloud server. To achieve this, it first creates a token using its key, and then it asks the CS for access to the data by giving the CS the token.

Proposed CP-IBE-based ECC approach. The algorithms utilized in the data-sharing system are discussed in detail below.

- **Setup** $(S_p, A) \rightarrow (M_K, P_G)$

The setup procedure takes a security parameter S_p and an attribute universe A as inputs and produces global public parameters P_G and a master key M_K for the system as outputs.

- **Encryption** $(P_G, M, \omega) \rightarrow CT$

The encryption method is given a message M, a collection of descriptive qualities, and the global parameters P_G. It generates cipher text CT.

- **KeyGen** $(P_G, M_K, T) \rightarrow SK$

The global parameters P_G, the master secret key M_K, and an access tree T are sent into the key creation method. For each authorized user, it produces a private key P_K.

- **TokenGen** $(P_G, D) \rightarrow TK$

The user executes this procedure to create a decryption token TK.

- **Partial Encryption** $(P_G, TK, CT) \rightarrow CTPartial$

The partial decryption algorithm accepts the global parameters P_G, the user token, and the ciphertext as input. It returns ciphertext that has been partly decrypted.

- **Decryption** $(CTPartial) \rightarrow (M, AC_M)$

The user executes the decryption algorithm, which uses the partly decrypted ciphertext. It displays the message M as well as the message authentication code, AC_M.

- **Elliptic Curve Cryptography**

ECC is public key cryptography. Assume p is a prime number and F_p is the field of integers modulo p. A cubic equation $y^2 = x^3 + ax + b$ defines an elliptic curve (EC) over a finite field (Galois Field) GF and each elliptic curve is formed by a distinct value of a and b. The collection of all locations (x, y) that satisfy the above equation, as well as a point in infinity, lies on the elliptic curve. The private key in the ECC is a random number, while the public key is a point in the curve formed by multiplying the private key by the generator point G in the curve.

Based on the preliminary results and system model, this paper presented a CP-IBE scheme in detail in this section owner O develops a collection of expressive qualities before encrypting a message M with an algorithm and sending cipher text to the CS provider. This work employs lightweight operations of ECC in conjunction with an algorithm of symmetric encryption. If user U_s gets provided authority and wishes to get data stored on CS, user U_s must construct and deliver a decryption token to the CS. When the CS receives the token, it partly gets back the saved cipher text, and transmits consequential communication to the U_s. Ciphertext can then be readily decrypted by the user. Setup, Encryption, KeyGen, TokenGen, Partial Decryption, and Decryption are the six function modules of the proposed data-sharing method. These are their descriptions:

- **Setup** $(S_p, A_u) \to (M_K.P_G)$

The trusted attribute authority executes the setup procedure, input as security parameter S_p and the attribute universe A_u. It generates Global Public Parameters P_G as well as the Master Key M_K. To do this, the authority first chooses a random integer R_i from ZP^* and computes $PK = R_i.G$. Let $A_u = \{1,...,n\}$ be a collection of all attributes in the system; the authority selects a random number $s_i \in ZP^*$ and computes the public key of each attribute I as $P_i = s_i.G$. The CS chooses a random number μ for its secret key and calculates $PKCS = \mu.G$ for its public key. The authority does not know μ, and the CS proves the authority's knowledge of μ using a zero-knowledge proof procedure. The authority assigns the secret key $M_K = \{R_i, \{s_1,...,s_i\}\ i \in A_u$ and publishes the global parameters $P_G = \{M_K, M_K CS,\ \{P_1,...,P_i\}\ i \in A_u\}.o$.

- **Encryption** $(P_G, M, \omega) \to CT$

As input, the encryption method receives a message M, a collection of descriptive qualities ω, and the global parameters P_G. When the owner O wishes to encrypt a message using the set of characteristics, he or she selects a value at random from the set of values Z_P^* and computes K and F as follows:

$$K = d \cdot PK = (k_1, k_2)\ ;\quad F = d \cdot PKCS = (f_1, f_2)\,.$$

If $K = O$, the authority re-selects d at random from Z_P^* to compute K until $K = O$. Then use the points (k_1, k_2) as the encryption and integrity keys, and construct ciphertext C and AC_M for message M as follows:

$$C = ENC(M, k_1)\,;$$

$$C' = ENC(C, f_1)\,; \tag{1}$$

$$AC_M = HAC_M(M, k_2)\,. \tag{2}$$

$ENC\,()$ in Equation (1) is a symmetric encryption method such as AES. The message M is encrypted with key k_1 and then re-encrypted with key f_1, obscuring the ciphertext C from the authority's view. $HAC_M()$ is a cryptographic hash function in Equation (2) that creates the hash-based message authentication code for message M based on the integrity key k_2. Finally, the owner O computes $C_i = d.P_i$ for all of the attributes in ω and uploads $CT = (\omega, C', AC_M, \{C_i\}\ i \in \omega, H = d \cdot G)$ to the cloud service provider.

- **KeyGen** $(P_G, M_K, T) \rightarrow SK$

The keyGen algorithm calculates the decryption key for a user's request, U_B, in the manner shown below. In the access tree, it selects a polynomial q_x for each node x. Starting with the root node *R*, these polynomials are selected from top to bottom. The authority determines the degree *dx* of the polynomial q_x for each node x in the tree to be one less than the threshold *kx* of that node, that is, $dx = kx_1$. In order to fully fix qR, it first sets $qR(0) = R_i$ for the root node *R* (keep in mind that R_i is the authority's secret value). The remaining points are then set at random. It sets $q_x(0) = q_{\text{parent}(x)}(index(x))$ for every other node x and picks dx additional random locations like $qR.(x)$. The unique index number assigned to x by its parent is called $Index(x)$.

Let Y represent the collection of leaf nodes in the tree, and $att(y)$ represents the attribute related to leaf node y. The KeyGen method produces the following values once the polynomials are finished for all leaf nodes *y*:

$$D_y = q_y(0) / s_i,\ i = att(y).$$

Finally, the authority sends $D = (D_x,\ i = att(x),\ \text{and}\ i \in \omega)$ as the private key for U_B.

- **TokenGen** $(P_G, D) \rightarrow TK$

At this phase, a user U_B generates a token TK_B based on his/her private key and sends it to the *CS* to convey most of the decoding computational load to the *CS*. For this purpose, U_B first selects a random number b from Z_P^* and computes $D' = \{D_x \cdot\ b = (q_x(0) \cdot\ b) / s_i\} i \in \omega$. After that, the user U_B computes the point $Q = b.PK_{CS} = b.\beta.G = (q_1, q_2)$ and $B = b.G$ and sets the token TK_B as follows:

$$TK_B = \{B, TB = ENCq_1(D')\ i = att(x),\ i \in \omega, T)\}\ .$$

To protect against a DOS attack in this case, a timestamp T is employed. The *CS* examines the time stamp T and decrypts TB after receiving the token. The *CS* continues the partial decryption step if it is valid. An symmetric encryption function like $AES\ is\ ENC(.)$.

- **Partial Decryption** $(P_G, TK_B, CT) \rightarrow CTPartial$

When the CS receives the token TK_B, it uses its secret key to compute the decryption key q_1 as $Q = \beta.B = \beta.b.\ G = (q_1, q_2)$ and then decrypts T_B. The *CS* declines the request for partial decryption if the timestamp $T's$ validation is unsuccessful. Otherwise, the partial decryption process is carried out by the *CS* as follows. Let x be a node of tree T, first this algorithm defines a recursive algorithm Decrypt Node (CT, D', x). Let $i = att(x)$, if x is a leaf node, then Decrypt Node (CT, D', x) is computed as follows:

$$D'x.Ci = Dx.b.Ci = q_x(0).s-1i.b.d.s_i.G = q_x(0).b.d.G.$$

This recursive procedure returns an element from the ECC group or $\perp$.

The algorithm $Decrypt\,Node(CT, D', x)$ searches for all nodes z that are offspring of x if x is a non-leaf node, and the result is saved as F_z. Let L_x be a random k_x-sized collection of child nodes z that satisfy $Fz \neq \perp$. $Decrypt\,Node\,(CT, D', x)$ returns if there is no such L_x, indicating that the node was not fulfilled. Otherwise, assuming $i = index(z)$ and $L'x = \{index(z),\ z \in Lx\}$ it is possible to calculate $DecryptNode(CT, D', x)$ as follows:

$$\Sigma Xz \in Lx\Delta i, L'_x(0).DecryptNode(CT, D', z) = \sum_{z \in L'_x} \Delta i, L'_x(0).qz(0).b.d.G =$$

$$= \sum_{z \in L'_x} \Delta i, L'_x(0).qparent(z)\big(index\ (z)\big).b.d.G =$$

$$= \sum_{z \in L'_x} \Delta i, L'_x\ (0).q_z(i).b.d.G = q_z(0).b.d.G. \qquad (3)$$

The outcome of $Decrypt\,Node\,(CT, D', R) = qR$ for the root node R of the access tree T is based on the information given $qR(0).R_i.b.d.G = R_i.b.d.G$ The CS calculates F = after computing the $Decrypt\,Node\,(CT,\ TK_B, R).C = DEC(C', f_1)$ and $H = .d.G = (f_1)(f_1, f_2)$. The CS then gives the user $UB\,CT_{Partial} = C,\ AC_M,\ N = Decrypt\,Node(CT, D', R)$. Equation (3) demonstrates that the cloud service provider cannot decipher the ciphertext since they are unaware of the value of b and can only assist the receiving users in doing so.

- **Decryption** $(CT_{Partial}) \rightarrow (M, AC_M)$

The user UB may quickly determine the decryption and integrity keys after receiving the $CT_{Partial}.N = R_i.b.d.G$ means that the decryption method just needs to divide b to retrieve the keys as Equation

$$C'_1 = N / b = \alpha.b.d.G.b^{-1} = \alpha.d.G = (k'_1\ , k'_2)\ .$$

The points $(k'_1\ , k'_2)$ serve as the message $M's$ integrity key and decryption key, respectively. When the user enters $M' = DEC(C, k'_1)$, the message M may be decrypted. If $HAC_M\,(M', k'_2) = AC_M$, the message M is accurate.

In order to revoke a user *UB*, the authority securely transfers all of the revoked user *UB*'s attributes to the *CS*. When the *CS* receives the token, it first determines if it has all of the *UB*'s properties, and if so, it rejects it.

Secrecy Revocation Algorithm

The system has the power to cancel a user's access if they leave the system. In other words, the authority securely transmits to the *CSP* all of a user's *UB*'s revoked characteristics. When the *CSP* receives the token, it first determines if it has all of the *UB*'s properties, and if so, it rejects it. Even if the user still possesses a working secret key after the revocation procedure, they will not be able to access the saved data. The revocation procedure is effective since no need to be updated. Cipher text Based Encryption (CBE), which presents an encryption access control (EAC) technique to satisfy *User* revocation that includes both user revocation and attribute, is seen here as a means of ensuring data security. The authorized users should be updated right once if the data owner modifies one of the attributes in an access *User* since the revoked users who had previously received access to the *User* can see the ciphertext. Four different sorts of update *User* levels are established, specifically for data owners. All secret token keys are uniquely produced at all levels by categorizing those levels. As a result, the secret token key is hashed to create a new secret key (Fig. 3).

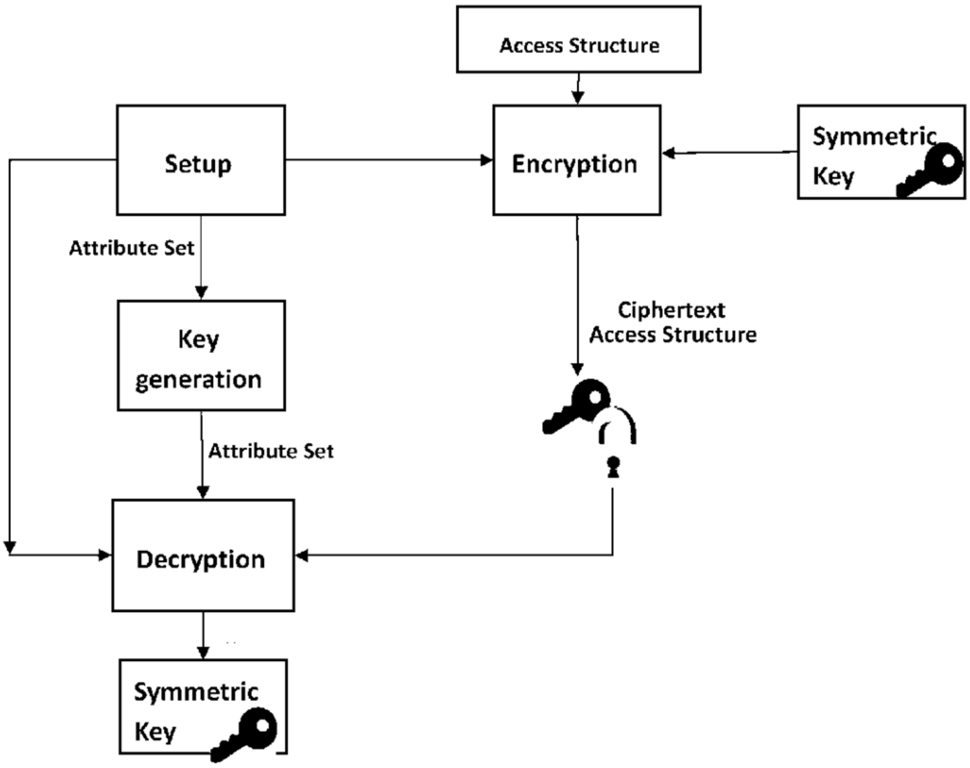


Fig. 3. Flow chart encrypting *User* algorithm

The proposed method has four basic algorithms to handle user secrecy revocation. They are, in order, the encrypting *User* algorithm, re-encryption algorithm, update key generation algorithm and decryption algorithm.

Encrypting *User* algorithm. Each access *User* is represented by a distinct identity, which stands for the traits that are crucial for identifying authorized individuals. The data owner's *(DO)* present *User* identity is documented in the algorithm for updating policies as $User_{id}$, and the new *User* identity the *DO* wants to modify is marked as $New\ User_{id}$. The four status numbers are specified for each level of the individual updating *User*. The status number should be. The identification of the material for plaintext is recorded in the system and is specified as α in accordance with each Upd_{level}.

Input: $User_{id}$, New_User_{id}.

Output: Upd_{level}

1) perform one of four *User* levels for updating;

2) recognize Upd_{level} based on choice in step 1;

3) obtain α by formative all Upd_{level};

4) locate New_User_{id} as a present *User* identity and cancel the old $User_{id}$;

5) go again Upd_{level}.

Re-Encryption Algorithm. The plaintext message (M) related to new_User_{id} is was re-encrypted using a special method. The Upd_{SK} is the encryption key used by the DO. The generated cipher text is C.

Input: M New_User_{id}, Upd_{SK}.

Output: C

1) perform updates $User_{id}$ to, New_User_{id} for M;

2) obtain $Upd_{level} = UpdateUser(Userid, New\_User_{id})$;

3) acquire Upd_{SK} from the TA;

4) $C = ReEncrypt\ (M, New\_User_{id}\ Upd_{SK})$;

5) go back C.

Update Key Generation Algorithm. *TA* generates a default key string () and uses *M'*s unique identity, which is to be accessible as, in the update key creation procedure. The status number (α) is originally determined by *TA* using the Upd_{level}. The TA then changes the numbers for both strings. The *TA* then joins the strings of the key with the other keys. As a secret token key, this concatenation is transformed into the *Base64String format* (SK_{token}). By obtaining the hash code from the MD5 function for the SK_{token}, the *TA* then constructs Upd_{SK}. The *Base64String format* output is likewise consistently applied to the final Upd_{SK}. According to each user's request identification, the *TA* also creates $a\$$ for them.

Input: $¥, \beta, \gamma, Upd_{level}, New\_User_{id}$.

Output: Upd_{SK}, \$

1) start α based on Upd_{level};

2) TA obtains the and β since the users need a directory;

3) TA gets ¥, and converts both β and α to strings;

4) $SK_{token} = Concatenate(¥, \beta, \alpha)$;

5) $SK_{token} = ToBase64String(SK_{token})$;

6) . $Upd_{SK} = GetHashCode(SK_{token})$;

7) $Upd_{SK} = ToBase64String(Upd_{SK})$;

8) TA generates \$;

9) records the $New_{User_{id}}$ to filter unauthorized users;

10) return Upd_{SK}, \$.

Decryption Algorithm. The data user (*DU*) uses his $User_{id}$ to prove his identification as a *User holder* in the decryption procedure to decode *C*. The *TA* will offer the *DU* the Upd_{SK} and if his $User_{id}$ is acceptable to the New $User_{id}$ of the *DO*. In order to obtain the original ciphertext C, the *DU* uses it to decode the proxy ciphertext (*C'*) from the cloud. The *DU* can then use this Upd_{SK} to decode the *C*.

Input: *C*, $, Upd_{SK}, $User\_User_{id}$.

Output: *M*

1) *DU* enters the system by demonstrating his abilities;

2) the system records his $User\_User_{id}$ according to his attributes;

3) *DU* needs the. Upd_{SK} and $ to the TA;

4) the system records his $User\_User_{id}$ based on this attributes;

5) *DU* requests the Upd_{SK} and $ to the *TA*;

6) *DU* verifies inbox and acquires the Upd_{SK}, $ from *TA*;

7) *DU* obtains the C' from the cloud;

8) $C = Decrypt\ (C', \$, User\_User_{id})$;

9) $M = Decrypt\ (C, Upd_{SK}, User\_User_{id})$;

10) return M.

Threat detection model

Resource depletion attack detection algorithm. In order to distinguish between legitimate and illicit sources, first specify a set of parameters (or rules) for a source $S_1 (y \le j \le x)$. The first argument is *atj* $(atj \le h)$, which stands for the active time of a source S_1. This field demonstrates that source S_1 has not delivered any attack traffic to the system while acting as a cloud provider, in addition to indicating how long a source has been connected to a cloud service. As a result, *atj* is regarded as a crucial factor in making informed choices about attack detection. The second parameter, N_{in}, indicates how many incoming flows have been established to a cloud service and is the number of flows from a source S_1, where $N_{in} \ge 1$. Whenever return traffic constantly takes a different route for a better service response or for other objectives, it is fair to just take into account request flows. In the Internet protocol suite, the *ICMP* protocol is referred to as a supporting protocol and is only used to determine whether a requested service is unavailable or whether a host or router could not be reached. As a result, a typical source S_1 typically transmits a small number of request packets to a destination, and each source S_1 only creates one flow in an *SDN* switch's flow tables. A malicious source S_1, on the other hand, can produce one or many flows in an *SDN* switch. In order to distinguish an abnormal source S_1 from its regular counterparts based on the aforementioned studies, define the third parameter, Pf (the average number of packets per flow of the source S_1), as follows:

$$Pf = tpktj;\quad Attack\,Type\,I\ ;$$

$$Pf = \frac{tpktj}{nconj};\quad Attack\,Type\,II\ .$$

Where *tpktj* represents the source $S_1's$ transmitted packets to the cloud and *Nin* represents the source $S_1's$ flow number. The source $S_1's$ traffic protocol is a crucial factor in dividing the two primary assault methods. The following assessed parameter is the priority of a source S_1, indicated as *Prij*, which makes a distinction between reliable $(Prij = 1)$, typical $(Prij = 2)$, and unidentified $(Prij = 3)$ sources. The $flag, Flag_j$, which displays the status of the source S_1, is the final parameter introduced. Source S_1, has two statuses: attack and normal, which correspond to $Flag_j = -1$ and 1, respectively. At first, a new source is designated as a standard source. It should be noted that the Update Agent updates a database after collecting values from each active source $S_1's\, atj, Nin, Pf, Proj$, and $Prij$ at each observation. The tuple of parameters for a source S_1 is constructed using the definitions given above and is an entry in the database. $S_1 = (atj, Nin, Pf, Proj, PrijFlag_j)$. The search engine uses S_1 and $Proj$ matches to separate the *IP* sources for each search operation. Then examine the effective classification of normal and attack sources using the *IP-based* technique utilizing a history of *IP* database (*HIP Database*). Based on the statistics gathered at each observation, first extract and update all active source *IP* addresses by protocol to the *IP* Database using the Update Agent when the system is functioning. Utilize the *Ini* sets to categorize normal and abnormal sources for each sort of traffic protocol for our first observations and keep the *Ini* sets separate until the proposed system picks up any attack sources. The *IP* Database could provide some more sources for the observation t. Using pre-made *Ini* sets, these sources are further validated to determine their *Flag*. The suggested Algorithm will replace the value of the associated *Ini* set for the subsequent observation $(t+1)$. Then, using the $B(tC_1)ICMP$ comparison, identify these sources as either regular sources or attack sources using the Anomalous Source Detection Algorithm.

Step 1: $B(t+1)i = \{(ati; w1i); (Pfi; w2i); (nconi; w3i)\} \leftarrow e$ The boundary set of the protocol I have given at the t^{th} observation

Step 2: $(atj, Nin, Pf, Proj, PrijFlag_j)$ A set of attributes of a source $S1$ that is collected at the (tC_1) observation

Step 3: $Xi\, D\, ati * w1i + C\, Pfi * w2i + C\, nconi * w3i$

Step 4: $Xj\, D\, atj * w1i + C\, Pf * w2i + C\, Nin * w3i$

Step 5: if $Xi \le Xj$ **then**

Step 6: $Flag_j = -1$ {Attack source}

Step 7: else

Step 8: $Flag_j = 1$ {Normal source}

Step 9: end if

Step 10: return: $Flag_j$

Normally attacks create a large number of flows to the target with a small number of packets. In order to counteract this attack, the Mitigation Agent notifies the *SDN* controller's forwarding engine to ignore packets in messages from attacking sources that demand the installation of new flows at the Open Flow switch and sends a flow mod message with a delete action to the edge Open Flow switch. In the event of an assault, this strategy eliminates all anomalous flows and stops fresh attack flows. As a result, it prevents the system from being overload-edand guarantees efficient cloud operation. Additionally, the attack flow deletion from the switch will result in a considerable drop in the number of collected flows at the time of the subsequent observation. Therefore, this approach can conserve as many computing resources as feasible from the cloud control plane.

Sybil attacks detection. Sybil Attack Check employs encryption in order to offer secure communication, and it also safeguards the data against Sybil assaults by utilizing a detection algorithm. The cluster's nodes are given secure communication during the secure phase. Utilizing one-time authentication (OTA), the secret key and the node id provide authentication. A node needs a one-time authentication from the Cluster head (CH) in order to communicate, and the CH node only sends the OTA following a successful verification. The CEA cryptography algorithm is a novel one that is suggested to offer secrecy (Cipher encryption algorithm). Traditional encryption techniques are resource-intensive; they use a lot of storage space, power, and processing time.

CEA is a straightforward encryption algorithm with promising security. It just requires a small number of rounds and basic operations. The precise stages of the CEA algorithm are shown in the algorithm below. It accepts 64-bit plain text as input, and 32-bit random keys, and outputs a 64-bit encrypted text. Either eight or sixteen rounds of operation can be used to implement CEA. Four fundamental operations—*XOR, pair swap, encoding, and 1s complement*—are carried out in each round. In the encryption procedure, the plaintext is divided into the left plaintext (L_i) denotes the round, and the right plaintext (R_i), and an *XOR* operation is performed with two 32-bit random keys. On the resulting L_i and R_i, a paired exchange is performed in the second phase. Two-bit pair swapping is done after the plain text has been divided into two-bit pairs. These two-bit pairs are then substituted in the following phase with the equivalent encoding bits. As an illustration, the bit 00 is encoded as 01, 01 as 10, and so on. The final step involves performing a 1s complement operation to obtain the necessary 64-bit cipher text. The operations are carried out backward throughout the decryption process; the first stage involves doing the 1s complement. The next phase involves decoding using the decoding bits followed by a pair swap operation. To obtain the necessary plain text, an *XOR* operation is performed in the last stage.

The cryptographic symbols used in Sybil Attack Check include: G_g — Group generator of prime order; G_g — a bilinear group output; A — set of all features; SK_s — secret key; CA — cipher text with access User. Set of integer integers A and Z_p. While security typically rises, composite-order group pairing performance sharply declines. Since the assault, in this case, is selective, there is a need to utilize a prime-order group since it is less difficult and more affordable. Prime-order groups can only provide selective security in their respective security models. Definition: Assume that G_g operates on the input variable as a bilinear prime order group generator. $G_g 1$, and $G_g 2$, which have additive and multiplication properties, are cyclic groups of prime order.

The output of G_g is defined as $(p, G_g 1, G_g 2, G_g T, e)$.

Key generation $(PK,\ M_K, S) \rightarrow SK_s$: The secret key SK_s are produced using the key generation algorithm using the inputs of the public key PK, master key M_K, and set of characteristics S, as illustrated in Fig. 4.

$SK_s = (S, K, K', \{K_i\}\ i \in Is)$, where $S = (Is, S\ )$ with $Is \subseteq Zp$ and $S = \{S_i\}$ $i \in Is$, $K = g^{\alpha} g^{at} R$ choose $tR\,Zp, R, R', Ri \in R\,G2, K' = gt\,R'$ and $Ki = (Ui\,si\,h\ )t\,Ri.$

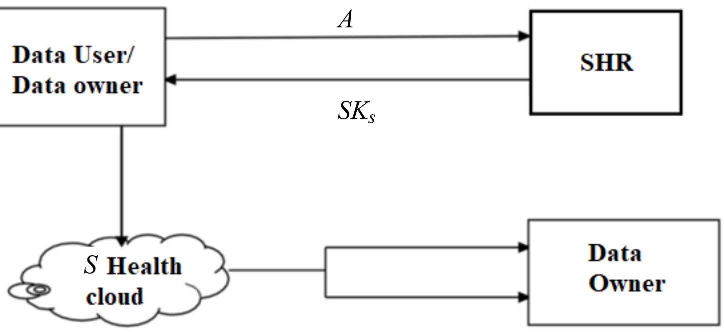


Fig. 4. Smart Health Records (SHR) encryption and decryption process in the cloud

Where $e : G_g 1 \times G_g 2 \rightarrow G_g T$ is a bilinear map with the following property.

- Non-degenerate: $e(g, g)$ has the order p where $g \in G_g$.
- Bilinear: $e(ga, hb) = e\ (ga, hb)$ for all $a, b \in Zp$ and $g \in G_g 1,\ h \in G_g 2$.

The following set of algorithms is used in the proposed approach.

Setup $(1\lambda) \rightarrow (PK, M_K)$*:* The setup algorithm takes the input parameter λ as input and runs the group generator $\mathbf{G}(1\lambda)$ to get $(p, G_g 1, G_g 2, G_g T, e)$. It outputs the Master Key M_K and Public Key PK of the system. The master key is $M_K = (,h,)$ where $R\,G2$ and the system public key is $PK = (p, g, U_1, U_2, ..., U_p, ga, Y)$ where Y is defined as $e\,(g, g)$, $H = {}$ $= YZ, g, h, U_1, U_2, ..., U_p$ and an $R\,Zp$.

SW. Encryption $(PK, M, A) \rightarrow CA$: The message M, the access structure A, and the system's public parameters PK are all inputs to the encryption method (A, P, T). The cipher text CA is produced.

$$CA = (A, P), C', \{TI\}\ i \in S\},$$

where

$$C' = SW.LEA(M)\ .Y^S \ \text{ and } \ \{TI\} = g^{t1}, g^{t2}, \ldots, g^{t|u|}.$$

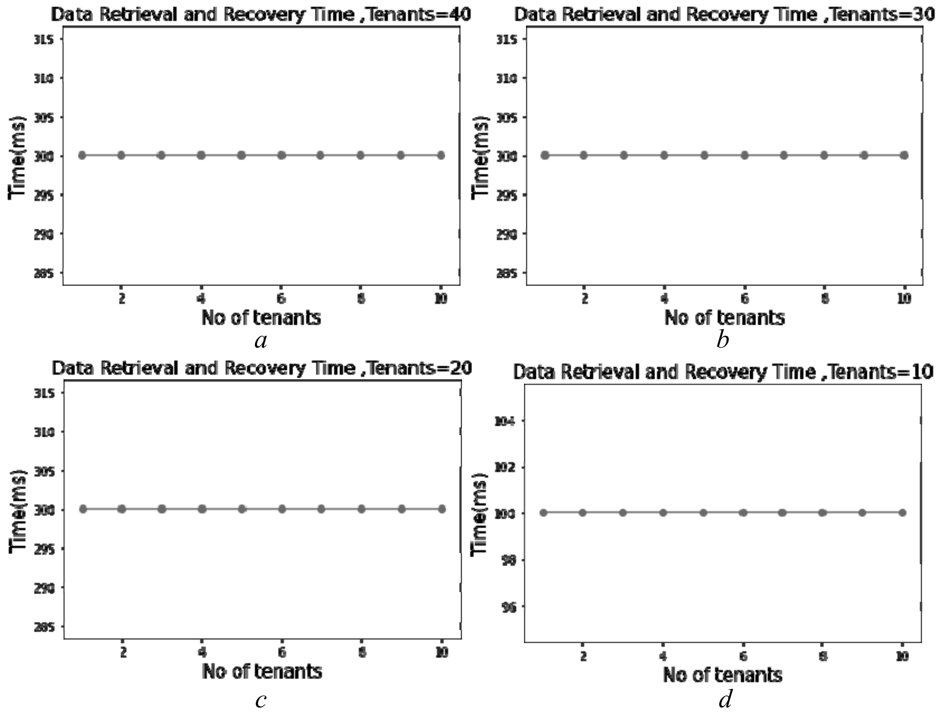
SW. Decryption $(PK, CA, SKS) \rightarrow M$ or $\perp$: the decryption algorithm takes the public key, plain text, and secret key as inputs and produces M if the access structure is satisfied otherwise. The real *SHR* is only decrypted if the characteristics match the access structure defined for that specific *SHR*; otherwise, access is denied to that particular user. This is done at the initial stage of the decryption process. From (A, P) derive IA, P, where IA, P stands for the minimal subset of

$1,...,l$ that meets the (A,P). Here, examine the possibility of a $I\,IA,P$ satisfying $C1 = e$ $(SKS,I$ over (A,P).

Where $\Sigma\omega_i A_i = (1,0,\ldots,0)$ otherwise it returns $\perp$. Here (A,P,T) is hidden, and in which (A,P), is revealed. If $i \in x$ satisfied the plain text M is given as output. $M = SW.LEA(C_1)/Y^S$, where M is in plain text, and C_1 is M's encrypted form.

PERFORMANCE ANALYSIS

This section addresses the results of the implementation and the performance of our proposed system. The proposed system is implemented using the Cloud Sim framework. The fundamental goal of the suggested technique is to ensure security and privacy. Novel algorithms are employed to accomplish this goal of policy concealing and the performance of the model is discussed here (Fig. 5).



a *b* *c* *d*

Fig. 5. Data retrieval and recovery time: *a* — tenants 40; *b* — tenants 30; *c* — tenants 20; *d* — tenants 10

The user's terminal will return and decrypt a variety of matching files using a data sharing scheme. There are 0 to 10 files that have been matched. Tenants range in number from 10 to 40. The value of recovery time, which ranges from 100 to 300, is discovered to remain constant for any number of renters.

The amount of time a user must wait after making a keyword token query in order to receive the retrieved health records must be carefully considered. All the data are calculated for 10 tenants. The test time of 1800 ms is observed for 10 tenants. The trace of the system is range from 50 to 500 ms. The time for token generation is observed to be constant at 30 ms. The decryption time for the model was found as 25 ms. and the encryption time for the model is estimated as 50 milliseconds. The key generation time for the model is nearly 25 ms and it increases to 140 ms for 10 tenants (Fig. 6).

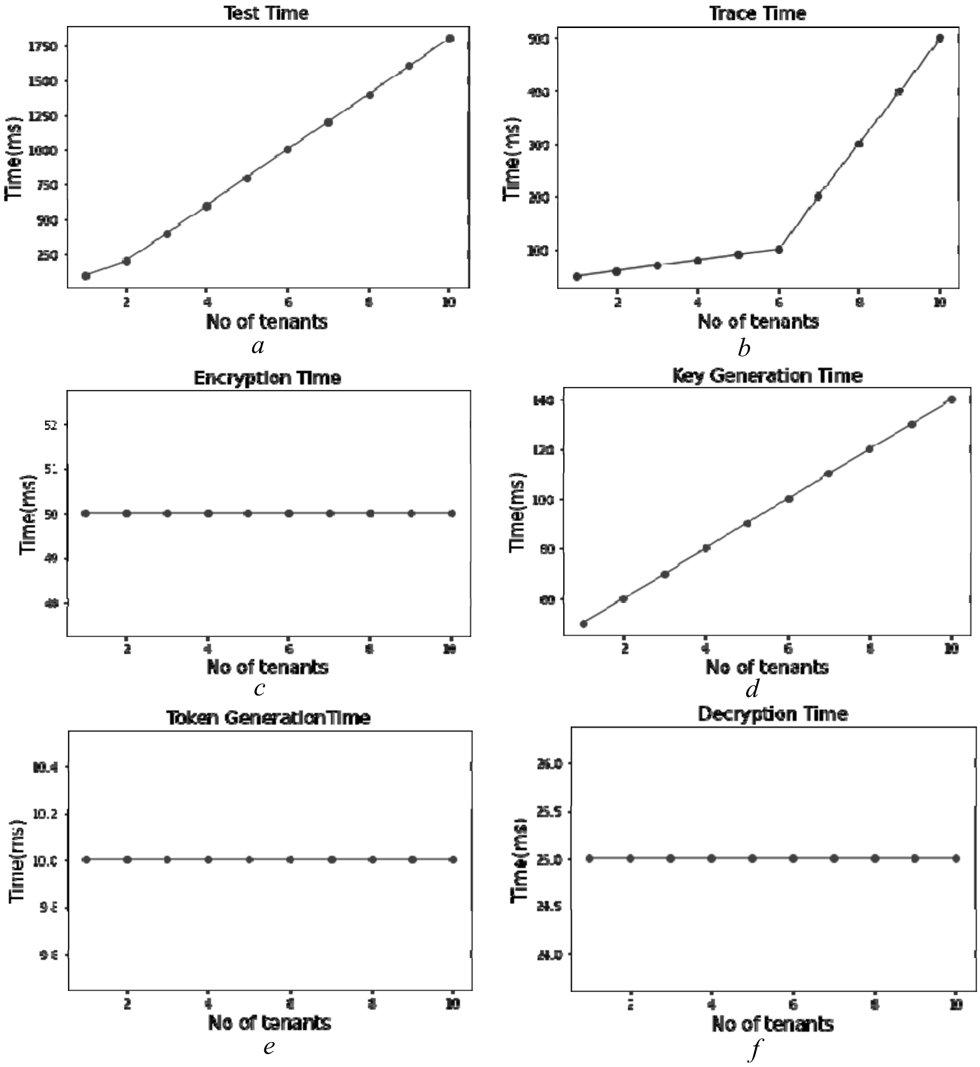


*Fig. 6***.** Computation efficiency

To evaluate the storage and transmission overhead, there is a need to identify the performance based on public and private keys and the token size in Fig. 7. The research found that for the public parameter 7, *a* requires substantially limited space and communication expenses. The key parameter size is 5000 bits, regardless of how many tenants are allowed in the system. For *b* secret key size is increasing from 1000 to 10000 bits with an increase in the number of tenants.

It is shown that 7, *c* requires the least amount of cipher text storage, which is especially useful for saving money under the pay-for-use cloud model. Additionally, by transferring the ciphertext to the public cloud using less battery power, the data owner may increase the lifespan of the user's mobile devices. The token size obtained in 7, *d* illustrates that the size remains constant as still figure of tenants rises.

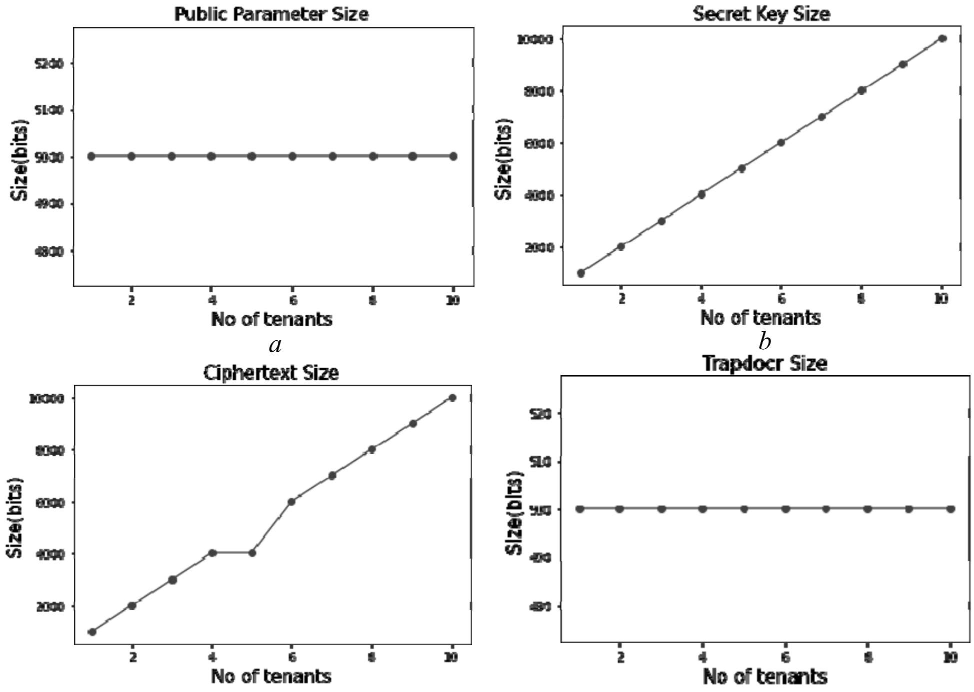


Fig. 7. Storage and transmission efficiency

The performance of the scheme in terms of accuracy, similarity, sensitivity, error, and false positive rate is depicted in the above graph. A heterogeneous, time-series set of data with ten classes of data was utilized. It is observed that the accuracy is obtained to be 90 %. The similarity range is obtained as 21%. The sensitivity of the system for any input change is depicted to be 98% for any given number of attributes. The calculation of the false positive rate (FPR) with the following formulae. The error is determined to be low as 10% and the FPR was found to be 1% for the proposed system (Fig. 8).

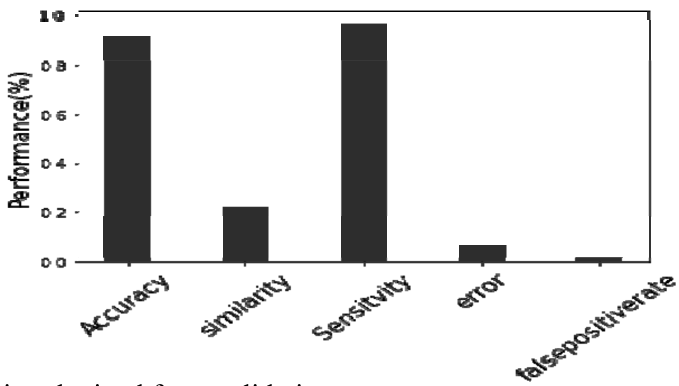


Fig. 8. Metrics obtained from validation

$$FPR = \frac{FP}{FP + TN}.$$

Comparison Metrics

In this part, the validity of the recommended methodology is assessed with the working of various traditional methodologies (Fig. 9).

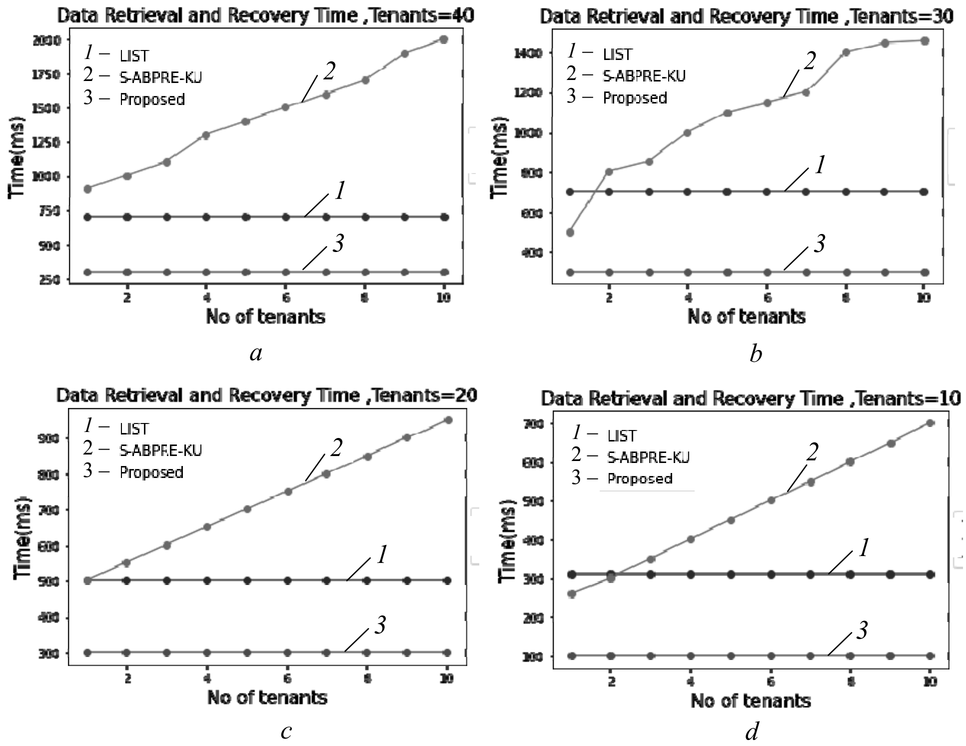


Fig. 9. Comparison of data retrieval and recovery time: *a* — Tenants=40; *b* — Tenants = 30; *c* — Tenants=20; *d* — Tenants=10

The data retrieval and recovery time of the proposed method is evaluated with other existing methods like LIST [30] and SABPRE [27]. For all the other methods the value of time with respect to the increase in tenants is observed to be comparatively high. The maximum time observed for 10 tenants is 710 ms for the LIST model whereas for the proposed model it is 100. The recovery time for 20 tenants is found to range from 500 to 1000 ms for the other methods and for the proposed method it is found to be 300 ms. The time is constant for 30 tenants for LIST as 700 and for SABPRE it is ranging from 500 to 1500 ms. For 40 tenants also the proposed method achieved less recovery time comparatively which is 350 ms.

The computation time for the model is compared for test time, trace time, token generation time, decryption time, encryption time and key generation time with existing models like LIST [30], SABPRE [27], ABKS [26], TCPABE [29] and LUCP ABE [28] etc.

For test and trace time the comparison is with LIST, SABPRE and ABKS. It is observed that the test duration from 1 to 1750 ms and trace duration from 1 to 200 ms which is minimum for the proposed model. The maximum values for the test time ranges from 1 to 2500 for the ABKS method and for trace the value ranges from 1 to 5000 ms for LUCPABE (Fig. 10).

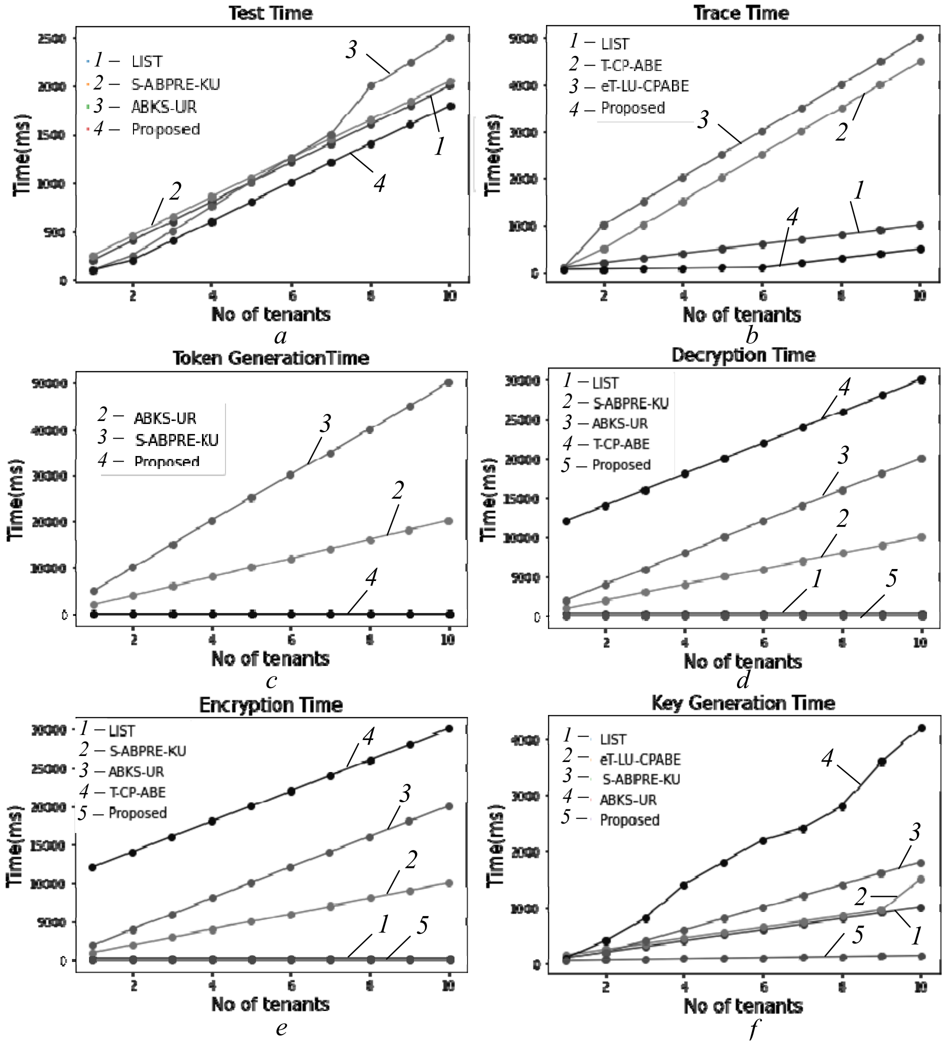


Fig. 10. Comparison of computation time

The token generation time for the proposed method is minimum when compared with others. The maximum values for the same with 10 tenants are found to be 50000 ms for the SABPRE method whereas for the proposed method it is 1 ms. The encryption and decryption times are compared in which the encryption and decryption time for the proposed method is constant for any number of tenants for other methods it linearly varies. The maximum value is obtained as 30000ms whereas for the proposed method it is 1 ms. For the key generation time, the value is 1 ms for any tenant value. But for other methods, it is constantly increasing with the increase in tenant number. The maximum value is 4500 ms for the ABKS method.

Storage and transmission overhead is compared with existing models in which the proposed model makes constant size irrespective of the tenant number, whereas all other methods vary with tenant size. The proposed method has the least size for all the storage and transmission sizes. All the size values range from 1 to 300000 bits with the proposed system (Fig. 11).

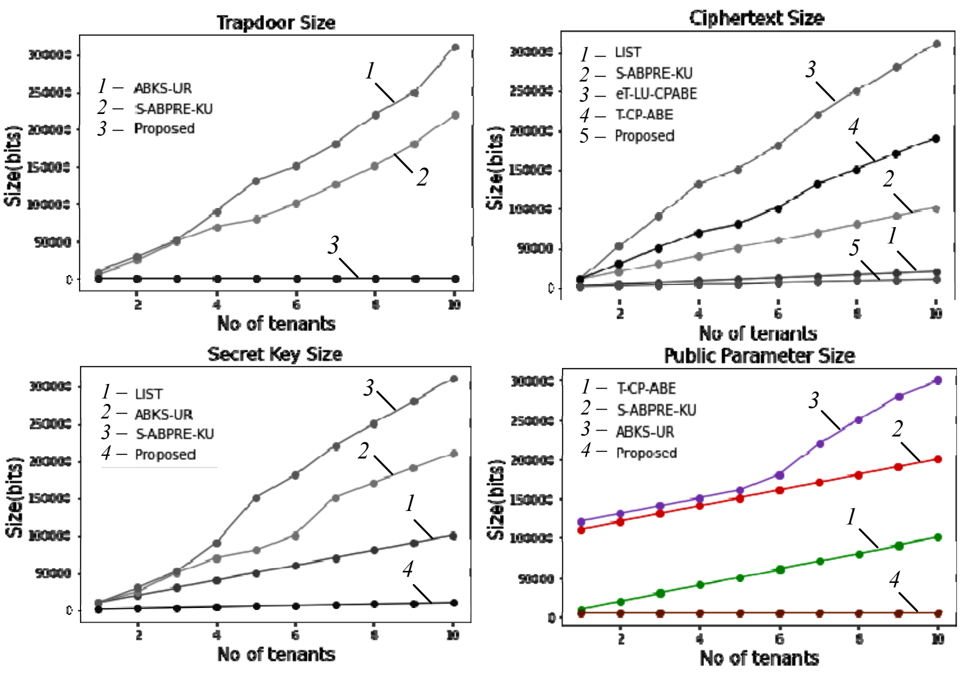


Fig. 11. Storage and transmission comparison

CONCLUSION

This study presented a data-sharing and revocation system based on ciphertext and ECC. Without giving any personal information to the cloud service provider, the majority of the decryption computational burden was transferred. Each authorized user employs a one-time password to decode the cipher text and a private secret key to decode the original cipher text. Thus, only an authorized user's secret key and session key may be used to view the plaintext communication. To tackle Sybil and resource depletion threats in a cloud context, another novel methodology was presented. This method not only prevents Sybil attacks within the planes of control and data from overwhelming cloud infrastructure, but it in turn enhances the level of service delivered to cloud users. Identification of Sybil nodes created as a result of the Sybil attack is done, and all authorized entities detected in the smart health system are informed of the updated revocation list. The system performed 90% accurately, according to the results with 21% of similarity found. The sensitivity of the system for any input change is depicted to be 98% for any given number of attributes. The error is determined to be low as 10% and the false positive rate is found to be 2% for the system and the system performance was compared with other existing techniques.

REFERENCES

1. Remya Sivan and Zuriati Ahmad Zukarnain, "Security and privacy in cloud-based e-health system," *Symmetry,* vol. 13, no. 5, pp. 742, 2021.
2. Umar Zaman et al., "Towards Secure and Intelligent Internet of Health Things: A Survey of Enabling Technologies and Applications," *Electronics,* vol. 11, no. 12, pp. 1893, 2022.
3. Tuan Minh Nguyen Dang, Nhan Nguyen Thanh, and Loc Le Tuan, "Applying a mindfulness-based reliability strategy to the Internet of Things in healthcare–A busi-

ness model in the Vietnamese market," *Technological Forecasting and Social Change,* vol. 140, pp. 54–68, 2019.

4. Kavita Jaiswal and Anand Veena, "A survey on IoT-based healthcare system: potential applications, issues, and challenges," *Advances in Biomedical Engineering and Technology*. Springer, Singapore, pp. 459–471, 2021.
5. Ying Jin et al., "Identifying human body states by using a flexible integrated sensor," *npj Flexible Electronics,* vol. 4, no. 1, pp. 1–8, 2020.
6. L. Minh Dang et al., "A survey on internet of things and cloud computing for healthcare," *Electronics,* vol. 8, no. 7, pp. 768, 2019.
7. Sherali Zeadally et al., "Smart healthcare: Challenges and potential solutions using internet of things (IoT) and big data analytics," *PSU research review,* 2019.
8. Shahidinejad Ali, Mostafa Ghobaei-Arani, and Mohammad Masdari, "Resource provisioning using workload clustering in cloud computing environment: a hybrid approach," *Cluster Computing,* vol. 24, no. 1, pp. 319–342, 2021.
9. M. Saleh Altowaijri, "An architecture to improve the security of cloud computing in the healthcare sector," *Smart Infrastructure and Applications.* Springer, Cham, pp. 249–266, 2020.
10. L. Pallavi, A. Jagan, and B. Thirumala Rao, "Mobility Management Challenges and Solutions in Mobile Cloud Computing System for Next Generation Networks," *International Journal of Advanced Computer Science and Applications,* vol. 11, no. 3, 2020.
11. Li Xiaoqing et al., "Building the Internet of Things platform for smart maternal healthcare services with wearable devices and cloud computing," *Future Generation Computer Systems,* vol. 118, pp. 282–296, 2021.
12. G. Lavanya et al., *IoT Enabled Assisting Device for Seizures Monitoring*. 2019, pp. 10–14.
13. M. Anuradha et al., "IoT enabled cancer prediction system to enhance the authentication and security using cloud computing," *Microprocessors and Microsystems*, vol. 80, pp. 103301, 2021.
14. Thilakarathne Navod Neranjan, Mohan Krishna Kagita, and Thippa Reddy Gadekallu, "The role of the internet of things in health care: a systematic and comprehensive study," *Available at SSRN,* pp. 3690815, 2020.
15. Mehrotra Preeti, Preeti Malani, and Prashant Yadav, "Personal protective equipment shortages during COVID-19—supply chain–related causes and mitigation strategies," *JAMA Health Forum*, vol. 1, no. 5, 2020.
16. Verma Deepak Kumar and Tanya Sharma, "Issues and challenges in cloud computing," *Int. J. Adv. Res. Comput. Commun. Eng.,* vol. 8, pp. 188–195, 2019.
17. Siyal Asad Ali et al., "Applications of blockchain technology in medicine and healthcare: Challenges and future perspectives," *Cryptography,* vol. 3, no. 1, pp. 3, 2019.
18. Kumar Y. Kiran and R. Mahammad Shafi, "An efficient and secure data storage in cloud computing using modified RSA public key cryptosystem," *International Journal of Electrical and Computer Engineering,* vol. 10, no. 1, pp. 530, 2020.
19. Hans-Georg Eichler et al., "Data rich, information poor: can we use electronic health records to create a learning healthcare system for pharmaceuticals?," *Clinical Pharmacology & Therapeutics,* vol. 105, no. 4, pp. 912–922, 2019.
20. Yan Shengguang et al., "Concurrent healthcare data processing and storage framework using deep-learning in distributed cloud computing environment," *IEEE Transactions on Industrial Informatics,* vol. 17, no. 4, pp. 2794–2801, 2020.
21. Hao Meng et al., "Privacy-aware and resource-saving collaborative learning for healthcare in cloud computing," *ICC 2020-2020 IEEE International Conference on Communications (ICC).* IEEE, 2020.
22. Yudi Zhang et al., "Efficient identity-based distributed decryption scheme for electronic personal health record sharing system," *IEEE Journal on Selected Areas in Communications,* vol. 39, no. 2, pp. 384–395, 2020.
23. Akhilendra Pratap Singh et al., "A novel patient-centric architectural framework for blockchain-enabled healthcare applications," *IEEE Transactions on Industrial Informatics,* vol. 17, no. 8, pp. 5779–5789, 2020.

24. Xu Yang et al., "Efficient and anonymous authentication for healthcare service with cloud based WBANs," *IEEE Transactions on Services Computing,* 2021.
25. Seunghwan Son et al., "Design of secure authentication protocol for cloud-assisted telecare medical information system using blockchain," *IEEE Access,* vol. 8, pp. 192177–192191, 2020.
26. Wenhai Sun et al., "Protecting your right: Verifiable attribute-based keyword search with fine-grained owner-enforced search authorization in the cloud," *IEEE Transactions on Parallel and Distributed Systems,* vol. 27, no. 4, pp. 1187–1198, 2014.
27. Kaitai Liang and Susilo Willy, "Searchable attribute-based mechanism with efficient data sharing for secure cloud storage," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1981–1992.
28. Jianting Ning et al., "White-box traceable ciphertext-policy attribute-based encryption supporting flexible attributes," *IEEE Transactions on Information Forensics and Security,* vol. 10, no. 6, pp. 1274–1288, 2015.
29. Liu Zhen, Zhenfu Cao, and S. Duncan Wong, "White-box traceable ciphertext-policy attribute-based encryption supporting any monotone access structures," *IEEE Transactions on Information Forensics and Security,* vol. 8, no. 1, pp. 76–88, 2012.
30. Yang Yang, et al., "Lightweight sharable and traceable secure mobile health system," *IEEE Transactions on Dependable and Secure Computing,* vol. 17, no. 1, pp. 78–91, 2017.
31. Q. Zhang, Q. Liu, and G. Wang, "PRMS: A personalized mobile search over encrypted outsourced data," *IEEE Access*, 6, pp. 31541–31552, 2018.

Received 21.10.2022

INFORMATION ON THE ARTICLE

Dipa D. Dharmadhikari, ORCID: 0000-0001-7206-5573, CSIT Dr. Babasaheb Ambedkar Marathwada University, Aurangabad, India, e-mail: scholar.dipaddharmadhikari@gmail.com

Dr. Sharvari Chandrashekhar Tamane, ORCID: 0000-0001-9350-1549, MGM's Jawaharlal Nehru Engineering College, MGM University, Aurangabad, India

РОЗШИРЕНА СХЕМА БЕЗПЕКИ ДЛЯ СПІЛЬНИХ ДИНАМІЧНИХ ДАНИХ З ЕФЕКТИВНОЮ ЛЕГКОЮ ЕЛІПТИЧНОЮ КРИПТОГРАФІЄЮ / Діпа Д. Дхармадхікарі, Шарварі Чандрашекхар Тамане

Анотація. Технологія хмарних обчислень прогресує, тому Cloud Computing (CC) створює різноманітні хмарні сервіси (CS). Користувачі можуть отримати простір для зберігання від постачальника, оскільки послуги хмарного зберігання досить практичні; багато користувачів і компаній зберігають свої дані у хмарному сховищі. Конфіденційність даних стає більшим ризиком для постачальників послуг, коли більше інформації передається CS. У роботі підхід до шифрування тексту та еліптичної кривої (ECC) із шифруванням на основі ідентифікації (CP-IBE) використовується у хмарному середовищі для забезпечення безпеки даних у середовищі охорони здоров'я. Проблема відкликання стає складною, оскільки характеристики використовуються для створення шифрованих текстів і секретних ключів, отже, вводиться алгоритм відкликання користувача, для якого секретний ключ маркера унікально створюється для кожного рівня, що забезпечує безпеку. Початкова операція, включаючи підпис, публічні перевірки, динамічні дані, чутливі до атак Sybil, для подолання цього вводиться алгоритм перевірки атак Sybil, який ефективно захищає систему. Крім того, умови для публічного аудиту з використанням спільних даних і типових стратегій, включаючи аналітичну функцію, безпеку та умови продуктивності, аналізуються щодо точності, чутливості та подібності.

Ключові слова: зашифрований текст, відкликання користувача, обмін даними, CC, ECC, питання безпеки.

UDC 004.896
DOI: 10.20535/SRIT.2308-8893.2023.3.03

RESEARCH ON HYBRID TRANSFORMER-BASED AUTOENCODERS FOR USER BIOMETRIC VERIFICATION

M.P. HAVRYLOVYCH, V.Y. DANYLOV

Abstract. Our current study extends previous work on motion-based biometric verification using sensory data by exploring new architectures and more complex input from various sensors. Biometric verification offers advantages like uniqueness and protection against fraud. The state-of-the-art transformer architecture in AI is known for its attention block and applications in various fields, including NLP and CV. We investigated its potential value for applications involving sensory data. The research proposes a hybrid architecture, integrating transformer attention blocks with different autoencoders, to evaluate its efficacy for biometric verification and user authentication. Various configurations were compared, including LSTM autoencoder, transformer autoencoder, LSTM VAE, and transformer VAE. Results showed that combining transformer blocks with an undercomplete deterministic autoencoder yields the best performance, but model performance is significantly influenced by data preprocessing and configuration parameters. The application of transformers for biometric verification and sensory data appears promising, performing on par with or surpassing LSTM-based models but with lower inference and training time.

Keywords: biometric verification, transformers, variational autoencoder, transformer autoencoder.

INTRODUCTION

The usage of various deep learning algorithms boosted and enabled various AI and machine learning fields and applications. The biometric field was no exception, specifically with the growth and significant adoption of various electronic devices such as smartphones, bracelets, watches, etc. One of the important areas where biometric data is utilised is security, verification and authentication. Much research was conducted in this field to discover and provide deep learning architectures that will be able to build efficient and reliable systems feasible for usage in real life.

Traditional methods, such as passwords and PINs, are prone to breaches and hacking, as well as are challenging to manage, which lead us to the exploration of more secure and user-friendly alternatives. However, the effectiveness of biometric verification is contingent on the ability to process and interpret complex biometric data accurately. Deep learning approached, which can generalize over large data samples and be high-performant, is a solution to solve the problem. Specifically, combining autoencoders and transformer attention layers, a novel deep learning approach, has shown promise in enhancing the performance of biometric verification systems. However, this approach is still not widely presented in biometric verification and continuous authentication research.

The relevance of this research lies in developing more secure and efficient user authentication methods. By enhancing the performance of biometric verification systems, we can provide a more secure and convenient alternative to traditional authentication methods.

The object of this research is the application of autoencoders combined with transformer attention layers in biometric verification and continuous authentication.

This study investigates the effectiveness of autoencoders combined with transformer attention layers for biometric verification and continuous authentication. We aim to assess whether this novel approach can improve the performance and efficiency of biometric verification systems, thereby contributing to the development of more secure and user-friendly authentication methods.

LITERATURE REVIEW

In [1], the authors convey an in-depth survey on which deep learning and machine learning models are used for biometric verification. There is extensive research on hybrid models, such as extracting features with the CNN model and conducting authentication with some machine learning models, such as SVM or One-Class SVM or LSTM block with further Stochastic Gradient Descent (SGD) classifier. Another quite popular solution is using LSTM model architecture, which is self-explainable as biometric in many cases is sensory data with a sequential structure. Specifically for the motion or gait patterns, the hybrid architecture LSTM + CNN is popular, which outperforms the LSTM or CNN separately [1; 2]. Overall it is noticed that hybrid architectures provide a boost in performance and are widely adopted in biometric authentication. It is worth noting that there is no clear distinction between supervised and unsupervised approaches in the paper, and all of them are compared altogether, which is essential for the context of the constraints and limitations of the implemented verification system. Our interest is in unsupervised approaches as they provide a solution in real cases when there is no access to other users' data (as it will be due to data privacy), contrary to supervised models.

The data nature causes the popularity of LSTM applications for sensory data, but not only recurrent architectures can handle sequences. The transformer architecture [3] was initially adopted in natural language processing (NLP) tasks and almost replaced the recurrent neural networks in that field [4].

Transformers' way of consuming sequences provided faster training and inference and better generalization capabilities for sequences as it does not have an issue of forgetting input in case of long input, as the sequence was consumed as a whole instantly and not chunk by chunk. On the other hand, the architecture requires fixed sequence length and sequences with lengths higher than the model support will not be processed. As the transformers were great with sequence data — they slowly started being used in other fields, such as CV and time series. In the [5; 6], authors review the effectiveness of transformers for time series data and compare various transformer types, which show pretty decent results.

Nevertheless, RNNs are still holding their place in the time series field, as they are better at capturing the autoregressive nature of time series signals. Both models have pros and cons, and at the end of the day, each can bring something to the table. In [7], authors show that LSTM with attention layer outperforms the transformer-based model for time series tasks, which supports the idea that hybrid models create more performant and robust deep learning systems.

In biometric fields, transformers were used for human activity recognition (HAR) problems [8; 9]. The authors proposed a HAR transformer, which solves the time series classification problem.

The choice of the approach and model architecture for biometric verification depends on which type of authentication system we want to build. Authentication can be implicit and explicit, as well as continuous or more discrete. We are interested in implicit continuous authentication, generally the unsupervised approach. The overall model architecture used for such tasks is autoencoder. We have reviewed and experimented with the usage of autoencoders for biometric verification tasks in our previous research [10]. In another paper, we reviewed which sensor data signal contributes the most to creating a distinctive user pattern [11].

In [12], the VAE-based system was proposed to solve the text keystroke authentication when the training is done on the English typing data and evaluating the Korean typing data from the same users. This may show that the model learns the pattern of the user uniqueness and not the different patterns related to activities. A deep LSTM-based autoencoder is proposed in [13] for anomaly detection in ECG signals. In contrast, in [14], adversarial autoencoder [15], which is the combination of autoencoder with generative adversarial networks (GAN), was used for the health monitoring of ECG and for detecting abnormal data points, which by the authors outperformed LSTM and VAE architectures. The autoencoder with attention mechanism, placed between encoder and decoder blocks to learn relations on the latent space feature representations, is proposed in [16] for ECG data anomaly detection.

However, the LSTM-based architecture still was more performant and better at capturing time series data. In [17], the authors proposed the attentive adversarial autoencoder for user authentication. Compared to approaches like one-class SVM, LSTM and HMM, the autoencoder-based solution achieved the highest performance in terms of qualitative metrics and time performance. In [18], the purely transformer-based architecture is used for detecting anomalies in ECG series, which is also shown to be a viable option.

Autoencoder and its various modification of it are widely used and researched in the area of intelligent fault diagnosis and prognosis for industrial systems [19]. In this area, autoencoders help to prevent system failure processing, like wind turbine equipment or other complex systems, processing the multiple modality data, such as acoustic and vibration signals [20]. Stacked autoencoder architecture is quite famous for fault diagnosis, where multiple encoders and decoders are stacked on top of each other, which may help the neural network to recognize data trends and patterns better.

We want further review and experiment with various autoencoder-based architecture sand specifically review the possibility of incorporating elements from other architecture to see whether it will impact the performance. As the transformer-based architecture is still state-of-the-art in many fields, though it was proposed some time ago, and multiple other research incorporate it for various biometric-related tasks, such as health monitoring – we would like to experiment with how it will impact metrics in biometric verification tasks, and whether it will reduce the inference time, as a transformer, due to the way how they process sequence should be faster than RNN.

MATERIALS AND METHODS

As a baseline model with which we will compare other experiments, an LSTM autoencoder will be used. The autoencoder is an artificial neural network for learning hidden internal representations and features of input data. It consists of two main parts: an encoder that compresses the input into a latent-space representation and a decoder which reconstructs the input from the latent space. During training autoencoder learns to minimize the difference between the input and the reconstructed output. The optimization task objective is to minimize this difference, called the reconstruction error:

$$E = \sqrt{\sum_{i=1}^{n} \left\| x_i - d_{\varphi}(e_{\theta}(x_i)) \right\|} ,$$

where $x_1, \ldots, x_n$ is data rows, and the functions d_{φ} and e_{θ} represent the encoder and decoder, respectively, with some parameters φ and θ.

Autoencoder can be considered as a high-level neural network architecture, as it does not limit what architectural elements should or should not be in the encoder and decoder. However, there are some types of autoencoders that specify some limitations on the architecture of the autoencoder or some of its configurations. For example, a sparse autoencoder should have a dimension of latent space higher than the input dimension; the denoising autoencoder puts the requirement for adding noise to the input data; the contractive autoencoder specifies the optimization loss.

Variational Autoencoder (VAE) is somewhat different from other autoencoder types, as it maps the input data not to the fixed latent space representation, but the Gaussian distribution with some parameters (mean and variance). Thus, it allows us to present our input data points in probabilistic manner. This model architecture is close to the generative AI algorithms we reconstruct our data sampling it from out latent distribution, so in fact generating it [21].

The encoder part of the VAE is defined as:

1. Encoder:

$$\mu = W_{\mu} * h + b_{\mu} ;$$

$$\log(\sigma^2) = W_{\sigma} * h + b_{\sigma} .$$

2. Reparameterization Trick:

$$z = \mu + \sigma \odot \varepsilon , \text{ where } \varepsilon \sim N(0, I) .$$

3. Decoder:

$$p(x \mid z) = f_{\text{dec}}(z) .$$

4. Loss Function:

$$L = E[\log p(x / z)] - D_{KL}(Q(z / x) || P(z)) ,$$

where h is the output of the encoder's hidden layer; W_{μ}, W_{σ}, b_{μ}, and b_{σ} are the weights and biases for the mean and log-variance, respectively; μ and σ are the mean and standard deviation of the latent variable z; ε is a random variable sampled from a standard normal distribution; $\odot$ denotes element-wise multiplication;

f_{dec} is the decoder function; $p(x/z)$ is the probability of the data given the latent variable; $Q(z/x)$ is the approximate posterior distribution; $P(z)$ is the prior distribution (standard normal distribution in the case of VAEs); $D_{KL}(...)$ is the Kullback–Leibler divergence, which measures the difference between two probability distributions; $E[\ldots]$ denotes the expectation; L is the loss function that the VAE aims to minimize.

These formulas represent the core of the VAE. The encoder generates the parameters of the latent variable's distribution, the reparameterization trick is used to sample from this distribution, and the decoder generates the data from the sampled latent variable. The loss function consists of the reconstruction loss (the first term) and the regularization term (the second term).

Neural network building blocks. As autoencoder is a high-level architecture – it may be constructed from any neural network units which are suitable for the given problem and data input.

Long Short-Term Memory (LSTM). LSTM is a type of recurrent neural network (RNN) that can learn and remember over long sequences and is not that by the vanishing gradient problem, as just RNN. It achieves this by using a series of "gates". These blocks collectively decide what information should be kept or discarded.

The LSTM cell can be defined by the following set of equations:

Forget gate:

$$f_t = \sigma(W_f * (h_{t-1}, x_t) + b_f).$$

Input gate:

$$i_t = \sigma(W_i * (h_{t-1}, x_t) + b_i).$$

Cell update:

$$\widehat{C}_t = \tanh(W_c * (h_{t-1}, x_t) + b_c).$$

New cell:

$$C_t = f_t * C_{t-1} + i_t * \widehat{C}_t.$$

Output gate:

$$o_t = \sigma(W_o * (h_{t-1}, x_t) + b_o).$$

New hidden state:

$$h_t = o_t * \tanh(C_t).$$

Where σ is the sigmoid function, (h_{t-1}, x_t) denotes the concatenation of the input vector x_t and the previous hidden state h_{t-1}, and W and b are the weight matrices and bias vectors.

Transformers (attention unit). Transformers are a type of model that uses self-attention mechanisms and are particularly effective for tasks involving sequential data. Unlike RNNs, transformers do not require that the sequence data be processed in order, thus allowing for parallel processing of the data.

The self-attention mechanism in transformers can be defined as:

$$Q = W_q * X, K = W_k * X, V = W_v * X.$$

$$\text{Attention}(Q,K,V) = \text{soft}\max\left(\frac{Q * K^T}{\sqrt{d_k}}\right) * V .$$

Where Q, K, and V are the query, key, and value vectors, and d_k is the dimension of the key vector. The softmax function ensures that the weights of the different positions sum to 1.

EXPERIMENTS

Dataset. Open-source dataset [22–24], a large-scale user study with 100 volunteers to collect a wide spectrum of signals about smartphone user behaviors, including touch, gesture, and pausality of the user, as well as movement and orientation of the phone. Data from three usage scenarios on smartphones were recorded: 1) document reading; 2) text production; 3) navigation on a map to locate a destination.

The dataset contains multiple modalities input from various sensors. For our experimentation, we selected the accelerometer, gyroscope and magnetometer inputs in the dataset.

The dataset contains multiple activities, such as read and walking, read and sitting, write and walking, write and sitting, navigate the map and walking and navigate the map and sitting – overall 6 activity types. We have trained our models on some selected activity type, as well as on activity pair, like reading, navigating the map or writing and activity triplet, like sitting or walking.

For deep learning models, we split data in overlapping on 50 percent windows with a sampling of 100Hz and a length of 1s.

The original dataset is split into a 20% share for the test set and the rest for the train.

We preprocessed data in 2 ways: standart dcaling and min-max normalizing.

Sensors description. An accelerometer measures changes in velocity along one axis. The values reported by the accelerometers are measured in increments of the gravitational acceleration, with the value 1.0 representing an acceleration of 9.8 meters per second in the given direction. Depending on the direction of the acceleration, the sensor values may be positive or negative. A gyroscope measures the rate at which a device rotates around a spatial axis and is used to detect or measure direction. The magnetometer measures the strength of the magnetic field surrounding the device, allowing us to detect the device's orientation correctly [25; 26].

Metrics. The threshold formula was used as in [10]:

$$T = \sum_{i=1}^{N} MAE_i \Big/ N + std(MAE_i) ,$$

where MAE is the mean absolute error between ground truth and predicted sample; std – standard deviation; and N is the number of samples in the training dataset.

As model evaluation metrics [27], the EER (equal error rate); FAR (false accept rate) and FRR (false reject rate) were chosen, which are typical for assessing the biometric system quality:

$$\mathrm{FAR} = \mathrm{FPR} = \mathrm{FP/FP} + \mathrm{TN}\ ;$$

$$\mathrm{FRR} = \mathrm{FNR} = \mathrm{FN/TP} + \mathrm{FN}\ .$$

Equal error rate is obtained by adjusting the system's detection threshold to equalize FAR and FRR. The EER is calculated using the following formula:

$$\mathrm{EER} = \mathrm{FAR} + \mathrm{FRR}/2\ ,$$

where $|\,\mathrm{FAR} + \mathrm{FRR}\,|$ is the smallest value [27].

The models were coded and trained in Python using Keras library with Tensorflow backend.

All models were trained in 20 epochs with Adam optimiser on the GeForce RTX 2070 GPU.

The architecture of transformer-based hybrid autoencoder used for experiments illustrated in Fig. 1.

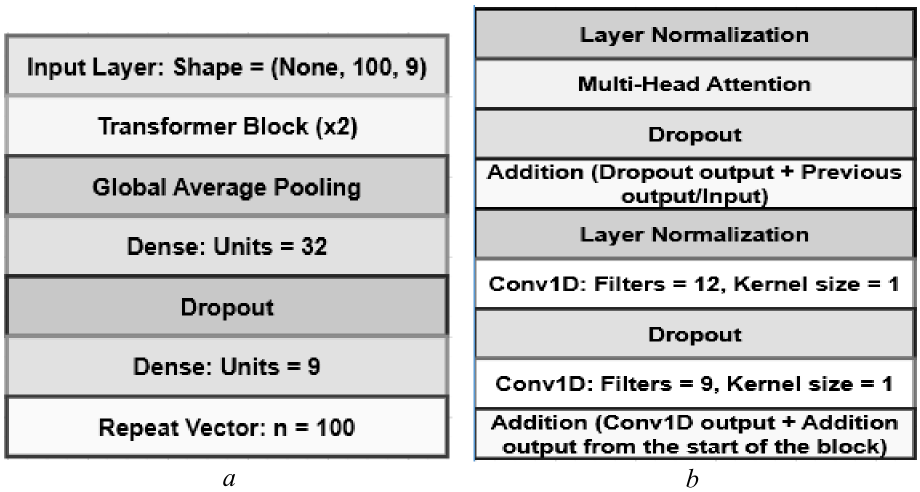


Fig. 1. Architecture of transformer-based hybrid autoencoder: *a* — the high-level autoencoder architecture with transformer encoder; *b* — the internal structure of transformer-based encoder with attention units

The LSTM autoencoder architecture with which the transformer-based autoencoder was compared is illustrated on Fig. 2.

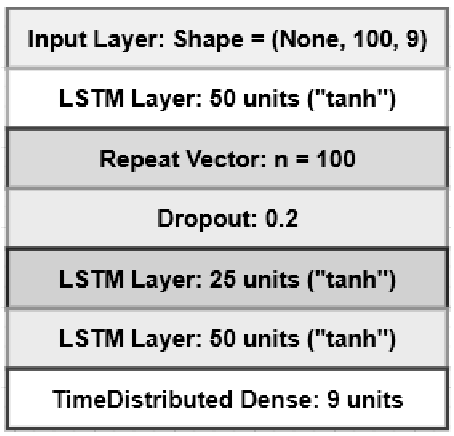


Fig. 2. LSTM autoencoder architecture

RESULTS

The experimentation results can be reviewed in the tables below.

The results for the single activity with standart scaling data preprocessing and variational-based autoencoders can be reviewed in Table 1. In Table 2 we can view the model performance for activity pairs, and results for acitivy triplet in Table 3. For the deterministic models the data was processed with min-max normalization.

As well we can view the performance time for training and inference for models in Table 4. Overall in each exeperiment for each chosen activity set data 100 model were trained.

Table 1. Average EER, FAR, FRR for 100 users for single activity

Model architecture	Average EER	Average FAR	Average FRR
Single activity — write and sitting			
LSTM VAE	5.10%	14.25%	3.28%
Transformer-VAE	4.20%	12.95%	1.72%

Table 2. Average EER, FAR, FRR for 100 users for activity pairs

Model architecture	Average EER	Average FAR	Average FRR
Activity Pair — write and walking, write and sitting			
LSTM AE	5.21%	13.30%	3.37%
Transformer AE	4.22%	13.93%	1.76%
Activity Pair – map and walking, map and sitting			
LSTM AE	6.74%	14.39%	5.00%
Transformer AE	5.87%	13.24%	3.42%

Table 3. Average EER, FAR, FRR for 100 users for activity triplet

Model architecture	Average EER	Average FAR	Average FRR
Activity Triplet — read and sitting, write and sitting, map and sitting			
LSTM AE	1.26%	12.38%	0.06%
Transformer AE	1.61%	10.97%	0.14%
Activity Triplet – read and walking, write and walking, map and walking			
LSTM AE	9.10%	12.66%	9.51%
Transformer AE	6.47%	12.37%	4.81%

Table 4. Average training and inference time for 100 users for sitting activity triplet

Model architecture	Training Time (*s*)	Inference Time (*s*)
LSTM AE (MSE loss)	90.67	43.32
Transformer-AE (MSE loss)	82.22	29.61
Difference	9.32%	31.65%

DISCUSSION

The obtained experiments results showed us that transformer architecture, specifically its central architectural unit as attention, provides performance improvement for the biometric verification task in the case of deterministic model version or generative (variational). The transformer-based autoencoder outperformed the LSTM based one in the case of training on single activity and activity pairs, which confirmed that the model performance is stable over different data inputs.

Though on sitting activity triplet, the LSTM AE slightly outperformed the Transformer AE in terms of EER and FRR but had a higher FAR rate. It shows us that the LSTM can generalize better with a larger train data sample. However, as well showing us that transformer-based autoencoders can generalize on smaller amounts of data.

It is worth noting that the transformer is significantly faster than the LSTM based model in terms of training and inference time; therefore, it is a much better fit for the edge devices like smartphones or smartwatches, where such models will be applied.

During the experimentation, we were also trying different losses and data preprocessing approaches and figured out that models are susceptible to the scale of the data input. The insightful observation was that deterministic models are great for generalization in the case of data normalization with min-max. However, in the case of standard scaling, the variational version generalizes better, which can happen due to multiple factors. First, min-max transformation can distort the data distribution in case of significant outliers in data; therefore, variational autoencoder that samples from Gaussian distribution with mean and variance will not be able to learn on the data that do not follow Gaussian distribution. On the other hand, the reason why deterministic models could not generalize well on standard scaled data was due to using as input multiple sensor signals, which may have different ranges and make it harder for neural networks that are sensitive to the range caused by the tanh activation function. Though this observation should be rigorously tested, it provides insights into how the data should be preprocessed for different architectures and how strongly the data format is coupled with the neural network.

CONCLUSIONS

We have conducted various experiments in this research and proposed and analysed the hybrid transformer-based autoencoder model architecture. The model was high-performant compared to the LSTM-based architecture and robust with different data inputs regarding amount and activity types.

Overall more than 800 neural networks were trained during the experimentation.

We have noticed that although the model architecture plays a significant part in the final metrics, the data pre-processing step is critical, and we cannot expect from deep learning model to generalise without preliminary steps. Depending on model internals, we should keep an eye on the validity of data distribution and the presence of noise and outliers in the dataset. Model type and data may also impact the selection of optimised losses, such as the used in our models' mean squared error or mean absolute error, which is more robust to the outliers, or the combination of both losses like Huber loss. During experimentation, we noticed that optimised loss may significantly add to the model's generalisation ability. However, this observation should be researched further to understand how model architecture connects with the different loss functions.

As further steps – we may consider creating the ensemble of the models in order to achieve the highest possible metric value. We can see that treating a neural network as a weak learner is possible. Though, it has a considerable amount of parameters – the discussion in the machine learning community makes us believe that it should be the auspicious direction in further neural network architecture development.

REFERENCES

1. S. Kokal, M. Vanamala, and R. Dave, "Deep Learning and Machine Learning, Better Together Than Apart: A Review on Biometrics Mobile Authentication," *Journal of Cybersecurity and Privacy*, vol. 3, no. 2, pp. 227–258, Jun. 2023. doi: 10.3390/jcp3020013.
2. Q. Cao, F. Xu, and H. Li, "User Authentication by Gait Data from Smartphone Sensors Using Hybrid Deep Learning Network," *Mathematics*, vol. 10, no. 13, p. 2283, Jun. 2022. doi: 10.3390/math10132283.
3. A. Vaswani et al., "Attention is All you Need," in *Advances in Neural Information Processing Systems (NIPS)*, vol. 30, 2017.
4. S.M. Lakew, M. Cettolo, and M. Federico, "A Comparison of Transformer and Recurrent Neural Networks on Multilingual Neural Machine Translation," in *International Conference on Computational Linguistics (COLING), 2018*.
5. A. Zeng, M. Chen, L. Zhang, and Q. Xu, "Are Transformers Effective for Time Series Forecasting?," in *Proceedings of the AAAI Conference on Artificial Intelligence, 2023*.
6. Q. Wen et al., "Transformers in time series: A survey," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 2023.
7. A. Katrompas, T. Ntakouris, and V. Metsis, "Recurrence and Self-attention vs the Transformer for Time-Series Classification: A Comparative Study," in Michalowski M., Abidi S.S.R., Abidi S. (eds) *Artificial Intelligence in Medicine (AIME)*, vol. 13263. Springer, Cham, 2022. doi: 10.1007/978-3-031-09342-5_10.
8. I. Dirgová Luptáková, M. Kubovčík, and J. Pospíchal, "Wearable Sensor-Based Human Activity Recognition with Transformer Model," *Sensors*, vol. 22, no. 5, p. 1911, Mar. 2022. doi: 10.3390/s22051911.
9. A. Raza et al., "Lightweight Transformer in Federated Setting for Human Activity Recognition," *ArXiv, arXiv:2110.00244*, Nov. 4, 2022. Accessed on: July 6, 2023. [Online]. Available: https://arxiv.org/abs/2110.00244
10. M. Havrylovych and V. Danylov, "Research of Autoencoder-Based User Biometric Verification with Motion Patterns," *System Research and Information Technologies*, no. 2, 2022, pp. 128–136. doi: 10.20535/srit.2308-8893.2022.2.10.
11. M. Havrylovych, V. Danylov, and A. Gozhyj, "Comparative Analysis of using Recurrent Autoencoders for User Biometric Verification with Wearable Accelerometer," *Proceedings of the 9th International Conference "Information Control Systems & Technologies" (ICST)*, pp. 358–370, Sept. 2020.
12. F. Trad, A. Hussein and A. Chehab, "Free Text Keystroke Dynamics-based Authentication with Continuous Learning: A Case Study," in *2022 IEEE 21st International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/*

SmartCNS), Chongqing, China, 2022, pp. 125–131. doi: 10.1109/IUCC-CIT-DSCI-SmartCNS57392.2022.00031S.

13. M. Roy, S. Majumder, A. Halder, and U. Biswas, "ECG-NET: A deep LSTM autoencoder for detecting anomalous ECG," *Engineering Applications of Artificial Intelligence*, vol. 124, p. 106484, 2023. doi: 10.1016/j.engappai.2023.106484.
14. L. Shan et al., "Abnormal ECG detection based on an adversarial autoencoder," *Frontiers in Physiology*, vol. 13, 2022. doi: 10.3389/fphys.2022.961724.
15. A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial Autoencoders," *arXiv, arXiv:1511.05644*, 2015. [Online]. Available: https://arxiv.org/abs/1511.05644.
16. A. Oluwasanmi, M.U. Aftab, E. Baagyere, Z. Qin, M. Ahmad, and M. Mazzara, "Attention Autoencoder for Generative Latent Representational Learning in Anomaly Detection," *Sensors*, vol. 22, no. 1, p. 123, Dec. 2021. doi: 10.3390/s22010123.
17. M. Hu, K. Zhang, R. You, and B. Tu, "Relative Attention-based One-Class Adversarial Autoencoder for Continuous Authentication of Smartphone Users," *arXiv, arXiv:2210.16819*, 2022. [Online]. Available: https://arxiv.org/abs/2210.16819.
18. A. Alamr and A. Artoli, "Unsupervised Transformer-Based Anomaly Detection in ECG Signals," *Algorithms*, vol. 16, no. 3, p. 152, Mar. 2023. doi: 10.3390/a16030152.
19. M. Ma, C. Sun and X. Chen, "Deep Coupling Autoencoder for Fault Diagnosis With Multimodal Sensory Data," in *IEEE Transactions on Industrial Informatics*, vol. 14, no. 3, pp. 1137–1145, March 2018. doi: 10.1109/TII.2018.2793246.
20. S. Qiu et al., "Deep Learning Techniques in Intelligent Fault Diagnosis and Prognosis for Industrial Systems: A Review," *Sensors*, vol. 23, no. 3, p. 1305, Jan. 2023. doi: 10.3390/s23031305.
21. L. Weng, "From Autoencoder to Beta-VAE," *Lil'Log*, Aug. 12, 2018. Accessed on: July 15, 2023. [Online]. Available: https://lilianweng.github.io/lil-log/2018/08/12/from-autoencoder-to-beta-vae.html
22. Qing Yang et al., "A Multimodal Data Set for Evaluating Continuous Authentication Performance in Smartphones," in *Proceedings of the 12th ACM Conference on Embedded Network Sensor Systems (SenSys '14). ACM, New York, NY, USA*, pp. 358–359. doi: 10.1145/2668332.2668366.
23. Zdeňka Sitová et al., "HMOG: New Behavioral Biometric Features for Continuous Authentication of Smartphone Users," in *IEEE Transactions on Information Forensics and Security*, vol.11, no.5, pp. 877–892, May 2016. doi: 10.1109/TIFS.2015.2506542.
24. "A Multimodal Data Set for Evaluating Continuous Authentication Performance," *H-MOG Dataset*. Accessed on: July 15, 2023. [Online]. Available: https://hmog-dataset.github.io/hmog/
25. A. Allan, *Basic Sensors in iOS: Programming the Accelerometer, Gyroscope, and More*; 1st ed., O'Reilly Media, Beijing, Aug. 30, 2011. Accessed on: July 8, 2023. [Online]. Available: https://www.amazon.com/Basic-Sensors-iOS-Programming-Accelerometer/dp/1449308465
26. S. Balli, E.A. Sağbaş, and M. Peker, "A Mobile Solution Based on Soft Computing for Fall Detection," in *Mobile Solutions and Their Usefulness in Everyday Life*, S. Paiva, Ed. Cham: Springer International Publishing, 2019, pp. 275–294. doi: 10.1007/978-3-319-93491-4_14.
27. M. El-Abed, C. Charrier, and C. Rosenberger, "Evaluation of Biometric Systems," *New Trends and Developments in Biometrics*. InTech, Nov. 28, 2012. doi: 10.5772/52084.

Received 14.07.2023

INFORMATION ON THE ARTICLE

Mariia P. Havrylovych, ORCID: 0000-0002-9797-2863, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: mariia.havrylovych@gmail.com

Valeriy Ya. Danylov, ORCID: 0000-0003-3389-3661, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: danilov1950@ukr.net

ДОСЛІДЖЕННЯ ГІБРИДНИХ АВТОКОДУВАЛЬНИКІВ З ВИКОРИСТАННЯМ ТРАНСФОРМЕРІВ ДЛЯ БІОМЕТРИЧНОЇ ВЕРИФІКАЦІЇ КОРИСТУВАЧА / М.П. Гаврилович, В.Я. Данилов

Анотація. У дослідженні розширено попередню працю з біометричної верифікації на основі руху з використанням сенсорних даних шляхом дослідження нових архітектур та більш складних даних від різних датчиків. Біометрична верифікація дає такі переваги, як унікальність для кожного користувача і захист від шахрайства. Архітектура трансформера, одна з найсучасніших у сфері штучного інтелекту, відома своїм юнітом уваги та застосуванням у різних сферах, включаючи NLP та CV. У праці досліджено її потенційну цінність для додатків, які обробляють сенсорні дані. Дослідження пропонує гібридну архітектуру, що об'єднує блоки уваги від трансформера з різними автокодувальниками, щоб оцінити її ефективність для біометричної верифікації та аутентифікації користувача. Порівняно різні конфігурації, включно з автокодувальником LSTM, автокодувальником на базі трансформера, LSTM VAE і VAE на основі трансформера. Результати показали, що поєднання блоків трансформера із неповним детермінованим автокодувальником дає найкращі метрики, але на показники моделі також значно впливають попереднє оброблення даних і параметри конфігурації алгоритму. Застосування трансформерів для біометричної верифікації та сенсорних даних виглядає багатообіцяльним, за метриками нарівні з моделями на основі LSTM або перевершуючи їх, проте з меншими часом обробленням сигналу і навчання моделі.

Ключові слова: біометрична верифікація, транформери, варіаційний автокодувальник, автокодувальник на основі трансфомера.

ПРОБЛЕМИ ПРИЙНЯТТЯ РІШЕНЬ ТА УПРАВЛІННЯ В ЕКОНОМІЧНИХ, ТЕХНІЧНИХ, ЕКОЛОГІЧНИХ І СОЦІАЛЬНИХ СИСТЕМАХ

UDC 519.925.51
DOI: 10.20535/SRIT.2308-8893.2023.3.04

INVESTIGATION OF COMPUTATIONAL INTELLIGENCE METHODS IN FORECASTING AT FINANCIAL MARKETS

Yu. ZAYCHENKO, He. ZAICHENKO, O. KUZMENKO

Abstract. The work considers intelligent methods for solving the problem of short- and middle-term forecasting in the financial sphere. LSTM DL networks, GMDH, and hybrid GMDH-neo-fuzzy networks were studied. Neo-fuzzy neurons were chosen as nodes of the hybrid network, which allows to reduce computational costs. The optimal network parameters were found. The synthesis of the optimal structure of hybrid networks was performed. Experimental studies of LSTM, GMDH, and hybrid GMDH-neo-fuzzy networks with optimal parameters for short- and middle-term forecasting have been conducted. The accuracy of the obtained experimental predictions is compared. The forecasting intervals for which the application of the researched artificial intelligence methods is the most expedient have been determined.

Keywords: optimization, GMDH, hybrid GMDH-neo-fuzzy network, LSTM, short- and middle-term forecasting.

INTRODUCTION

Problems of forecasting share prices and market indexes at stock exchanges pay great attention of investors and various money funds. For its solution were developed and for a long time applied powerful statistical methods, first of all ARIMA [1; 2]. Last years different intelligent methods and technologies were also suggested and widely used for forecasting in financial sphere, in particular among them neural networks and fuzzy logic systems.

The efficient tool of modelling and forecasting of non-stationary time series is Group method of data Handling (GMDH) suggested and developed by acad. Alexey Ivakhnenko [3; 4]. This method is based on self-organization and enables to construct optimal structure of forecasting model automatically in the process of algorithm run. Methods GMDH and fuzzy GMDH were successfully applied for forecasting at stock exchanges for long time.

As alternative approach for forecasting in finance is application of various types of neural network: MLP [5], fuzzy neural networks [6; 7], neo-fuzzy networks [8] and Deep learning (DL) networks [9].

New trend in sphere DL networks is a new class of neural networks – hybrid DL networks based on GMDH method [10]. The application of self-organization in these networks enables to train not only neuron weights but to construct optimal structure of a network. Due to a method of training in these networks weights are adjusted not simultaneously but layer after layer. That prevents the phenomenon of vanishing or explosion of gradient. It's very important for networks with many layers.

The first works in this field used as nodes of the hybrid network Wang-Mendel neurons with two inputs [10]. But drawback of such neurons is the necessity to train not only neural weights but the parameters of fuzzy sets in antecedents of rules as well. That needs a lot of calculation expenses and large training time as well. Therefore, later DL neo-fuzzy networks were developed in which as nodes were used neo-fuzzy neurons by Yamakawa [8; 11; 12]. The main property of such neurons is that it's necessary to train only neuron weights but not fuzzy sets. That demands less computation in comparison to Wang-Mendel neurons and significantly cuts training time as a whole. The investigation of both classes of hybrid DL networks was performed and their efficiency at forecasting in financial sphere was compared in [13].

At the same time for long term forecasting LSTM networks were developed [14–16] and successfully applied for forecasting in economy and financial sphere. LSTM networks have long memory where the information about preceding values of forecasted time series is stored and they are enabled to forecast at middle term and long term forecasting intervals. Therefore, it presents great interest to compare the efficiency of hybrid DL networks, GMDH and LSTM at the problems of short-term and middle-term forecasting at financial sphere.

The goal of this paper is to investigate the accuracy of intelligent methods – hybrid DL networks, GMDH and LSTM at the problem of forecasting market indices at the stock exchange at the different forecasting intervals (short-term and middle-term), compare their efficiency and to determine the classes of forecasting problems for which the application of corresponding computational intelligence methods is the most perspective.

THE DESCRIPTION OF THE EVOLVING HYBRID GMDH-NEO-FUZZY NETWORK

The evolving hybrid DL-network architecture is presented in Fig. 1. To the system's input layer a $n \times 1$-dimensional vector of input signals is fed. After that

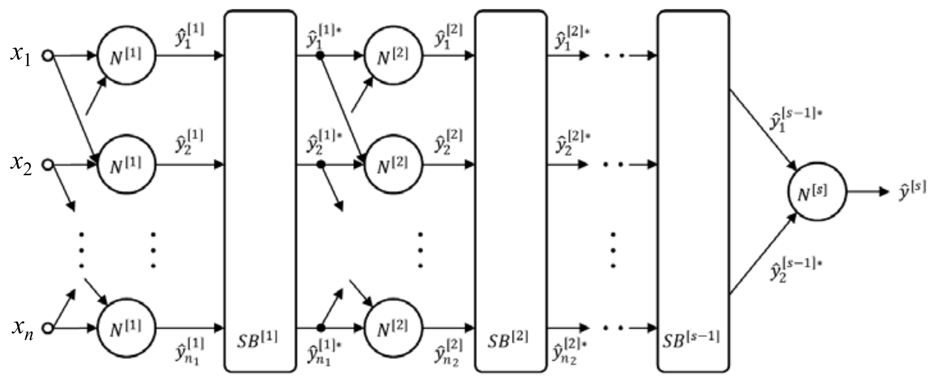


Fig. 1. Evolving GMDH-network

this signal is transferred to the first hidden layer. This layer contains $n_1 = c_n^2$ nodes, and each of these neurons has only two inputs.

At the outputs $N^{[1]}$ of the first hidden layer the output signals are formed. Then these signals are fed to the selection block of the first hidden layer.

It selects among the output signals $\hat{y}_l^{[1]} n_1 *$ (where $n_1 * = F$ is so-called freedom of choice) most precise signals by some chosen criterion (mostly by the mean squared error $\sigma^2_{y_l^{[1]}}$). Among these $n_1 *$ best outputs of the first hidden layer $\hat{y}_l^{[1]} * n_2$ pairwise combinations $\hat{y}_l^{[1]}*, \hat{y}_p^{[1]} *$ are formed. These signals are fed to the second hidden layer, that is formed by neurons $N^{[2]}$. After training these neurons output signals of this layer $\hat{y}_l^{[2]}$ are transferred to the selection block $SB^{[2]}$ which choses F best neurons by accuracy (e.g. by the value of $\sigma^2_{y_l^{[2]}}$) if the best signal of the second layer is better than the best signal of the first hidden layer $\hat{y}_1^{[1]} *$. Other hidden layers work similarly. The system evolution process continues until the best signal of the selection block $SB^{[s+1]}$ appears to be worse than the best signal of the previous *s-h* layer. Then it's necessary to return to the previous layer and choose its best node neuron $N^{[s]}$ with output signal $\hat{y}^{[s]}$. And moving from this neuron (node) along its connections backwards and sequentially passing all previous layers the final structure of the GMDH-neo-fuzzy network is constructed.

It should be noted that in such a way not only the optimal structure of the network may be constructed but also well-trained network due to the GMDH algorithm. Besides, since the training is performed sequentially layer by layer the problems of high dimensionality as well as vanishing or exploding gradient are avoided.

NEO-FUZZY NEURON AS A NODE OF HYBRID GMDH-SYSTEM

Let's consider the architecture of the node that is presented in Fig. 2 and is suggested as a neuron of the proposed GMDH-system. As a node of this structure a neo-fuzzy neuron (NFN) developed by Takeshi Yamakawa and co-authors in [9] is used. The neo-fuzzy neuron is a nonlinear multi-input single-output system shown in Fig. 2. The main difference of this node from the general neo-fuzzy neuron structure is that each node uses only two inputs.

It realizes the following mapping:

$$\hat{y} = \sum_{i=1}^{2} f_i(x_i),$$

where x_i is the input i $(i = 1, 2, \ldots, n)$, $\hat{y}$ is a system output. Structural blocks of neo- fuzzy neuron are nonlinear synapses NS_i which perform transformation of input signal in the form

$$f_i(x_i) = \sum_{j=1}^{h} w_{ji}\mu_{ji}(x_i)$$

and realize fuzzy inference: *if* x_i *is* x_{ji} *then the output is* w_{ji} ,where x_{ji} is a fuzzy set which membership function is μ_{ji} , w_{ji} is a synaptic weight in consequent [11].

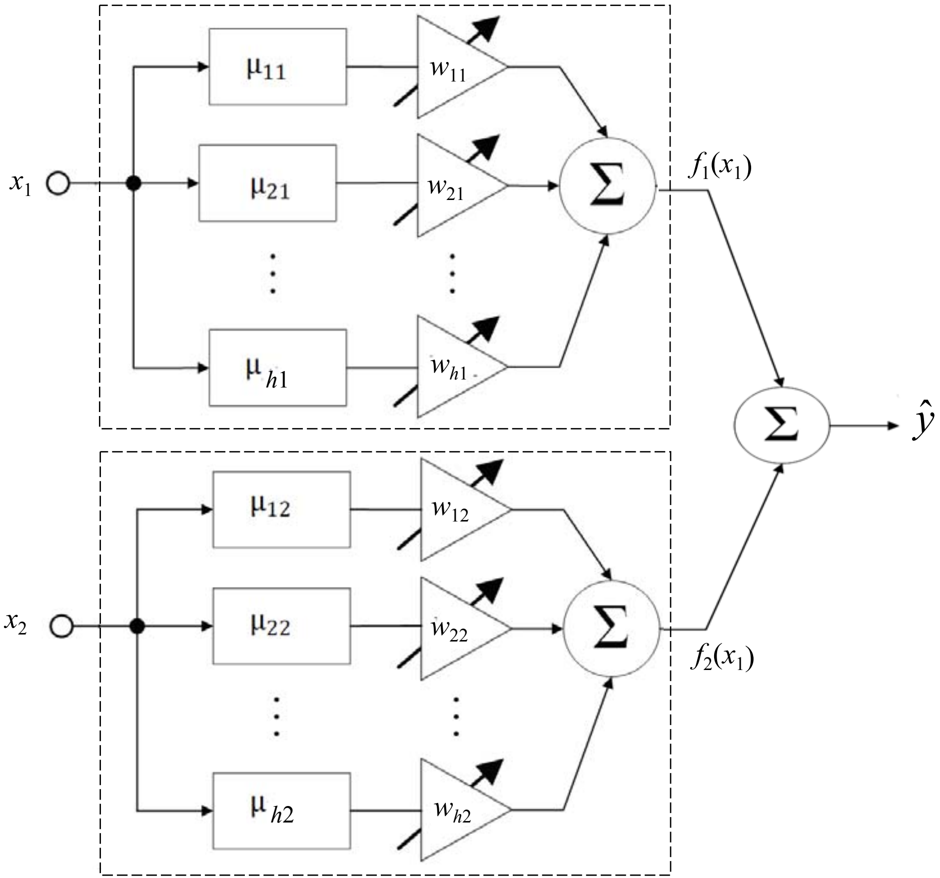


Fig. 2. Architecture of neo-fuzzy neuron with two inputs

THE NEO-FUZZY NEURON LEARNING ALGORITHM

The learning criterion (goal function) is the standard local quadratic error function:

$$E(k) = \frac{1}{2}(y(k) - \hat{y}(k))^2 = \frac{1}{2}e(k)^2 = \frac{1}{2}\left(y(k) - \sum_{i=1}^{2}\sum_{j=1}^{h} w_{ji}\mu_{ji}(x_i(k)) \right)^2 .$$

It is minimized via the conventional stochastic gradient descent algorithm.

In case we have a priori defined data set the training process can be performed in a batch mode at one epoch using conventional least squares method [12]

$$w^{[1]}(N) = \left(\sum_{k=1}^{N} \mu^{[1]}(k)\mu^{[1]T}(k) \right)^{+} \sum_{k=1}^{N} \mu^{[1]}(k)y(k) = P^{[1]}(N)\sum_{k=1}^{N} \mu^{[1]}(k)y(k) ,$$

where $(\bullet)^+$ means pseudo inverse of Moore–Penrose (here $y(k)$ denotes external reference signal (real value).

If training observations are fed sequentially in on-line mode, the recurrent form of the LSM can be used in the form:

$$\begin{cases} w_l^{ij}(k) = w_l^{ij}(k-1) + \dfrac{P^{ij}(k-1)(y(k) - (w_l^{ij}(k-1))^{\mathrm{T}} \varphi^{ij}(x(k)))\varphi^{ij}(x(k))}{1 + (\varphi^{ij}(x(k)))^{\mathrm{T}} P^{ij}(k-1)\varphi^{ij}(x(k))}, \\ P^{ij}(k) = P^{ij}(k-1) - \dfrac{P^{ij}(k-1)\varphi^{ij}(x(k))(\varphi^{ij}(x(k)))^{\mathrm{T}} P^{ij}(k-1)}{1 + (\varphi^{ij}(x(k)))^{\mathrm{T}} P^{ij}(k-1)\varphi^{ij}(x(k))}. \end{cases}$$

DATASET

As the data set for forecasting were taken close values of market index NASDAQ Composite in the period since 01.01.22 till 01.01.23. The whole sample consisted of 251 instances included Open values, minimal, maximal and Close values and volume in each day. The sample was divided into training and test subsamples. The dynamics of NASDAQ Close values is shown in the Fig. 3.

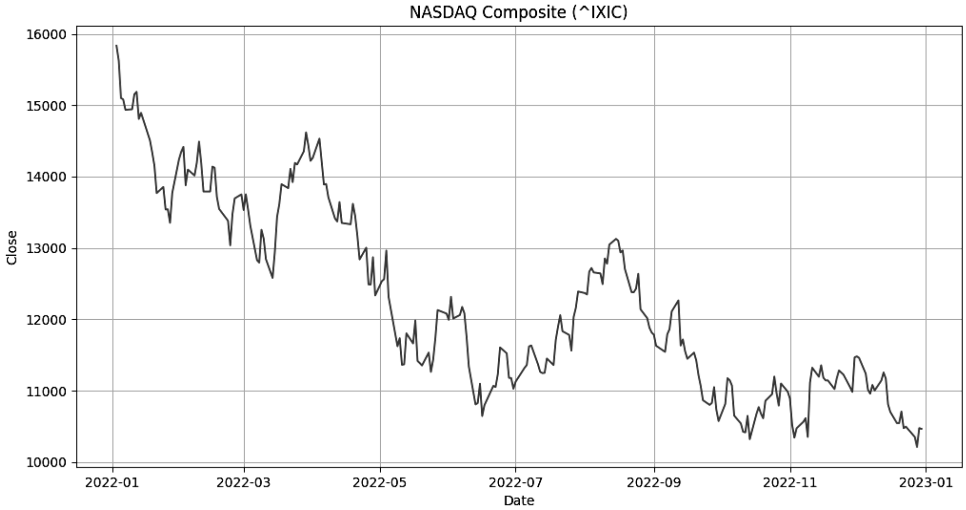


Fig. 3. Dynamics of the index Close

The correlogram of NASDAQ index is presented in the Fig. 4.

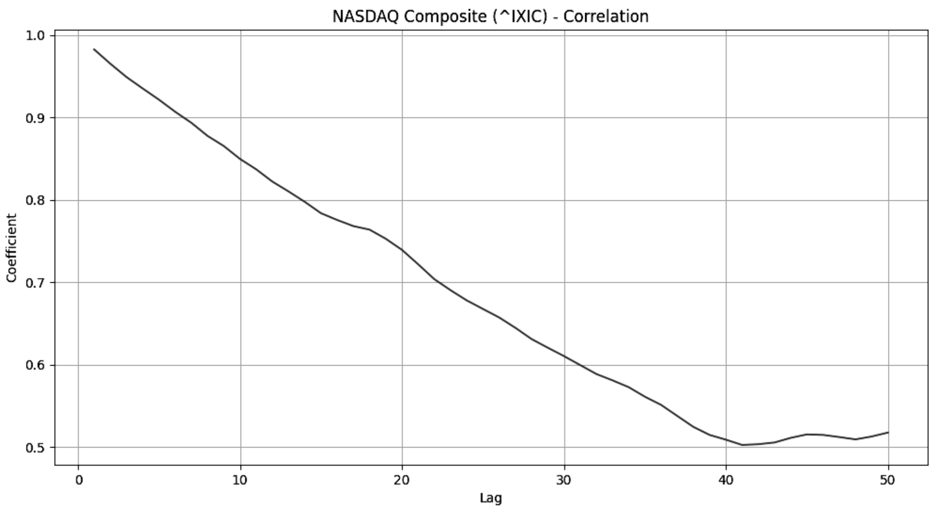


Fig. 4. Correlogram

Analyzing the presented curve, one may conclude that there is strong correlation between preceding and conceding values and even for lag 50 days the correlation is more than 0.5.

EXPERIMENTAL INVESTIGATIONS

In the investigations was explored the forecasting accuracy of hybrid DL neo-fuzzy networks at various forecasting intervals: short-term forecasting with intervals 1, 3, 5 and 7 days and middle-term forecasting with intervals 20 and 30 days. At the first step the variable experimental parameters of hybrid network were chosen which are presented in the Table 1.

Table 1. Experimental parameters

Parameter	Value
Membership functions	Gaussian
Number of inputs	3; 4; 5
Number of linguistic variables	3; 4; 5
Ratio (percentage of the training sample)	0.6 (60%); 0.7 (70%); 0.8 (80%)
Criterion	MSE; MAPE
Forecast interval	1; 3; 5; 7; 20; 30

The optimization of these parameters was performed in result the following optimal values were determined inputs: 3; linguistic variables: 3; ratio: 0.7.

After that the structure optimization of hybrid DL neo-fuzzy network was performed using GMDH method. The process of structure generation is presented in the Table 2.

Table 2. Structure generation (inputs: 3; variables: 3; ratio: 0.7)

Nodes	SB1	SB2	SB3
(0, 1)	2.6152319		
(0, 2)	5.6112545		
(1, 2)	3.8828252		
((0, 1), (0, 2))		0.03519317	
((0, 1), (1, 2))		0.0357832	
((0, 2), (1, 2))		0.05844182	
(((0, 1), (0, 2)), ((0, 1), (1, 2)))			0.09281185
(((0, 1), (0, 2)), ((0, 2), (1, 2)))			0.11276198
(((0, 1), (1, 2)), ((0, 2), (1, 2)))			0.08893768

In result the optimal structure of three layers: at the first layer 3 inputs, second layer – two neurons, third layer – one output neuron.

Further the training of the best hybrid network was carried out using method SGD (stochastic gradient descent) with variable step. Flow chart of forecasting results for interval 20 in presented in the Fig. 5. The values of MSE and MAPE for this experiment are shown in the Table 3.

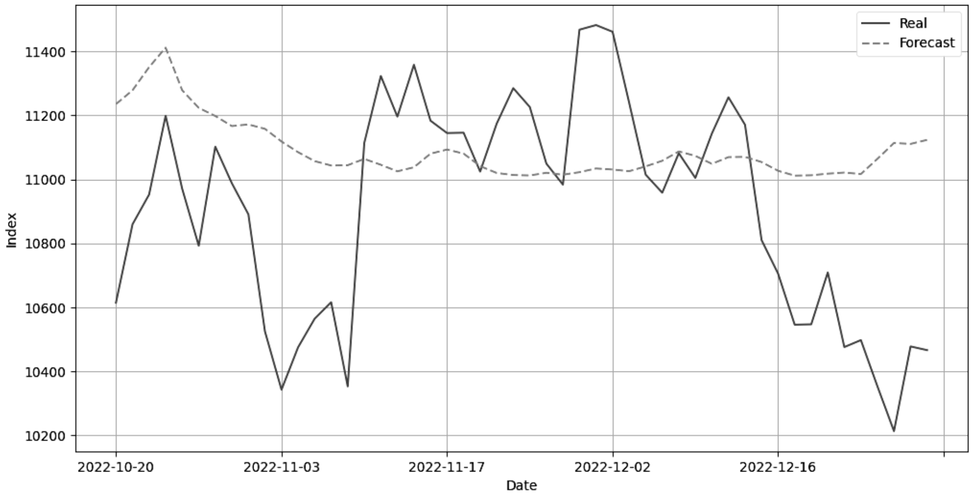


Fig. 5. The best forecast (inputs: 3; variables: 3; ratio: 0.7)

Table 3. Forecasting accuracy of hybrid neo-fuzzy network at forecasting interval 20 days

Criterion	MSE	MAPE
min	30.68518	0.049986
average	158515.7	3.024738
maximal	811272.4	8.818966

In the Fig. 6. flow chart of MAPE values for the best model of hybrid network is shown.

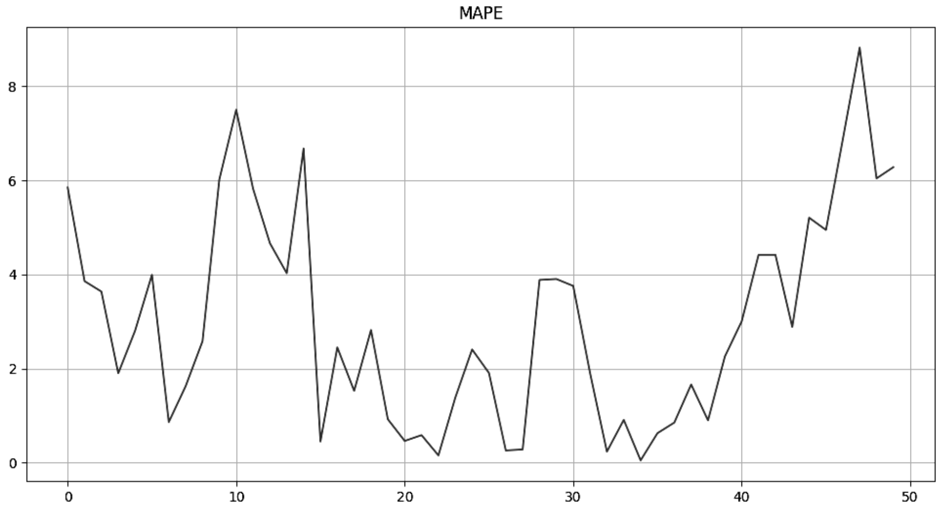


Fig. 6. MAPE for the best forecast (inputs: 3; variables: 3; ratio: 0.7)

Further the similar experiments of hybrid network were performed with forecasting interval 30 days. After optimization the parameters and structure of hybrid network it was trained using training subsample. The forecasting accuracy y at the test sample is presented at the Table 4.

Table 4. Forecasting accuracy of hybrid neo-fuzzy network at interval 30 days

Criterion	MSE	MAPE
min	177.865	0.120699
average	164611	3.07087
maximal	840641.8	8.977178

For estimating forecasting accuracy of hybrid DL network, it was compared with alternative methods: GMDH and LSTM. For GMDH algorithm the following parameters values were set after preliminary explorations: linear partial descriptions, number of inputs 5, ratio training/test 0.6. Flow chart of the best forecast is shown in the Fig. 7 and Fig. 8.

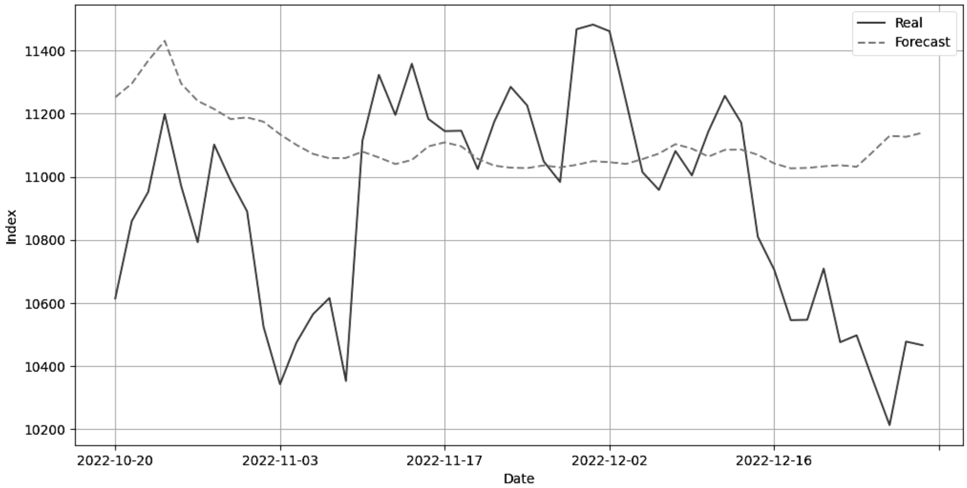


Fig. 7. The best forecast (inputs: 3; variables: 3; ratio: 0.8) for interval 30 days

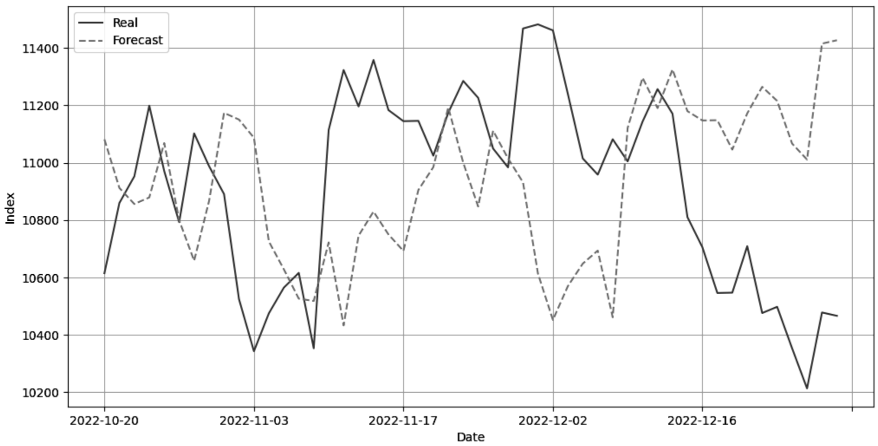


Fig. 8. The best forecast by GMDH (function: linear; inputs: 5; ratio: 0.6) 20 days

After that the experiments were performed with LSTM network. LSTM was trained and tested at the different forecasting intervals 1, 3, 5, 7, 20 and 30 days. The goal of experiments was to find the optimal parameters. The following parameters varied: number of inputs 3-5, ratio training/test 0.6, 0.7, 0.8. After that the LSTM with optimal parameters was applied for forecasting.

In the Table 5 forecasting accuracy of LSTM network at interval 3 days and in the Fig. 9 forecasting results are presented. The optimal parameters values were found number of inputs 5, ratio training/test 0.6.

Table 5. Forecasting accuracy of LSTM network at forecasting interval 3 days

Criterion	MSE	MAPE
min	113.4100292	0.098063438
average	117981.36	2.652192244
maximal	517650.7403	6.953914724

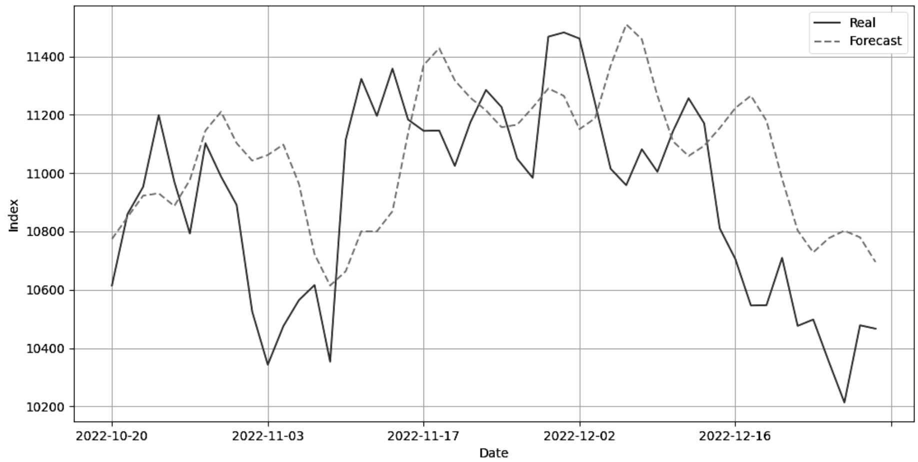


Fig. 9. The best forecast by LSTM (inputs: 5; ratio: 0.6) 3 days

The values of MSE and MAPE for forecasting with an interval of 20 days are shown in Table 6. The forecasting results are presented in Fig. 10.

Table 6. Forecasting accuracy of LSTM network at forecasting interval 20 days

Criterion		MSE	MAPE
min		49.56215352	0.06300144
average		327754.696	4.11679646
maximal	1545745.838	12.17316133	

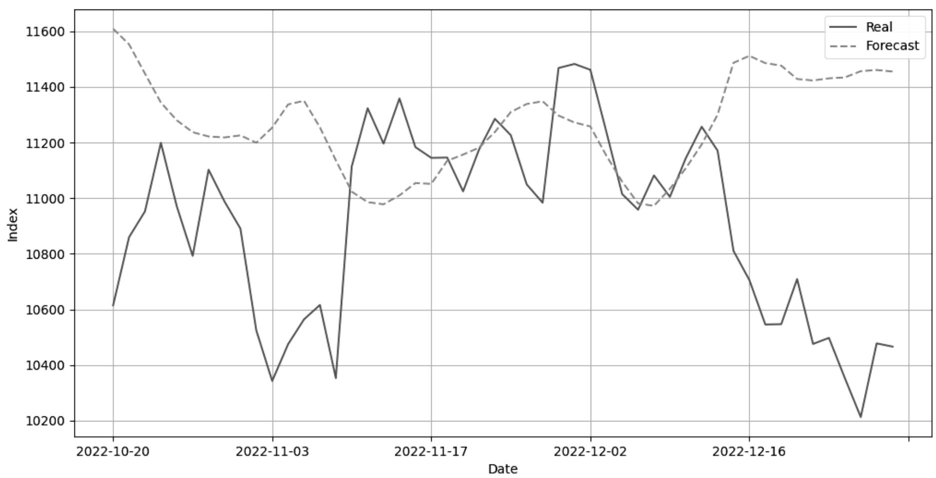


Fig. 10. The best forecast by LSTM (inputs: 5; ratio: 0.6) 20 days

The comparative experiments were performed in which the accuracy of forecasting by hybrid DL network, GMDH and LSTM at the different forecasting intervals was estimated and compared. The corresponding results are presented in the Tables 7, 8 and Fig. 11, 12.

Table 7. Average MSE values of the best models for different intervals

Interval	GMDH-neo-fuzzy	GMDH	LSTM
interval 1	97865.41363	44462.69	55461.3459
interval 3	104012.245	122615	117981.36
interval 5	155308.7139	151131.5	220850.108
interval 7	156023.0308	191982.4	241535.576
interval 20	158515.6721	243991.7	327754.7
interval 30	164610.9742	245615.6	327216.9

Table 8. Average MAPE values of the best models for different intervals

Interval	GMDH-neo-fuzzy	GMDH	LSTM
interval 1	2.483877618	1.557535	1.76242389
interval 3	2.544556353	2.623422	2.65219224
interval 5	2.889892779	3.035898	3.56067021
interval 7	2.867433998	3.428108	3.73361624
interval 20	3.02473808	3.710976	4.116796
interval 30	3.070870375	3.870127	4.25219

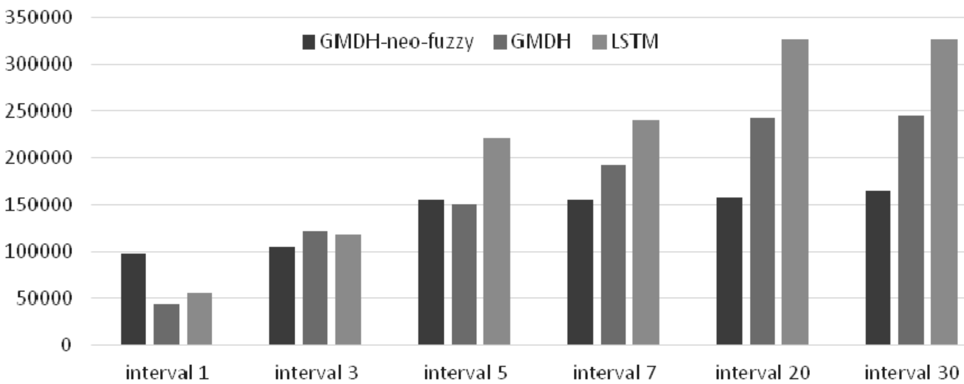


Fig. 11. Average MSE values of the best models for different intervals

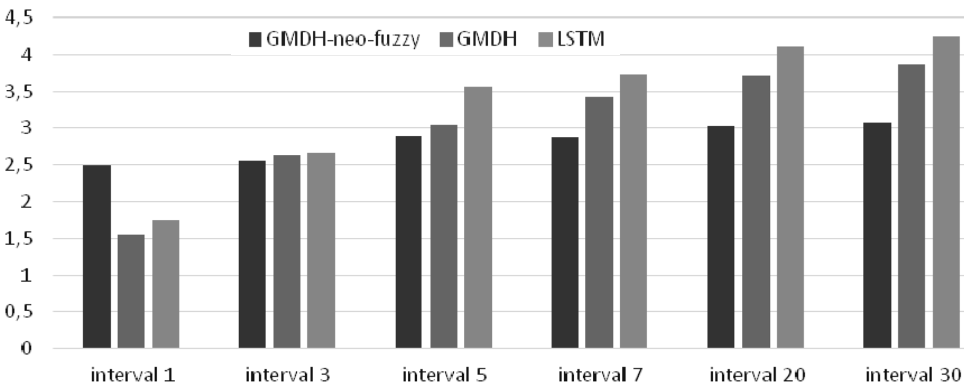


Fig. 12. Average MAPE values of the best models for different intervals

Analyzing the presented results in the Fig. 11 one may conclude that GMDH method appears to be the best at short term forecasting 1, 3 days which complies the theory.

Hybrid deep learning neo-fuzzy networks are the best at middle-term forecasting 7, 20, 30 days. LSTM networks appeared to be the worst by accuracy as compared with intelligent methods – hybrid DL networks and GMDH.

CONCLUSION

In this paper the investigations of artificial intelligence methods: hybrid Deep learning networks and GMDH were carried out in the problem of forecasting NASDAQ close prices.

During the experiments the optimal structure and optimal parameters: number of inputs, number of linguistic values, ratio training/test samples of hybrid neo-fuzzy networks were determined.

After optimization of hybrid neo-fuzzy networks and parameters of GMDH method the experiments on forecasting NASDAQ Close were performed at different intervals: 1, 3, 5, 7 (short-term forecast) and 20, 30 days (middle-term forecast).

The accuracy of forecasting by Hybrid DL networks and GMDH was compared with alternative method – LSTM networks.

The analysis of obtained results have shown that GMDH method is the best at short term forecasting 1, 3 days while hybrid deep learning neo-fuzzy networks are the best at middle-term forecasting 7, 20, 30 days. LSTM networks appeared to be the worst by accuracy as compared with intelligent methods – hybrid DL networks and GMDH.

REFERENCES

1. Peter J. Brockwell and Richard A. Davis, *Introduction to time series and forecasting*; 2nd ed. Springer, 2002, 429 p.
2. Robert H. Shumway and David S. Stoffer, *Time Series Analysis and its Applications with R Examples*; 4-th edition. Springer, 2017, 562 p.
3. A.G. Ivakhnenko, G.A. Ivakhnenko, and J.A. Mueller, “Self-organization of the neural networks with active neurons,” *Pattern Recognition and Image Analysis*, vol. 4, no. 2, pp. 177–188, 1994.
4. A.G. Ivakhnenko, D. Wuensch, and G.A. Ivakhnenko, “Inductive sorting-out GMDH algorithms with polynomial complexity for active neurons of neural networks,” *Neural Networks*, 2, pp. 1169–1173, 1999.
5. S.S. Haykin, *Neural networks: a comprehensive foundation*; 2nd ed. Upper Saddle River, N.J: Prentice Hall, 1999.
6. S. Ossovsky, *Neural networks for information processing*. M.: Finance and Statistics, 2002, 344 p.
7. F. Wang, “Neural Networks Genetic Algorithms and Fuzzy Logic for Forecasting,” *Proc. Intern. Conf. Advanced Trading Technologies*. New York, 1992, pp. 504–532.
8. T. Yamakawa, E. Uchino, T. Miki, and H. Kusanagi, “A neo-fuzzy neuron and its applications to system identification and prediction of the system behavior,” *Proc. 2nd Intern. Conf. Fuzzy Logic and Neural Networks «LIZUKA-92», Lizuka, 1992*, pp. 477–483.
9. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT PRESS, 2016. Available: http://www.deeplearningbook.org

10. Yuriy Zaychenko, Yevgeniy Bodyanskiy, Oleksii Tyshchenko, Olena Boiko, and Galib Hamidov, "Hybrid GMDH-neuro-fuzzy system and its training scheme," *Int. Journal Information theories and Applications*, vol. 24, no. 2, pp. 156–172, 2018.
11. Yu. Zaychenko and Galib Hamidov, "The Hybrid Deep Learning GMDH-neo-fuzzy Neural Network and Its Applications," *Proceedings of 13-th IEEE Intern Conference Application of Information and Communication Technologies-AICT2019*, *23–25 October 2019, Baku*, pp. 72–77.
12. Evgeniy Bodyanskiy, Yuriy Zaychenko, Olena Boiko, Galib Hamidov, and Anna Zelikman, "Structure Optimization and Investigations of Hybrid GMDH-Neo-fuzzy Neural Networks in Forecasting Problems," *System Analysis & Intelligent Computing; eds. Michael Zgurovsky, Natalia Pankratova (Book Studies in Computational Intelligence, SCI)*, vol. 1022. Springer, 2022, pp. 209–228.
13. Yuriy Zaychenkoa, Helen Zaichenkoa, and Galib Hamidov, "Hybrid GMDH Deep Learning Networks – Analysis, Optimization and Applications in Forecasting at Financial Sphere," *System Research and Information Technologies*, no. 1, pp. 73–86, 2022. doi: 10.20535/SRIT.2308-8893.2022.1.06.
14. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation,* vol. 9, pp. 1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735.
15. B. Hammer, "On the approximation capability of recurrent neural networks," *Neurocomputing*, vol. 31, pp. 107–123, 1998. doi: 10.1016/S0925-2312(99)00174-5.
16. C. Olah, *Understanding LSTM networks*, 2020. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/
17. A. Graves, "Generating sequences with recurrent neural networks," *CoRR*, vol. abs/1308.0850, 2013. doi: 10.48550/arXiv.1308.0850.

Received 08.05.2023

INFORMATION ON THE ARTICLE

Yuriy P. Zaychenko, ORCID: 0000-0001-9662-3269, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: zaychenkoyuri@ukr.net

Helen Yu. Zaichenko, ORCID: 0000-0002-4630-5155, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: syncmaster@bigmir.net

Oleksii V. Kuzmenko, ORCID: 0000-0003-1581-6224, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: oleksii.kuzmenko@ukr.net

ДОСЛІДЖЕННЯ МЕТОДІВ ОБЧИСЛЮВАЛЬНОГО ІНТЕЛЕКТУ У ПРОГНОЗУВАННІ НА ФІНАНСОВИХ РИНКАХ / Ю.П. Зайченко, О.Ю. Зайченко, О.В. Кузьменко

Анотація. Розглянуто інтелектуальні методи для короткострокового та середньострокового прогнозування у фінансовій сфері. Досліджувалися DL мережі LSTM, МГУА та гібридні МГУА неофаззі мережі. Як вузли гібридної мережі обрано неофаззі нейрони, що дозволяє зменшити обчислювальні витрати. Знайдено оптимальні параметри мереж. Виконано синтез оптимальної структури гібридних мереж. Проведено експериментальні дослідження мереж LSTM, МГУА та МГУА неофаззі з оптимальними параметрами для короткострокового та середньострокового прогнозування. Порівняно точність отриманих експериментальних прогнозів. Визначено інтервали прогнозування, для яких застосування досліджених методів штучного інтелекту є найбільш доцільним.

Ключові слова: оптимізація, МГУА, гібридна мережа МГУА-неофаззі, LSTM, короткострокове та середньострокове прогнозування.

UDC 004.942:519.216.3
DOI: 10.20535/SRIT.2308-8893.2023.3.05

UNCERTAINTIES IN DATA PROCESSING, FORECASTING AND DECISION MAKING

L.B. LEVENCHUK, O.L. TYMOSHCHUK, V.H. GUSKOVA, P.I. BIDYUK

Abstract. Forecasting, dynamic planning, and current statistical data processing are defined as the process of estimating an enterprise's current state on the market compared to other competing enterprises and determining further goals as well as sequences of actions and resources necessary for reaching the goals stated. In order to perform high-quality forecasting, it is proposed to identify and consider possible uncertainties associated with data and expert estimates. This is one of the system analysis principles to be hired for achieving high-quality final results. A review of some uncertainties is given, and an illustrative example showing improvement of the final result after considering possible stochastic uncertainty is provided.

Keywords: mathematical model, statistical data uncertainties, system analysis principles, forecasting, decision support system.

INTRODUCTION

Analysis of dynamic processes in forecasting and planning procedures is an urgent problem not only for financial organizations and companies but for all industrial enterprises, small and medium business, investment and insurance companies etc. Forecasting, dynamic planning (DP) and current data processing could be defined as the process of estimation by an enterprise of its current state on the market in comparison with other competing enterprises, and determining further goals as well as sequences of actions and resources that are necessary for reaching the goals stated. The process of forecasting and planning is performed continuously (or quasi-continuously) with acquiring new information (knowledge) about market, technologies, forecast estimates of necessary variables, current and future situations. All this knowledge is used for correcting actions and activities of an enterprise and supporting its competitiveness with flow of time.

Formally DP could be presented in the form:

$$DSP = \{\mathbf{X}_0, \mathbf{G}, \mathbf{R}, \mathbf{D}(t), \mathbf{K}, \mathbf{T}, \mathbf{F}, \Delta\mathbf{D}(t), \Delta\mathbf{R}(t)\},$$

where $\mathbf{X}_0$ is initial state of an enterprise; $\mathbf{G}$ are the goals stated by the enterprise management; $\mathbf{R}$ are resources that are necessary for reaching the goals stated. $\mathbf{D}(t)$ is a sequence of actions that should be performed on the interval of planning; $\mathbf{K}$ is a new knowledge about environment; $\mathbf{T}$ are new technologies. Symbol $\mathbf{F}$ designates possible results of forecasting and foresight; $\Delta\mathbf{D}(t)$ are corrections that are to be performed for reaching the goals; $\Delta\mathbf{R}(t)$ are necessary extra resources. One of the main problems that are to be solved within the DP paradigm is high quality forecasting of relevant processes.

Adequate models of the process and the forecasts generated with them are helpful for taking into consideration a set of various influencing factors and make based on objective planning managerial decisions. Another purpose of the studies is in estimating possible risks using forecasts of volatility. There are several types of processes that could be described with mathematical models in the form of appropriately constructed equations or probability distributions. Among them are the processes with deterministic and stochastic trends, and heteroscedastic processes. As of today the following mathematical models are widely used for describing nonlinear dynamics of processes relevant to planning: linear and nonlinear regression (logit and probit, polynomials, splines), autoregressive integrated moving average (ARIMA) models, autoregressive conditionally heteroscedastic models (ARCH), generalized ARCH (GARCH), dynamic Bayesian networks, support vector machine (SVM) approach, neural networks and neuro-fuzzy techniques as well as combinations of the approaches mentioned [1–5].

All types of mathematical modeling usually need to cope with various kinds of uncertainties associated with statistical/experimental data, structure of the process under study and its model, parameter uncertainty, and uncertainties relevant to the quality of models and forecasts. Reasoning and decision making are very often performed with leaving many facts unknown or rather vaguely represented in processing of data and expert estimates. To avoid or to take into consideration the uncertainties and improve this way quality of the final result (estimates of processes forecasts and planning of decisions based upon them) it is necessary to construct appropriate computer based decision support systems (DSS) for solving multiple specific problems.

Selection and application of a specific model for process description and forecasts estimation depends on application area, availability of statistical/experimental data, qualification of personnel, who work on the data analysis problems, and availability of appropriate applied software. Better results for estimation of processes forecasts are usually achieved with application of ideologically different techniques combined in the frames of one specialized computer system. Such approach to solving the problems of quality forecasts estimation can be implemented in the frames of modern decision support systems. DSS today (especially intellectual DSS) create a powerful instrument for supporting user's (managerial) decision making as far as it combines a set of appropriately selected data and expert estimates processing procedures aiming to reach final result of high quality: objective high quality alternatives for a decision making person (DMP). Development of a DSS is based on modern theory and techniques of system analysis principles, data processing systems, estimation and optimization theories, mathematical and statistical modeling and forecasting, decision making theory as well as many other results of theory and practice of processing data and expert estimates [6–8].

The paper considers the problem of adequate models constructing for solving the problems of modeling and estimating forecasts for selected types of dynamic processes with the possibility for application of alternative data processing techniques, modeling and estimation of parameters and states for the processes under study in conditions of availability possible uncertainties.

PROBLEM FORMULATION

The purpose of the study is as follows: 1) analysis of uncertainty types characteristic for model building and forecasting dynamic processes; 2) selection of techniques for taking into consideration the uncertainties detected; 3) selection of mathematical modeling and forecasting techniques for nonstationary and nonlinear heteroscedastic processes; 4) illustration of the methodology application to solving selected problem of forecasts estimation using appropriate statistical data.

COPING WITH UNCERTAINTIES

All types of mathematical modeling with the use of statistical/experimental data usually need to consider various kinds of uncertainties associated with data, informational structure of a process under study and its model, parameter estimate uncertainty, and uncertainties relevant to the quality of models and forecasts. In many cases a researcher has to cope with the following basic types of uncertainties: structural, statistical and parametric. Structural uncertainties are encountered in the cases when structure of the process under study (and respectively its model) is unknown or not clearly enough defined, in other words known partially only. For example, when the functional approach to model constructing is applied usually we do not know details of an object (or a process) structure and it is estimated with appropriate model structure estimation techniques: correlation analysis, estimation of mutual information, lags, testing for nonlinearity and nonstationarity, identification of external disturbances etc. Uncertainty could also be introduced by an expert who is studying the process and provides its estimates for model structure, parameter restrictions, selection of computational procedures etc. The sequence of actions necessary for identification, processing and taking into consideration of uncertainties could be formulated as follows: – identification and reduction of data uncertainty; – model structure and parameters estimation; – reduction of uncertainties related to the model structure and parameters estimation; – reduction of uncertainties relevant to expert estimates; – estimation of forecasts and reduction of respective uncertainties; – selection of the best final result using appropriate set of quality statistics. All the tasks mentioned above are usually solved sequentially (in an adaptive loop) with appropriately designed and implemented DSS.

Here we consider uncertainties as the factors that influence negatively the whole process of mathematical model constructing, forecasts and possible risk estimating and generating of alternative decisions. These factors lead to lower quality of intermediate and final results of computations performed within selected or designed system. They are inherent to the process being studied due to incomplete or noise corrupted data, complex stochastic external influences, incompleteness or inexactness of our knowledge regarding the objects (systems) structure, incorrect application of computational procedures etc. The uncertainties very often appear due to incompleteness of data, noisy measurements or they are invoked by sophisticated stochastic external disturbances with complex unknown probability distributions, poor estimates of model structure or by a wrong selection of parameter estimation procedure. The problem of uncertainty identification is solved with application of special statistical tests, visual studying of available data, using appropriate expert estimates.

As far as we usually work with stochastic data, correct application of existing statistical techniques provides a possibility for approximate estimation of a system (and its model) structure. To find "the best" model structure it is recommended to apply adaptive estimation schemes that provide automatic search in a pre-defined range of possible model structures and parameters (model order, time lags, and possible nonlinearities). It is often possible to perform the search in the class of regression type models with the use of information criterion of the following type [2]:

$$N\ \log(FPE) = N\log(V_N(\hat{\theta})) + N\log\left(\frac{N+p}{N-p}\right), \qquad (1)$$

where $\hat{\theta}$ is a vector of model parameters estimates; N is a power of time series used; *FPE* is final prediction error term; $V_N(\hat{\theta})$ can be determined by the sum of squared errors; p is a number of model parameters. The value of the criteria (1) is asymptotically equivalent to the Akaike information criterion with $N \to \infty$. As the amount of data *N* may be limited, then an alternative, the minimum description length (MDL) criterion

$$MDL = \log(V_N(\hat{\theta})) + p\frac{\log(N)}{N}$$

could be hired to find the model that adequately represents available data with the minimum amount of available information.

There are several possibilities for adaptive model structure estimation: 1) application of statistical criteria for detecting possible nonlinearities and the type of nonstationarity (integrated or heteroskedastic process); 2) analysis of partial autocorrelation for determining autoregression order; 3) automatic estimation of the exogenous variable lag (detection of leading indicators); 4) automatic analysis of residual properties; 5) analysis of data distribution type and its use for selecting correct model estimation method; 6) adaptive model parameter estimation with hiring extra data; 7) optimal selection of weighting coefficients for exponential smoothing, nearest neighbor and other techniques. The development and use of a specific adaptation scheme depends on the volume and quality of data, specific problem statement, requirements to forecast estimates etc.

The adaptive estimation schemes also help to cope with the model parameter uncertainties. New data are used to re-compute model parameter estimates that correspond to possible changes in the object under study. In the cases when model is nonlinear, alternative parameter estimation techniques (say, MCMC) could be hired to compute alternative (though admissible) sets of parameters and to select the most suitable of them using statistical quality criteria.

Processing some types of possible stochastic uncertainties. While performing practical modeling very often statistical characteristics (covariance matrix) of stochastic external disturbances and measurement noise (errors) are unknown. To eliminate this uncertainty optimal filtering algorithms are usually applied that provide for a possibility of simultaneous estimation of object (system) states and the covariance matrices. One of the possibilities to solve the problem is application of optimal Kalman filter. Kalman filter is used to find optimal estimates of system states on the bases of a system model represented in a widely used convenient state space form as follows:

$$\mathbf{x}(k)=\mathbf{A}(k,k-1)\mathbf{x}(k-1)+\mathbf{B}(k,k-1)\mathbf{u}(k-1)+\mathbf{w}(k)\,, \qquad (2)$$

where $\mathbf{x}(k)$ is n-dimensional vector of system states; $k=0,1,2,...$ is discrete time; $\mathbf{u}(k-1)$ is m- dimensional vector of deterministic control variables; $\mathbf{w}(k)$ is n- dimensional vector of external random disturbances; $\mathbf{A}(k,k-1)$ is $(n\times n)$-matrix of system dynamics; $\mathbf{B}(k,k-1)$ is $(n\times m)$ matrix of control coefficients. The double argument $(k,k-1)$ means that the variable or parameter is used at the time moment k, but its value is based on the former (earlier) data processing including moment ($k-1$). Usually the matrices $\mathbf{A}$ and $\mathbf{B}$ are written with one argument like $\mathbf{A}(k)$ and $\mathbf{B}(k)$ to simplify the text. Besides the main task, optimal state estimation, Kalman filter can be used to solve the following problems: computing of short-term forecasts, estimation of unknown model parameters including statistics of external disturbances and measurement errors (adaptive extended Kalman filter), estimation of state vector components that cannot be measured directly, and fusion of data coming from various external sources (combining of available data directed towards enhancement of its information content).

Obviously stationary system model is described with constant parameters like $\mathbf{A}$ and $\mathbf{B}$. As far as matrix $\mathbf{A}$ creates a link between two consequent system states, it is also called state transition matrix. Discrete time k and continuous time t are linked to each other via data sampling time T_s: $t=kT_s$. In the classic problem statement for optimal filtering the vector sequence of external disturbances $\mathbf{w}(k)$ is supposed to be zero mean white Gaussian noise with covariance matrix $\mathbf{Q}$, i.e. the noise statistics are as follows:

$$E[\mathbf{w}(k)]=0,\ \ \forall k;\quad E[\mathbf{w}(k)\mathbf{w}^{\mathrm{T}}(j)]=\mathbf{Q}(k)\,\delta_{kj},$$

where δ_{kj} is Kronecker delta-function: $\delta_{kj}=\begin{cases}0, & k\neq j\\ 1, & k=j\end{cases}$; $\mathbf{Q}(k)$ is positively defined covariance ($n\times n$) matrix. The diagonal elements of the matrix are variances for the components of disturbance vector $\mathbf{w}(k)$. Initial system state $\mathbf{x}_0$ is supposed to be known and the measurement equation for vector $\mathbf{z}(k)$ of output variables is described by the equation:

$$\mathbf{z}(k)=\mathbf{H}(k)\mathbf{x}(k)+\mathbf{v}(k)\,, \qquad (3)$$

where $\mathbf{H}(k)$ is $(r\times n)$ observation (coefficients) matrix; $\mathbf{v}(k)$ is r-dimensional vector of measurement noise with statistics: $E[\mathbf{v}(k)]=0,\ E[\mathbf{v}(k)\,\mathbf{v}^{\mathrm{T}}(j)]=$ $=\mathbf{R}(k)\delta_{kj}$, where $\mathbf{R}(k)$ is $(r\times r)$ positively defined measurement noise covariance matrix, the diagonal elements of which represent variances of additive noise for each measurable variable. The noise of measurements is also supposed to be zero mean white noise sequence that is not correlated with external disturbance $\mathbf{w}(k)$ and initial system state. For the system (2), (3) with state vector $\mathbf{x}(k)$ it is necessary to find optimal state estimate $\hat{\mathbf{x}}(k)$ at arbitrary moment k as a linear combination of estimate $\hat{\mathbf{x}}(k-1)$ at the previous moment $(k-1)$ and the last measurement available $\mathbf{z}(k)$. The estimate of state vector $\hat{\mathbf{x}}(k)$ is computed as an optimal one with minimizing the expectation of the sum of squared errors, i.e.:

$$E[(\hat{\mathbf{x}}(k)-\mathbf{x}(k))^T(\hat{\mathbf{x}}(k)-\mathbf{x}(k))]=\min_K, \qquad (4)$$

where $\mathbf{x}(k)$ is an exact value of state vector that can be found using deterministic part of the state equation (2); $\mathbf{K}$ is optimal matrix gain that is determined as a result of minimizing quadratic criterion (4).

Thus, the filter is constructed to compute optimal state vector $\hat{\mathbf{x}}(k)$ in conditions of influence of external random system disturbances and measurement noise. Here one of possible uncertainties arises when we don't know estimates of covariance matrices $\mathbf{Q}$ and $\mathbf{R}$. To solve the problem an adaptive Kalman filter is to be constructed that allows for computing estimates of $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$ simultaneously with the state vector $\hat{\mathbf{x}}(k)$. Another choice is in constructing separate algorithm for computing the values of $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$. A convenient statistical algorithm for estimating the covariance matrices was proposed in [11]:

$$\hat{\mathbf{R}}=\frac{1}{2}[\hat{\mathbf{B}}_1+\mathbf{A}^{-1}(\hat{\mathbf{B}}_1-\hat{\mathbf{B}}_2)(\mathbf{A}^{-1})^{\mathrm{T}}];$$

$$\hat{\mathbf{Q}}=\hat{\mathbf{B}}_1-\hat{\mathbf{R}}-\mathbf{A}\hat{\mathbf{R}}\mathbf{A}^{\mathrm{T}},$$

where

$$\hat{\mathbf{B}}_1=E\{[\mathbf{z}(k)-\mathbf{A}\mathbf{z}(k-1)][\mathbf{z}(k)-\mathbf{A}\mathbf{z}(k-1)]^{\mathrm{T}}\};$$

$$\hat{\mathbf{B}}_2=E\{[\mathbf{z}(k)-\mathbf{A}^2\mathbf{z}(k-2)][\mathbf{z}(k)-\mathbf{A}^2\mathbf{z}(k-2)]^{\mathrm{T}}\}.$$

The matrices $\hat{\mathbf{Q}}$ and $\hat{\mathbf{R}}$ are used in the optimal filtering procedure as follows:

$$\mathbf{S}(k)=\mathbf{A}\mathbf{P}(k-1)\mathbf{A}^{\mathrm{T}}+\hat{\mathbf{Q}}; \quad \Delta(k)=\mathbf{S}(k)[\mathbf{S}(k)+\hat{\mathbf{R}}]^{\#};$$

$$\mathbf{P}(k)=[\mathbf{I}-\Delta(k)]\mathbf{S}(k), \quad k=0,1,2,\ldots,$$

where $\mathbf{S}(k)$ and $\mathbf{P}(k)$ are prior and posterior covariance matrices of estimate errors respectively; the symbol "$\#$" denotes pseudo-inverse; $\mathbf{A}^{\mathrm{T}}$ means matrix transposition; $\Delta(k)$ is a matrix of intermediate covariance results. The algorithm was successfully applied to the covariance estimating in many practical applications. The computation experiments showed that the values of $\Delta(k)$ become stationary after about 20–25 periods of time (sampling periods) in a scalar case, though this figure is growing substantially with the growth of dimensionality of the system under study. It was also determined that the parameter estimators are very sensitive to the initial conditions of the system. The initial conditions should differ from zero enough to provide stability for the estimates generated.

Other appropriate instruments for taking into consideration possible statistical uncertainties are fuzzy logic, neuro-fuzzy models, Bayesian networks, appropriate types of distributions etc. Some of statistical data uncertainties, such as missing measurements, extreme values and high level jumps of stochastic origin could be processed with appropriately selected statistical procedures. There exists a number of data imputation schemes that help to complete the sets of the data collected with improving its quality. For example, very often missing measurements for time series could be generated with appropriately selected distributions or in the form of short term forecasts. Appropriate processing of jumps and extreme values helps with adjusting data nonstationarity and to estimate correctly the probability distribution for the stochastic processes under study.

Processing data with missing observations (data are in the form of time series). As of today for the data in the time series form the most suitable imputation techniques are as follows: simple averaging when it is possible (when only a few values are missing); generation of forecast estimates with the model constructed using available measurements; generation of missing estimates from distributions the form and parameters of which are again determined using available part of data and expert estimates; the use of optimization techniques, say appropriate forms of EM-algorithms (expectation maximization); exponential smoothing etc. It should also be mentioned that optimal Kalman filter can also be used for imputation of missing data because it contains “internal” forecasting function that provides a possibility for generating quality short-term forecasts [12]. Besides, it has a feature of fusion the data coming from various external sources and improving this way the quality (information content) of state vector and its forecasts.

Further reduction of this uncertainty is possible thanks to application of several forecasting techniques to the same problem with subsequent combining of separate forecasts using appropriate weighting coefficients. The best results of combining the forecasts are achieved when variances of forecasting errors for different forecasting techniques do not differ substantially (at any rate the orders of the variances should be the same).

Coping with uncertainties of model parameters estimates. Usually uncertainties of model parameter estimates such as bias and inconsistency result from low informative data or data do not correspond to normal distribution, what is required in the case of least squares (LS) application for parameter estimation. This situation may also take place in a case of multi-collinearity of independent variables and substantial influence of process nonlinearity that for some reason has not been taken into account when the model structure was estimated. When power of the data sample is not satisfactory for model construction it could be expanded by applying special techniques, or simulation can be hired, or special model building techniques, such as group method for data handling (GMDH), are applied. Very often GMDH produces results of acceptable quality with rather short samples. If data do not correspond to normal distribution, then maximum likelihood technique could be used or appropriate Monte Carlo procedures for generating Markov Chains (MCMC) [13]. The last techniques could be applied with quite acceptable computational expenses when the number of parameters is not very high.

Generally, model structure and parameters estimation problems are at the core of modeling and they should be paid appropriate attention. Here several techniques should be applied to generate alternative sets of parameter estimates and this way to get a possibility for selecting the best alternative. Statistical criteria indicating model adequacy are helpful to select the best estimates. Also parameter estimates exhibiting the lowest variances are better than others having higher variance. Usually parameter estimation techniques provide the possibility for estimating the variances.

Dealing with model structure uncertainties. When considering mathematical models it is convenient to use proposed here a unified notion (representation) of a model structure which we define as follows: $S = \{r, p, m, n, d, w, l\}$, where r is model dimensionality (number of equations); p is model order (maximum order of differential or difference equation in a model); m is a number of independent variables in the right hand side of a model; n is a nonlinearity and its type (nonlinearity with respect to variables and parameters); d is a lag or output reaction delay time; w is stochastic external disturbance and its type; l are possible restrictions imposed on a model variables and/or parameters. When using DSS, the model structure can practically always be estimated using data. It means that elements of the model structure accept almost always only approximate values.

When a model is constructed on the purpose of forecasting we build several candidates and select the best one of them with a set of model quality statistics. Generally we could define the following techniques to fight structural uncertainties: gradual improvement of model order (AR(p) or ARMA(p, q)) applying adaptive approach to modeling and automatic search for the "best" structure using complex statistical quality criteria; adaptive estimation (improvement) of input delay time (lag) and data distribution type with its parameters; describing detected process nonlinearities with alternative analytical forms with subsequent estimation of model adequacy and forecast quality. As another example of complex statistical model adequacy and forecast quality criterion could be the following:

$$J = \left|1 - R^2\right| + \alpha \ln\left[\sum_{k=1}^{N} e^2(k)\right] + \left|2 - DW\right| + \beta \ln\left(1 + MAPE\right) + U \rightarrow \min_{\hat{\theta}_i},$$

where R^2 is a determination coefficient; DW is Durbin-Watson statistic; $MAPE$ is mean absolute percentage error for estimated forecasts; $\sum_{k=1}^{N} e^2(k) = \sum_{k=1}^{N} [y(k) - \hat{y}(k)]^2$ is the sum of squared model errors; U is Theil coefficient that measures forecasting characteristic of a model; α, β are appropriately selected weighting coefficients (their sum should be equal to 1); $\hat{\theta}_i$ is parameter vector for the i-*th* candidate model. A criterion of this type is used for automatic selection of the best candidate model. The criterion also allows operation of DSS in an automatic adaptive mode. Obviously, other forms of the complex criteria are possible. While constructing the criterion it is important not to overweigh separate members in the right hand side of the expression. As a general recommendation for model structure estimation can be application of appropriate adaptation scheme.

Coping with uncertainties of a level (amplitude) type. The use of random (i.e. with random amplitude or a level) and/or non-measurable variables results in necessity of hiring fuzzy sets for describing such situations. The variable with random amplitude can be described with some probability distribution if the measurements are available or they come for analysis in acceptable time span. However, some variables cannot be measured (registered) in principle, say amount of shadow capital that "disappears" every month in offshore, or amount of shadow salaries paid at some company, or a technology parameter, relative to control system, that cannot be measures on-line due to absence of appropriate sensor. In such situations we could assign to the variable a set of possible values in the linguistic form as follows: *capital amount* = { *very low*, *low*, *medium*, *high*, *very high* }. There exists a complete necessary set of mathematical operations to be applied to such fuzzy variables. Finally fuzzy value could be transformed into usual exact form using known techniques.

Appropriately constructed optimal Kalman filter can also be applied for estimating non-measurable variables using known covariances between measurable and non-measurable variables.

Processing probabilistic uncertainties. To fight probabilistic uncertainties it is possible to hire Bayesian approach that helps to construct models in the form of conditional distributions for the sets of random variables. Usually such models represent the process (under study) variables themselves, stochastic disturbances and measurement errors or noise. The problem of distribution type identification also arises in regression modeling. Each probability distribution is characterized by a set of specific values that random variable could take and the probabilities for these values. The problem is in the distribution type identification and estimating its parameters. The probabilistic uncertainty (will some event happen or not) could be solved with various models of Bayesian type. This approach is known as Bayesian programming or paradigm. The generalized structure of the Bayesian program application includes the following steps: 1) problem description and statement with putting the question regarding estimation of conditional probability in the form: $p(X_i \mid D, Kn)$, where X_i is the main (goal) variable or event; the probability p should be found as a result of application of some probabilistic inference procedure; 2) statistical (experimental) data D and knowledge Kn are to be used for estimating model and parameters of specific type; 3) selected and applied probabilistic inference technique should give an answer to the question put above; 4) analysis of quality of the final result using appropriate statistics. The steps given above are to some extent “standard” regarding model constructing and computing probabilistic inference using statistical data available. This sequence of actions is naturally consistent with the methods of cyclic structural and parametric model adaptation to the new data and operating modes (and possibly expert estimates).

One of the most popular Bayesian approaches today is created by the models in the form of static and dynamic Bayesian networks (BN). Bayesian networks are probabilistic and statistical models graphically represented in the form of directed acyclic graphs (DAG) with vertices as variables of an object (system) under study, and the arcs showing existing causal relations between the variables. Each variable of BN is characterized with complete finite set of mutually excluding states. Formally BN could be represented with the four following components: $\mathbf{N} = \langle \mathbf{V}, \mathbf{G}, \mathbf{P}, \mathbf{T} \rangle$, where $\mathbf{V}$ stands for the set of model variables; $\mathbf{G}$ represents directed acyclic graph; $\mathbf{P}$ is joint distribution of probabilities for the graph variables (vertices), $\mathbf{V} = \{X_1, ..., X_n\}$; and $\mathbf{T}$ denotes conditional and unconditional probability tables for the graphical model variables [14; 15]. The relations between the variables are established via expert estimates or applying special statistical and probabilistic tests to statistical data (when available) characterizing dynamics of the variables hired to construct the model.

The procedure of constructing BN is generally the same as for models of other types, say regression models. The set of the model variables should satisfy the Markov condition that each variable of the network does not depend on all other variables but for the variable's parents. In the process of BN constructing first the problem is solved of computing mutual information values between all variables of the net. Then an optimal BN structure is searched using acceptable quality criterion, say well-known minimum description length (MDL) that allows for analyzing and improving the graph (model) structure on each iteration of computing of the learning algorithm applied. Bayesian networks provide the following advantages for modeling: the model may include qualitative and quantitative variables simultaneously as well as discrete and continuous ones; number of the variables could be very large (thousands); the values for conditional probability tables could be computed with the use of statistical data and expert estimates; the methodology of BN constructing is directed towards identification of actual causal relations between the variables hired what results in high adequacy of the model; the model is also operable in conditions of missing data.

To reduce an influence of probabilistic and statistical uncertainties on models quality and the forecasts based upon them it is also possible to use the models in the form of Bayesian regression based on analysis of actual distributions of model variables and parameters. Consider a simple two variables regression model:

$$y(k)|\, x(k)=\beta_1+\beta_2\, x(k) + u(k), \quad k=0,1,..., n\,.$$

It is supposed that of random values $u_1,..., u_n$ are independent and belong, for example, to normal distribution $\{u(k)\} \sim N(0,\sigma_u^2)$; here vector of unknown parameters includes three elements $\theta=(\beta_1, \beta_2, \sigma_u^2)^T$. The likelihood function for dependent variable $\mathbf{y}=(y_1,..., y_n)^{\mathrm{T}}$ and predictor $\mathbf{x}=(x_1,..., x_n)^{\mathrm{T}}$ without proportion coefficient is determined as follows:

$$L(\mathbf{y}\,|\,\mathbf{x}, \beta_1, \beta_2,\sigma_u) = \frac{1}{\sigma_u^N}\, \exp\left\{ -\frac{1}{2\sigma_u^2} \sum_{k=1}^{N} [\, y(k) - \beta_1 - \beta_2\, x(k)]^2 \right\}.$$

Using simplified (non-informative) distributions for the model parameters

$$g(\beta_1,\beta_2,\sigma_u)=g_1(\beta_1)\, g_2(\beta_2)\, g_3(\sigma_u)\,;$$

$$g_1(\beta_1) \propto \text{const}\,; \;\; g_2(\beta_{21}) \propto \text{const}\,; \;\; g_3(\sigma_u) \propto 1/\sigma_u\,,$$

and Bayes theorem it is possible to find joint posterior distribution for the parameters in the form [16]:

$$h(\beta_1,\beta_2,\sigma_u\,|x,y) \propto \frac{1}{\sigma}\,\frac{1}{\sigma^N} \exp\left[-\frac{1}{2\sigma^2} \sum_{k=1}^{N} \left(y(k) - \beta_1 - \beta_2\, x(k) \right)^2 \right],$$

$$-\infty < \beta_1, \beta_2 < +\infty, \quad 0 < \sigma_u < \infty\,.$$

Maximum likelihood estimates for the model parameters are determined as follows:

$$\hat\beta_1 = \bar y - \hat\beta_2\, \bar x; \quad \hat\beta_2 = \frac{\sum_{k=1}^{N} [x(k)-\bar x][\,y(k)-\bar y\,]}{\sum_{k=1}^{N} [\,x(k)-\bar x\,]\; \sum_{k=1}^{N} [\,y(k)-\bar y\,]}\,,$$

where $\bar{x}=N^{-1}\sum_{k=1}^{N} x(k)$, $\bar{y}=N^{-1}\sum_{k=1}^{N} y(k)$, with unbiased sample estimate of variance:

$$\hat{\sigma}_u^2 = s^2 = \frac{1}{N-2}\sum_{k=1}^{N} [y(k) - \hat{\beta}_1 - \hat{\beta}_2 x(k)] .$$

Joint posterior density for the model parameters corresponds to two-dimensional Student distribution:

$$h_1(\beta_1, \beta_2 | \mathbf{y}, \mathbf{x}) \propto \Big\{(N-2)s^2 + N(\beta_1 - \hat{\beta}_1)^2 + (\beta_2 - \hat{\beta}_2)^2 \sum_{k=1}^{N} x(k)^2 + $$
$$+ 2(\beta_1 - \hat{\beta}_1)(\beta_2 - \hat{\beta}_2) \sum_{k=1}^{N} x(k) \Big\}^{-0,5N} .$$

This way we get a possibility for using more exact distributions of models variables and parameters what is necessary to enhance model quality. Using new observation x^* and prior information regarding particular model it is possible to determine the forecast interval for the dependent variable y^* as follows:

$$p(y^* | x^*) = \iiint L(y^* | x^*, \beta_1, \beta_2, \sigma)\, h(\beta_1, \beta_2, \sigma) | \mathbf{x}, \mathbf{y})\, d\beta_1, d\beta_2, d\sigma .$$

Another useful Bayesian approach is in hierarchical modeling that is based on a set of simple conditional distributions comprising one model. The approach is naturally combined with the theory of computing Bayesian probabilistic inference using modern computational procedures [17]. The hierarchical models belong to the class of marginal models where the final result is provided in the form of a distribution $P(\mathbf{y})$, where $\mathbf{y}$ is available data vector. The models are formed from the sequence of conditional distributions for selected variables including the hidden ones. The hierarchical representation of parameters usually supposes that data $\mathbf{y}$ is situated at the lower (first) level, model parameters (second level) $\theta = (\theta_i, i{=}1, 2,..., n)$, $\theta_i \sim N(\mu, \tau^2)$, determine distributions of dependent variables $y_i \sim N(\theta_i, \sigma^2)$, $i{=}1,2,..., n$, and parameters $\{\theta_i\}$ are determined by the pair (μ, τ^2) of the third level. Supposing the parameters σ^2 and τ^2 accept known finite values, and parameter μ is unknown with the prior π_μ, then joint prior density for (θ, μ) could be presented in the form: $\pi_\mu(\mu)\prod_i \pi_\theta(\theta_i | \mu)$, and the prior for parameter vector θ will be defined by the integral: $p(\theta) = \int \pi_\mu(\mu) \prod_i \pi_\theta(\theta_i | \mu)\, d\mu$.

Uncertainties associated with expert estimates. To decrease influence of the expert estimate uncertainties they are to be processed adequately before practical use. Possible uncertainties of the expert estimates can be caused by the following reasons: uncertainties associated with input information, and the knowledge and experience of an expert; uncertainties associated with the way of thinking used by specific expert and the methodology he hires as well as the information processing “analytic” machine that is functioning in his mind etc. Such uncertainties require application of special techniques to reduce their influence on the quality of final result.

DATA, MODEL AND FORECASTS QUALITY CRITERIA

To achieve reliable high quality final result of risk estimation and forecasting at each stage of computational hierarchy separate sets of statistical quality criteria have been used. Data quality control is performed with the following criteria:

– analysis of database for missing values using developed logical rules, and imputation of missed values with appropriately selected techniques;

– analysis of data for availability of outliers with special statistical tests, and processing of outliers to reduce their negative influence on statistical properties of the data available;

– normalizing data in the selected range in a case of necessity;

– application of low-order digital filters (usually low-pass filters) for separation of observations from measurement noise;

application of optimal (very often Kalman) filters for optimal state estimation and fighting stochastic uncertainties;

– application of principal component method to achieve desirable level of orthogonalization between the variables selected;

– computing of extra indicators for the use in regression and other models (say, moving average processes based upon measurements of dependent variables).

It is also useful to test how informative is the data collected. Very formal indicator for the data being informative is its sample variance. It is supposed formally that the higher is the variance the richer is the data with information. Another criterion is based on computing derivatives with a polynomial that describes data in the form of a time series. For example, the equation given below can describe rather complex process with nonlinear trend and short-term variations imposed on the trend curve:

$$y(k)=a_0+\sum_{i=1}^{p} a_i y(k-i)+c_1 k+c_2 k^2+...+c_m k^m+\varepsilon(k),$$

where $y(k)$ is basic dependent variable; a_i, c_i are model parameters; $k=0,1,2,...$ is discrete time; $\varepsilon(k)$ is a random process that integrates the influence of external disturbances to the process being modeled as well as model structure and parameters errors. The autoregressive part of model (1) describes the deviations that are imposed on a trend, and the trend itself is described with the *m*-th order polynomial of discrete time k. In this case maximum number of derivatives could be *m*, though in practice actual number of derivatives is defined by the largest number i of parameter c_i, that is statistically significant. To select the best model constructed the following statistical criteria are used: determination coefficient (R^2); Durbin-Watson statistic (DW); Fisher *F*-statistic; Akaike information criterion (*AIC*), residual sum of squares (*SSE*), and some others. The forecasts quality is estimated with hiring the criteria mentioned above in expressions (1) and (2). To perform automatic model selection the above mentioned combined criteria (1) could be hired. The power of the criterion was tested experimentally and proved with a wide set of models and statistical data. Thus, the three sets of quality criteria are used to insure high quality of final result.

ILLUSTRATIVE EXAMPLE OF DATA PROCESSING METHODOLOGY APPLICATION AND REDUCTION OF INFLUENCE UNCERTAINTIES

Consider closing stock prices in USD for IBM company given by the site Yahoo! Finance. The learning sample was taken in the period from August 8, 2016, to November 23, 2020. The test sample was taken in the period from November 24, 2020 to May 12, 2021. Short-term forecasting was performed with neural networks MLP and LSTM. For the sake of convenience the networks were designated as follows: MLP ($n1$, $n2$, w), where $n1$ is a number of neurons in the first hidden layer; $n2$ number of neurons in the second hidden layer; w is a size of window for input data, i.e. number of preceding measurements of a time series that influence current value. The LSTM network has the following representation: LSTM (n, w), where n is a number of neurons in hidden layer; w is size of data window. The best forecasting results were achieved with the networks MLP (32, 16, 5) and LSTM (64, 75). Statistical characteristics of the results are given in Table 1 below.

Table 1. Results of short-term forecasting with ARIMA, MLP and LSTM

Statistic	ARIMA(5,1,5)	MLP(32,16,10)	LSTM(64,75)
RSME	10.202	5.145	6.123
MAPE	5.702	3.129	4.099
MSE	104.094	26.498	37.621

Better forecasting results were achieved after application of exponential smoothing to initial data (Table 2). This way we reduced noise uncertainties and prepared the statistical data to further use by neural networks.

Table 2. Results of short-term forecasting with ARIMA, MLP and LSTM and preliminary data processing

Statistic	ARIMA(5,1,5)	MLP(32,16,10)	LSTM(64,75)
RSME	8.974	4.320	5.382
MAPE	4.682	2.935	3.116
MSE	95.169	22.806	33.127

It can be seen that MLP again produced the best result of short-term forecasting but all statistics are lower (better) than in previous case without preliminary data processing by exponential smoothing.

CONCLUSIONS

The general methodology was proposed for mathematical modeling and forecasting dynamics of economic and financial processes that is based on the system analysis principles. One of the main principles is identification and taking into consideration possible uncertainties associated with data and expert estimates. As instrumentation for fighting possible structural, statistic and parametric uncertainties the following techniques are proposed to use: digital filtering, optimal Kalman filter, various missing data imputation techniques, multiple methods for model parameter estimation, and Bayesian programming approach. The computational experiments carried out by the authors showed that the instrumentation for fighting uncertainties has always provided better results regarding model adequacy and quality of forecasts than processing data without these instruments. Thus, it is highly advisable to use these data processing instruments for improving quality of finale results of statistical and experimental data analysis. The illustrative example, given above, shows that appropriate preliminary data processing technique results in improvement of model adequacy and short-term forecasts.

REFERENCES

1. R.S. Tsay, *Analysis of financial time series*. Chicago: Wiley & Sons, Ltd., 2010, 715 p.
2. L. Harris, X. Hong, and Q.Gan, *Adaptive Modeling, Estimation and Fusion from Data*. Berlin: Springer, 2002, 323 p.
3. P. Congdon, *Applied Bayesian Modeling*. Chichester: John Wiley & Sons, Ltd., 2003, 472 p.
4. S.M. DeLurgio, *Forecasting Principles and Applications*. Boston: McGraw-Hill, 1998, 802 p.
5. S.J. Taylor, "Modeling stochastic volatility: a review and comparative study," *Mathematical Finance*, vol. 4, no. 2, pp. 183–204, 1994.
6. F. Burstein and C.W. Holsapple, *Handbook of Decision Support Systems*. Berlin: Springer-Verlag, 2008, 908 p.
7. C.W. Hollsapple and A.B. Winston, *Decision Support Systems*. Saint Paul (MN): West Publishing Company, 1996, 860 p.
8. P.I. Bidyuk, O.P. Gozhiy, *Computer decision support systems*. Mykolaiv: Petro Mohyla Black Sea National University, 2012, 380 p.
9. E. Xekalaki and S.Degiannakis, *ARCH Models for Financial Applications*. Chichester: Wiley & Sons, Inc., 2010, 550 p.
10. C. Chatfield, *Time Series Forecasting*. Boca Raton: Chapman & Hall/CRC, 2000, 267 p.
11. W.N. Anderson, G.B. Kleindorfer, P.R. Kleindorfer, and M.B. Woodroofe, "Consistent estimates of the parameters of a linear system," *The Annals of Mathematical Statistics*, vol. 40, no. 6, pp. 2064–2075, 1969.
12. B.P. Gibbs, *Advanced Kalman Filtering, Least-squares and Modeling*. Hoboken (New Jersey): John Wiley & Sons, Inc., 2011, 627 p.
13. W.R. Gilks, S. Richardson, and D.J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. New York: Chapman & Hall/CRC, 2000, 486 p.
14. F.V. Jensen and Th.D. Nielsen, *Bayesian Networks and Decision Graphs*. New York: Springer, 2007, 457 p.
15. M.Z. Zgurovsky, P.I. Bidyuk, O.M. Terentyev, and T.I. Prosyankina-Zharova, *Bayesian Networks in Decision Support Systems*. Kyiv: Edelweiss, 2015, 300 p.
16. J.M. Bernardo and A.F.M. Smith, *Bayesian theory*. New York: John Wiley & Sons, Ltd., 2000, 586 p.
17. W.M. Bolstad, *Understanding Computational Bayesian Statistics*. Hoboken (New Jersey): John Wiley & Sons, Ltd, 2010, 334 p.

Received 23.06.2023

INFORMATION ON THE ARTICLE

Liudmyla B. Levenchuk, ORCID: 0000-0002-8600-0890, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: lusi.levenchuk@gmail.com

Oxana L. Tymoshchuk, ORCID: 0000-0003-1863-3095, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: oxana.tim@gmail.com

Vira H. Huskova, ORCID: 0000-0001-7637-201X, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: guskovavera2009@gmail.com

Petro I. Bidyuk, ORCID: 0000-0002-7421-3565, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Ukraine, e-mail: pbidyuke_00@ukr.net

НЕВИЗНАЧЕНОСТІ В ОБРОБЛЕННІ ДАНИХ, ПРОГНОЗУВАННЯ І ПРИЙНЯТТЯ РІШЕНЬ / Л.Б. Левенчук, О.Л. Тимощук, В.Г. Гуськова, П.І. Бідюк

Анотація. Прогнозування, динамічне планування та оброблення поточних статистичних даних визначаються як процес оцінювання поточного стану підприємства на ринку порівняно з іншими конкуруючими підприємствами та визначення подальших цілей, а також послідовностей дій та ресурсів, необхідних для досягнення визначених цілей. Для здійснення прогнозування високої якості запропоновано визначити та врахувати можливі невизначеності, пов’язані з даними та експертними оцінками. Це один з принципів системного аналізу, який застосовується для досягнення високої якості кінцевого результату. Наведений огляд деяких невизначеностей та ілюстративний приклад, який показує поліпшення кінцевого результату після врахування можливої стохастичної невизначеності.

Ключові слова: математична модель, невизначеності статистичних даних, принципи системного аналізу, прогнозування, система підтримання прийняття рішень.

ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ ПРОБЛЕМИ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ ПІДТРИМАННЯ ПРИЙНЯТТЯ РІШЕНЬ

UDC 62-50
DOI: 10.20535/SRIT.2308-8893.2023.3.06

A CONCATENATION APPROACH-BASED DISEASE PREDICTION MODEL FOR SUSTAINABLE HEALTH CARE SYSTEM

K. THARAGESWARI, N. MOHANA SUNDARAM, R. SANTHOSH

Abstract. In the present world, due to many factors like environmental changes, food styles, and living habits, human health is constantly affected by different diseases, which causes a huge amount of data to be managed in health care. Some diseases become life-threatening if they are not cured at the starting stage. Thus, it is a complex task for the healthcare system to design a well-trained disease prediction model for accurately identifying diseases. Deep learning models are the most widely used in disease prediction research, but their performance is inferior to conventional models. In order to overcome this issue, this work introduces the concatenation of Inception V3 and Xception deep learning convolutional neural network models. The proposed model extracts the main features and produces the prediction result more accurately than traditional predictive models. This work analyses the performance of the proposed model in terms of accuracy, precision, recall, and f1-score. It compares the proposed model to existing techniques such as Stacked Denoising Auto-Encoder (SDAE), Logistic Regression (LR), MLP, MLP with attention mechanism (MLP-A), Support Vector Machine (SVM), Multi Neural Network (MNN), and Hybrid Convolutional Neural Network (CNN)-Random Forest (RF).

Keywords: feature extraction, disease prediction, deep learning, Inception V3, Xception.

INTRODUCTION

The World Health Organization defines the health care system as the organization of people that is mainly constructed to maintain, restore, and monitor the health details of the public. The health system improves people's health by providing personal care given by hospitals and doctors. As a result, the primary goal of the healthcare system is to keep people healthy by detecting diseases early and treating them appropriately. Obtaining an efficient health care system provides benefits to maintaining people's health.

In recent years, the development in the medical field has cured patients of various diseases, but still, people are affected by some diseases due to their unpredictable nature, and it causes a severe life-threatening problem for people. The early prediction of those diseases saves the lives of many people. In the health care system, doctors who use computer-aided diagnoses to quickly treat different diseases need to be able to recognize, analyze, and classify data [5]. Thus, the healthcare system requires a well-trained disease predictive model to accurately predict diseases. Based on the predicted disease, doctors can provide the correct treatment and maintain the patient's health. So, this proposed work designs a system that can efficiently determine the condition to predict the disease based on the given input data about the patient's health.

The accuracy of the prediction result depends on the technique used in training the model, and the training depends on the sample data used for training. So, the sufficiency of the data sample is very important in training the model to get a better prediction result. In order to overcome the deficiency of sample data, the transfer learning technique is used in the deep learning model. Deep learning is a subtype of the machine learning approach. The rapid growth of deep learning in various fields proposes various representative techniques. Because of the better ability to learn the features of input data, deep learning has progressively replaced traditional machine learning techniques [6]. Deep learning has the capacity to determine features automatically from a given dataset for each specific application.

Transfer Learning is a kind of deep learning technique that uses pre-trained knowledge from the past to train the new model and deploy the trained features of a large dataset into a small dataset. Hence, using this pre-trained transfer learning technique in the deep learning model will reduce the time taken for the classification and prediction of disease [5]. However, the deep learning technique used by the conventional model has several advantages, but the accuracy performance is not better. So, the main motive of this research work is to improve prediction accuracy by modifying the state-of-the-art method by concatenating networks. This research work uses a concatenation of two different pretrained Convolutional Neural Network (CNN) models called Inception V3 and Xception.

The Inception model consists of multiple convolution filters of various sizes, and hence it can improve the adaptability of the network and extract more copious features of different scales. Simultaneously, by using the Network in Network model, the Inception model can significantly reduce the parameters of the model. Hence, the network can minimize the number of convolution filters as much as possible without losing model feature representation, thus minimizing the complexity of the model [3]. Whereas, in the exception model, the term "exception" refers to an extreme. The exception model takes the rule of the Inception model to an extreme, so it provides the added advantage of feature extraction in a pointwise manner also. The concatenation of two different models makes the proposed system stronger due to the capability of multiple feature extraction [16]. The concatenation model reduces the complexity and provides accurate prediction results based on the given input data, which could be sustainable and reliable for health care.

The research contribution is summarized as follows:

- Presented a sustainable healthcare prediction system using a concatenation of Inception V3 and Xception to predict the disease based on input data.
- Presented an intense experimental analysis to validate the performance of the proposed model using standard benchmark data.
- Presented a comparative analysis of the proposed model with existing prediction models such as MLP, Logistic regression (LR), stacked denoising autoencoder (SDAE) and hybrid Convolutional Neural Network (CNN) – Random Forest (RF) for performance validation.

The further discussions are arranged in the following order: Section two comprises a detailed literature review and section three presents the proposed model, experimental analysis and its results comparisons are presented in section four and the conclusion is presented in the last section.

RELATED WORKS

In this section, research related to the existing disease prediction system was enumerated. Various techniques, feature advantages, and disadvantages are reviewed for investigation, and finally, the limitations are discussed to structure the research motivation. The healthcare system is rapidly becoming a necessary tool for offering a well-rounded chance to meet the requirements of public health. Integration of healthcare with a recommended system to meet the needs of older and chronically diseased people was reported in [1]. Deep learning is a subfield of machine learning that is broadly used in healthcare systems for the classification, identification, and prediction of clinical data. A feature selection algorithm-based prediction model reported in [3] uses a combination of the FCMIM-SVM approach to detect heart disease. The features for the prediction model are obtained using a fast-conditional mutual information feature selection algorithm, and local learning procedures are followed to remove the redundant and irrelevant features. Machine learning techniques have been broadly adopted in various fields, particularly in medical diagnosis. Multi-feature extraction using Inception-V3 and Densnet-201 is achieved to predict the brain tumor [4].

A feature selection algorithm based on Chronic Obstructive Pulmonary Disease prediction is reported in [6] uses an instance-based and feature-based transfer learning Balanced Probability Distribution (BPD) and cross-domain feature filtering algorithm. Various feature selection algorithms like the AdaBoost algorithm, TCA algorithm, and Multi-Task Learning (MTL) algorithm are compared with the BPD algorithm, and the result demonstrates the superior prediction performance of the BPD algorithm. A comparative analysis of machine learning techniques for disease prediction reported in [8] employs decision trees, K-nearest neighbor, and logistic regression algorithms to predict kidney disease. From the findings, it demonstrates that the decision tree approach and logistic regression give better performance in predicting kidney disease. The article [5] investigates the viability and effectiveness of various machine learning algorithms like REP Tree, Random Tree, Linear Regression, M5P Tree, Naive Bayes, J48, and JRIP to predict cardiovascular disease. And from the analysis, the result demonstrates the superior prediction performance of the Random Tree approach.

The computer-aided screening method reported in [17] utilizes a comparison of 13 pre-trained CNN models like AlexNet, GoogleNet, VGG16, VGG19, etc. and three deep learning based classifiers, namely K-Nearest Neighbor, Support Vector Machine, and Naive Bayes for the analysis of Covid19 X-Ray and CT Scan images. The result demonstrates that VGG19 with SVM classifier provides superior performance to other methods. A novel Leaf GAN (Generative Adversarial Network) method reported in [19] utilizes a data augmentation method to identify disease-affected grape leaves. Initially, for training the GAN model, four types of grape leaf disease images are generated using an image generator model, and secondly, to identify original and duplicate images, an image discriminator is used. Finally, a deep regret gradient penalty method is employed for the stabilization of the GAN model.

A combination of convolutional neural networks with recursive neural networks reported in [20] utilizes an automatic prediction method for identifying and categorizing the different kinds of blood cells. Due to a better understanding of the features of blood cell images, accurate prediction and classification of blood cells are attained. The segmentation of breast masses in mammograms reported in [21] utilizes a comparison of various kinds of deep learning models for the segmentation and classification of breast lesions. The result demonstrates that the VGG19 and ResNet50V2 models perform better in the classification of breast lesions than other models.

The identification of the Citrus Disease Severity reported in [19] trains and compares six different types of deep learning models to predict the Citrus Disease Severity. From the analysis, it demonstrates that GAN-based data augmentation with the Inception V3 model provides superior performance. A disease prediction model reported in [2] utilizes a multi-stage model by the combination of co-clustering and supervised machine learning methods to predict the intravenous immunoglobulin resistance in Kawasaki disease. Initially, co-clustering is used to cluster the missing data pattern blocks. Secondly, the selection of data features is obtained by group lasso. Finally, the prediction of immunoglobulin resistance in a patient is obtained using an explainable boosting method. The machine learning approach reported in the article [7] utilizes a cloud-centric IoT system for the prediction of skin disease. This research mainly focuses on the evaluation of six deep learning models, namely VGG16, Inception, Xception, MobileNet, ResNet50, and DenseNet161. Based on this evaluation, the article created a two-phase classification process by the Targeted Ensemble Machine Classification Model (TEMCM).

Apart from the classification of images, CNN are widely used for the segmentation of images. The segmentation model reported in [20] utilized a CNN along with an optimization technique to segment the various types of land in the Amazon. A deep learning model reported in [8] utilizes the fused outputs of ResNet50, Xception, and DenseNet in the FC layer of the super learner model for classifying the type of vehicle. From the above survey, it is observed that the performance of the prediction model depends on the selection of a suitable learning algorithm and classifiers. Various types of machine learning algorithms are mostly used in the research. From the above literature review, it is observed that the Dense Net and ResNet perform well. However, in some of the articles, concatenation of convolutional neural network models was utilized, but the performance can be improved further using novel architectures. In order to extend the performance of the concatenation-based CNN prediction system, a SoftMax discriminator algorithm is used along with the concatenation model, and it is discussed in the following section.

PROPOSED WORK

The proposed predictive model contains a concatenation of Inception V3 and Xception CNN (Fig. 1). The proposed work consists of three stages. The first stage involves Normalizing of the given medical input data. The second stage involves features extraction from the input-data using concatenation approach. The third stage involves the prediction of output using soft max discriminant classifier.

Data Normalizing Stage. This stage starts from the collection of medical datasets. Since the non-standardized input data’s take more time for learning process, the data in the dataset are subjected to Normalizing procedure. Based on the type of input data standardization will be applied to every input data to have the same dimension/size.

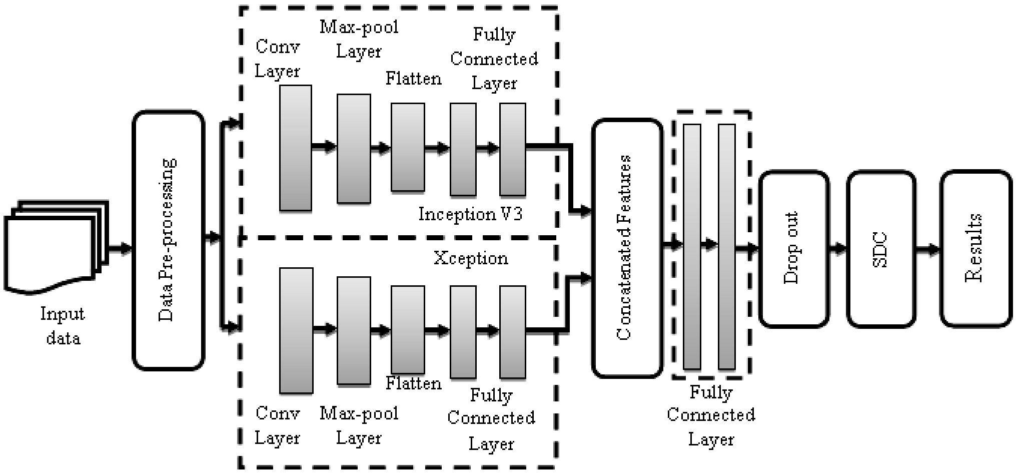


Fig. 1. Architecture of Proposed Concatenation Model

Feature Extraction from the Concatenation Model. Feature extraction is an essential step to be done before predicting the output, which is used to determine knowledge from datasets and it mainly focus on the selection of representative features from the given dataset, which will make the classifier to easily diffirentiate between the various classes. Additionally, it also provide benefit by reducing the computational time for training the classifier. A well-organized feature extraction method will improve the accuracy of predicted output. So, in this work, for attaining better feature extraction parallel deep feature extraction technique based on transfer learning method is applied to the concatenation of Inception V3 and Xception CNN models. The Normalized input data is given to the Xception and Inception V3 CNN model for feature extraction. To improve the quality of resultant feature map, the feature extracted from the two models are combined. Xception model convolutes the input data by both depth wise and pointwise manner (Fig. 2) and Inception model convolutes the input data by depth wise manner (Fig. 3).

Prediction of output using Soft-max Discriminant Classifier (SDC). The correct prediction of the given input data is done by the process of classification. The task of the classifier is to approximate a feature map function from input data to discrete output data. The important function of SDC is to identify the particular class to which the testing data belongs to. This can be done by using the calculation of weighing distance between the test data and train data. In the proposed work, the SDC is used as a binary classifier since it needs only two classes, where class 1 represents abnormal status and the class 2 represents the normal status. The following Figs. 2 and 3 show the architecture of Inception V3 Model and Xception model respectively.

Consider the input medical dataset $D = \{D_1, D_2, \ldots, D_n\}$. The normalization of the input data is given by

$$f(D) = x + y\frac{D - \min(D_i)}{\max(D_i) - \min(D_i)},$$

where D_i defines the input medical dataset of D, x and y represents the constant of normalizer.This standardized input datas are further given as input to the training and testing in the model. The standardized input data from the given medical dataset is given as input to both the Xception and Inception V3 model. The feature map expression of g_1 from the pre-trained Inception V3 model is given

$$q_1(i) = r_1(c_1 = c_{1i} \mid g_1(v_1, b_1)) .$$

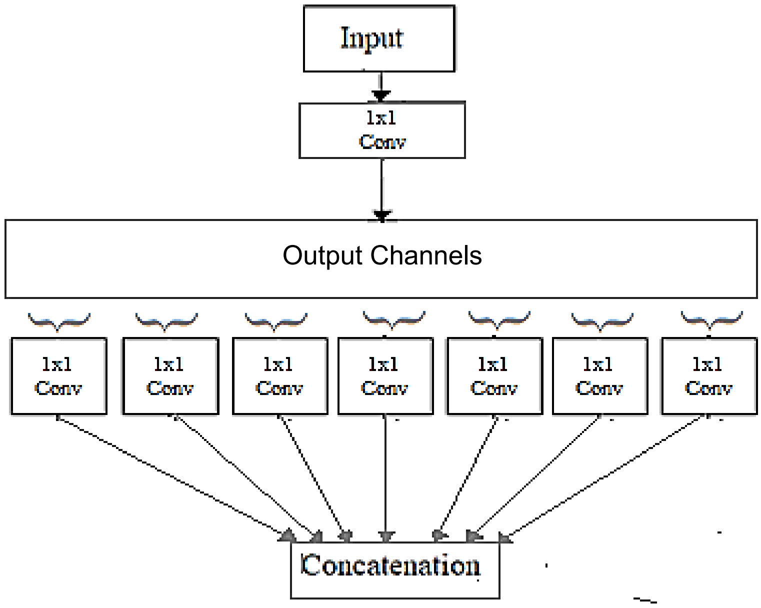


Fig. 2. Architecture of Xception

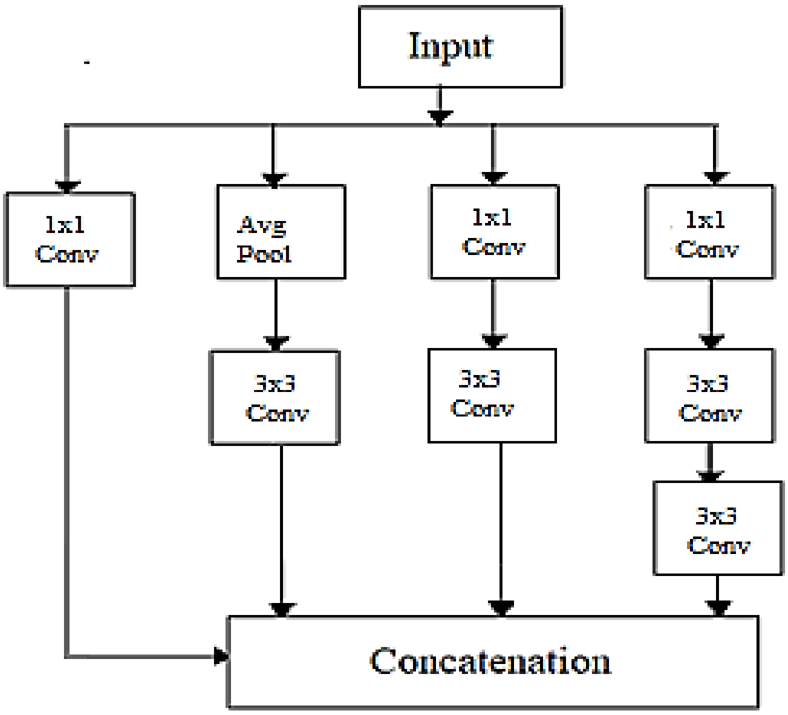


Fig. 3. Architecture of Inception V3

The feature map expression of g_2 from the pre-trained Xception model is given by

$$q_2(i) = r_2(c_2 = c_{2i} \mid g_2(v_2, b_2)),$$

where v_1 and v_2 are the weight vector and b_1 and b_2 are the bias vector of the Inception V3 and Xception model. The final feature map expression of G from the concatenation model is given by

$$Q(i) = R(C = c_i \mid G;(V, B)),$$

where $Q = q_1 + q_2$, $R = r_1 + r_2$, $C = c_1 + c_2$, $G = g_1 + g_2$, $V = v_1 + v_2$ and B= $B = b_1 + b_2$. The final feature map output obtained from the concatenation network is given as input to the FC layer of the proposed mode ,which flattens the input into a fixed length vector form. Next, the output of FC layer is given as input to the drop out layer which is used for regularizing the data by avoiding overfitting in the network. Finally, the output "*s*" from drop out layer is given to the SDC for the classification of data.

Training and Testing of Data's using SDC. Consider the training data set be $S = \{S_1, S_2, S_3, \ldots, S_m\} \in K^{e \times f}$ is determined from the *m* distinct classes:

$S_m = \{S_1^m, S_2^m, S_3^m, \ldots, S_{f_m}^m\} \in K^{e \times f_m}$ represents totally f_m data from the m^{th} class, where $\sum_{i=1}^{m} f_i = f$.

Consider the test data from the drop out layer as as $s \in K^{e \times 1}$ For representing the test data "*m*" class data's are used and by which a minimum reconstruction error is attained. The distance in between the *m* class data and the testing data helps SDC for classifying the input data. The expression of SDC will be given by the following equations:

$$l(s) = \text{argmax}\, T_s^i;$$

$$l(s) = \text{argmax} \log \left(\sum_{j=1}^{f_i} \exp\left(-\alpha \| s - s_j^i \|_2\right) \right),$$

where T_S^i, $l(s)$ defines the distance between the i^{th} class and the test data, from this *s* can be identified. α defines the penalty parameter, which can be used when $\alpha > 0$. If "*s*" corresponds to the i^{th} class, then s and s_j^i will have the same features and hence $\left\| s - s_j^i \right\|_2$ will goes to zero. This indirectly indicates that the features of the test data match with the features of the i^{th} class. Therefore T_s^i helps to attain the higher value asymptotically and hence T_s^i is maximized. By this way, SDC learns to predict the corresponding class of the given input data.

The prediction output of the SDC will be represented using binary classification as follows:

$$Output\,Z = \begin{cases} class1 \text{ if } T_s \le 0, \\ class2 \text{ if } T_s > 0. \end{cases}$$

Where *P* defines the decision condition of the SDC, which is predefined during the training period of SDC. *Class 1* defines the abnormal status and *class 2* defines the normal status. The decision condition will get change based on the type of disease.

Pseudocode for the proposed Concatenation of Inception V3 and Xception Model

$Input : D = \{D1, D2, D3, \ldots, Dn\}$

$$Output : Z = \begin{cases} class1 \text{ if } T_s \le 0 \\ class2 \text{ if } T_s > 0 \end{cases}$$

Begin
Initialize data normalization for the given input $f(D)$
Give the normalized data to both Inception V3 and Xception Pre-trained **Model**
Obtain the feature map of Inception V3 model as q_1
Obtain the feature map of Xception model as q_2
Concatenate both pre-trained model and obtain final feature map as Q
Flatten the final feature map using FC layer
Regularize the FC output data using drop out layer
Give the regularized data to SDC and obtain the classification
Obtain the classification of SDC from $l(S) = \operatorname{argmax} T_s^i$
If $T_s \le 0$
SDC predicts the output as abnormal status
else
SDC predicts the output as normal status
End if
End

RESULTS AND DISCUSSION

The proposed Concatenation of Inception V3-Xception prediction model is experimentally validated using MatLab 14.1 installed in an intel i3 processor 2.20 GHZ frequency with 8 GB memory. In order to obtain the accurate prediction of disease, the metrices such as accuracy, precision, recall, and f1-score are evaluated in the proposed work The first data set is collected from the Cardiology Department of Chinese PLA General Hospital [25]. The feature includes demographics, vital signs, lab tests, echocardiography, comorbidities, length of stay, and medications. Total 105 features are obtained from each patient and a total of 736 patient data are used in the dataset. The second data set is collected from the UCI machine learning repository [26]. The dataset includes 76 features which include demographics, vital signs, cholesterol, echocardiography, and medications. The benefit of using this dataset is that it allows investigating the classification with multiple features. The simulation parameters used in the proposed model is listed in Table 1.

T a b l e 1 . Simulation parameters

No	Parameter	Range/Value
1	Normalizer constant for input	x=0.1, y=0.7
2	Learning rate	0.9
3	Number of epochs	500

TP means the actual medical data containing abnormal value is correctly predicted as abnormal patient, FN means the actual medical data containing abnormal value is incorrectly predicted as Normal patient, FP means the actual medical data containing Normal value is incorrectly predicted as abnormal patient and TN means the actual medical data containing Normal value is correctly predicted as Normal data. The performance of the proposed work is determined in terms of the following metrices:

$$Recall = \frac{TP}{TP + FN};$$

$$Precision = \frac{TP}{TP + FP};$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN};$$

$$\text{F1-score} = \frac{2TP}{2TP + FP + FN}.$$

The Performance metrics of the proposed model is presented in the form of the following Table 2.

T a b l e 2 . Performance metrics of proposed model

No	Performance metrics	Dataset 1	Dataset 2
1	Recall	0.993	0.991
2	Precision	0.990	0.987
3	F1-score	0.991	0.914
4	Accuracy	0.985	0.983

Further the performance of the proposed concatenation model has been compared with existing techniques evaluated in Chen et al. [21] research work for dataset 1 and Sudarshan et al. [26] research work for dataset 2. For the data set 1, techniques like stacked denoising auto-encoder (SDAE), logistic regression (LR), MLP, MLP with attention mechanism (MLP-A) and Multi neural networks (MNN) are used to compare with proposed model. For the dataset 2 techniques like support vector machine (SVM), logistic regression (LR), Random Forest (RF), swarm artificial neural network (S-ANN) and multi neural networks are used to compared with proposed concatenation model.

Figs. 4 and 5 presents the precision analysis of proposed model and conventional models for dataset 1 and dataset 2 respectively. From the results, it is clear that the proposed concatenation model exhibits maximum precision which indicates the classification performance of proposed concatenation model has been increased due to the multi feature selection and processing using soft-max Discriminant classifier. Similarly, for dataset 2 the maximum performance is obtained by the proposed concatenation model whereas conventional methods obtain minimum precision values compared to proposed concatenation model. The average precision value attained by the proposed concatenation model for dataset 1 is 0.990 and for dataset 2 the obtained precision is 0.987 which is much better than the conventional methods.

The recall metrics of the proposed model and conventional models for dataset 1 and dataset 2 has been presented in Figs. 4 and 5 respectively.

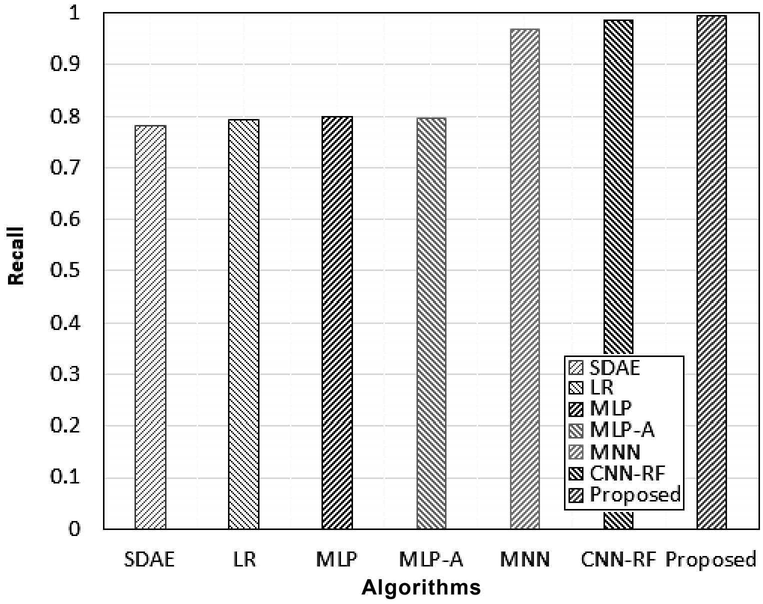


Fig. 4. Recall analysis for dataset 1

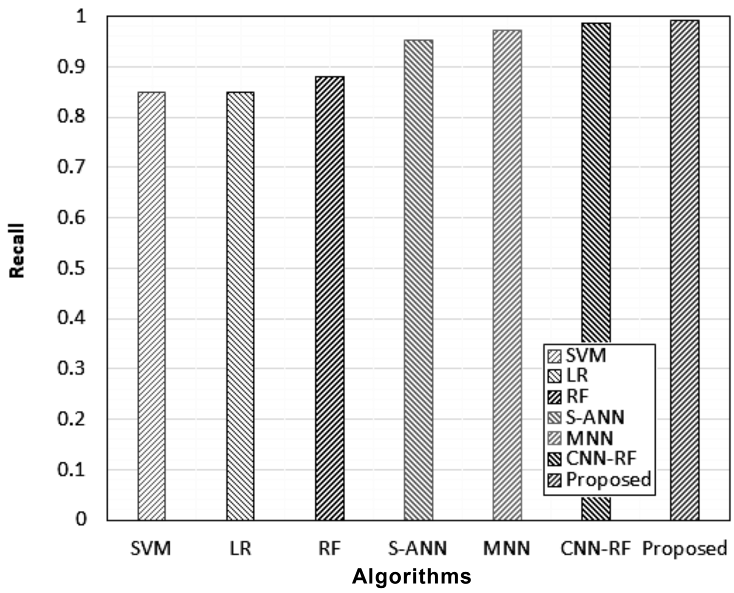


Fig. 5. Recall analysis for dataset 2

Results shows that the maximum recall obtained by the proposed model for both datasets. Though the performance of MNN is much better than other conventional techniques however it is lesser than the CNN-RF model and proposed concatenation model. The average recall value obtained by the proposed model for dataset 1 is 0.993 and for dataset 2 the obtained recall value is 0.991. The F1-score analysis for the proposed model and existing models are comparatively presented in Figs. 6 and 7 for dataset 1 and dataset 2 respectively. Based on the recall and precision values, the f1-score has been obtained and presented. From the results it clear that the maximum score is obtained by the proposed concatenation model when compared to other conventional techniques. The average f1-score obtained by the proposed concatenation model for dataset 1 is 0.991 and 0.914 for dataset 2.

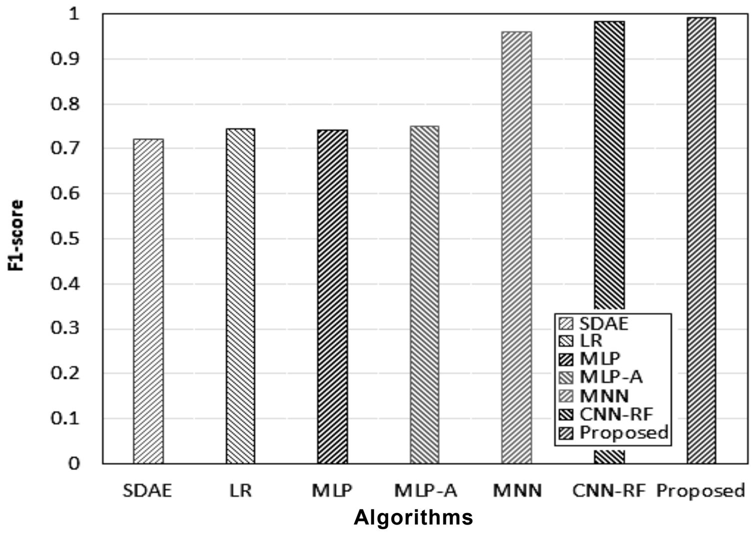


Fig. 6. F1-score analysis for dataset 1

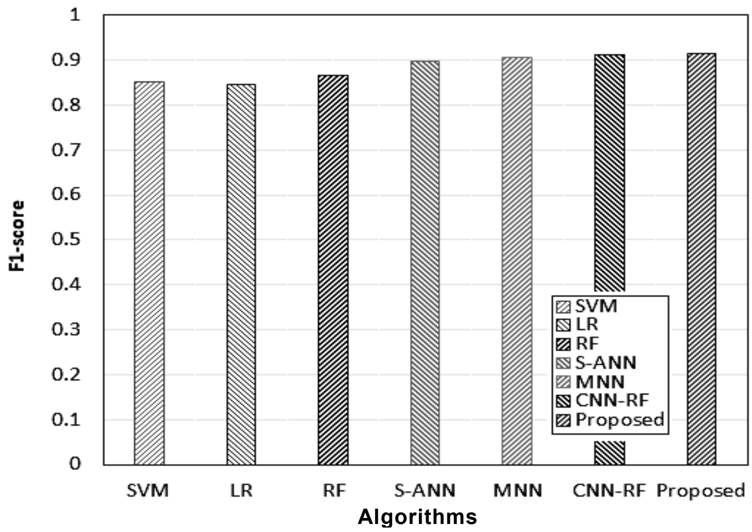


Fig. 7. F1-score analysis for dataset 2

The accuracy of the proposed concatenation model and conventional models are comparatively analyzed and depicted in Figs. 8 and 9 for dataset 1 and dataset 2 respectively. It can be analyzed from the results; the maximum accuracy is determined by the proposed concatenation model for both datasets whereas the performances of conventional models are lesser than the proposed model accuracy values. The maximum accuracy obtained by the proposed model is 0.985 for dataset 1 and 0.983 for dataset 2. The accuracy obtained by MNN for dataset 1 is 0.966 and dataset 2 is 0.968 which is nearly 2% lesser than the proposed concatenation model. The accuracy obtained by CNN-RF for dataset 1 is 0.973 and dataset 2 is 0.978 which is also nearly 1% lesser than the proposed concatenation model. The multi feature selection using concatenation model and prediction using SDC increases the prediction accuracy of the proposed model. Whereas conventional model performs less due to the improper feature selection and classification process.

Table 3 shows the performance comparative analysis of proposed concatenation model and conventional models in terms of accuracy, recall and precision. The average values from the results of dataset 1 and dataset 2 are presented in the tabulation. It can be observed from the results the performance of proposed concatenation model is much better than the conventional techniques. Thus, it is clear that the proposed concatenation model can be used for predicting disease in health cares to attain sustainable development.

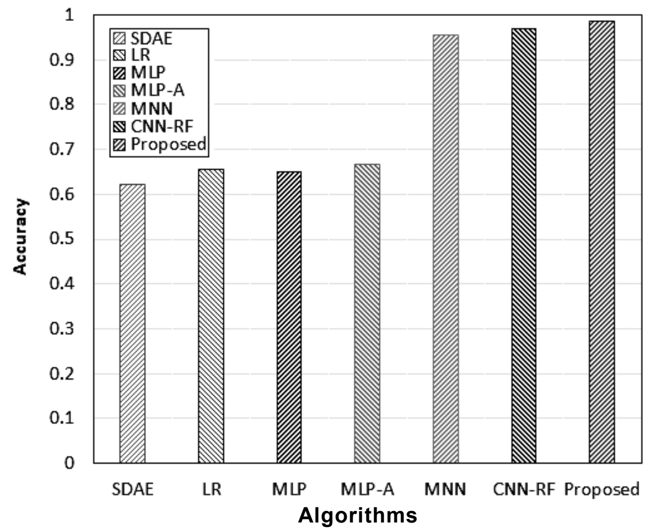


Fig. 8. Recall analysis for dataset 2

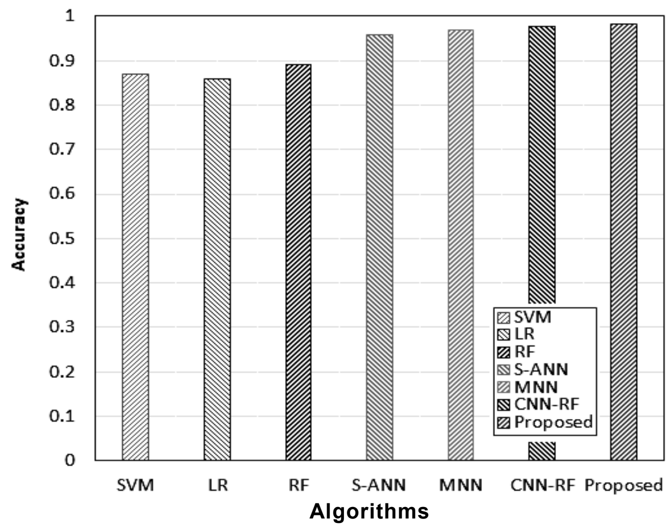


Fig. 9. Accuracy analysis for dataset 2

Table 3. Performance comparative analysis

No	Method	Accuracy	Precision	Recall
1	Stacked denoising auto-encoder (SDAE)	0.623	0.670	0.782
2	Logistic regression (LR)	0.655	0.700	0.792
3	MLP	0.651	0.692	0.799
4	MLP with attention mechanism	0.667	0.710	0.795
5	SVM	0.87	0.85	0.85
6	Logistic regression	0.86	0.84	0.85
7	Random forest	0.89	0.88	0.88
8	Swarm-ANN	0.957	0.952	0.952
9	MNN	0.966	0.962	0.97
10	CNN-RF	0.973	0.982	0.987
11	Proposed Inception V3 – Xception -SDC	0.985	0.988	0.992

CONCLUSION

A concatenation model for disease prediction in healthcare system is presented in this research work using InceptionV3-Xception model and Soft-max Discriminant Classifier. The proposed architecture utilizes the multi-features extracted from concatenation model and classify the data using Soft-max Discriminant Classifier. The novelty in the architecture enhances the classification performance of data analysis system compared to conventional CNN model. Standard healthcare datasets are used for experimentation and verified through performance metrics like accuracy, recall, precision and f1-score.To demonstrate the better performance, conventional techniques like stacked denoising auto-encoder (SDAE), logistic regression (LR), MLP, MLP with attention mechanism (MLP-A), support vector machine (SVM), Random Forest (RF), swarm artificial neural network (S-ANN) ,multi neural networks and Convolutional Neural Network-Random forest (CNN-RF) are compared with proposed concatenation model. Experimental results depicts that the performance of proposed model is much better than the conventional approaches. However, the performance of proposed concatenation model has several benefits, the prediction result is possible only for binary classes whether the data is normal or abnormal, that is considered as a minor limitation of this work. Further this research work can be extended using multi classification to identify the particular stage of the disease.

REFERENCES

1. Adekunle O. Afolabi and Pekka Toivanen, "Integration of Recommendation Systems into Connected Health for Effective Management of Chronic Diseases," *IEEE Access*, vol. 7, pp. 49201– 49211, 2019.
2. Haolin Wang , Zhilin Huang, Danfeng Zhang, Johan Arief, Tiewei Lyu, and Jie Tian,"Integrating Co-Clustering and Interpretable Machine Learning for the Prediction of Intravenous Immunoglobulin Resistance in Kawasaki Disease," *IEEE Access*, vol. 8, pp. 97064–97071, 2020.
3. Jian Ping Li, Amin Ul Haq, Salah Ud Din, Jalaluddin Khan, Asif Khan, and Abdus Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," *IEEE Access*, vol. 6, pp. 107562–107582, 2020.
4. Neelum Noreen, Sellappan Palaniappan, Abdul Qayyum, Iftikhar Ahmad, Muhammad Imran, and Muhammad Shoaib, "A Deep Learning Model Based on Concatenation Approach for the Diagnosis of Brain Tumor," *IEEE Access*, vol. 8, pp. 55135–55144, 2020.

5. Rajkumar Gangappa Nadakinamani et al., "Clinical Data Analysis for Prediction of Cardiovascular Disease Using Machine Learning Techniques," *Computational Intelligence and Neuroscience*, pp.1–13, 2022.
6. Qian Wang, Hong Wang, Lutong Wang, and Fengping Yu, "Diagnosis of Chronic Obstructive Pulmonary Disease Based on Transfer Learning," *IEEE Access*, vol. 8, pp. 47370–47383, 2020.
7. Lodewijk Brand, Kai Nichols, Hua Wang, Li Shen, and Heng Huang, "Joint Multi-Modal Longitudinal Regression and Classification for Alzheimer's Disease Prediction," *IEEE Transactions on Medical Imaging*, vol. 14, pp. 1–10, 2019.
8. Gazi Mohammed Ifraz, Muhammad Hasnath Rashid, Tahia Tazin, Sami Bourouis, and Mohammad Monirujjaman Khan, "Comparative Analysis for Prediction of Kidney Disease Using Intelligent Machine Learning Methods," *Computational and Mathematical Methods in Medicine*, pp.1–10, 2021.
9. Hong Qing Yu and Stephan Reiff-Marganiec, "Targeted Ensemble Machine Classification Approach for Supporting IoT Enabled Skin Disease Detection," *IEEE Access*, vol. 8, pp. 50244–50252, 2021.
10. Mohamed A. Hedeya, Ahmad H. Eid, and Rehab F. Abdel-Kader, "A Super-Learner Ensemble of Deep Networks for Vehicle-Type Classification," *IEEE Access*, vol. 8, pp. 98266–98280, 2020.
11. Cuiping Shi, Ruiyang Xia, and Liguo Wang, "A Novel Multi-Branch Channel Expansion Network for Garbage Image Classification," *IEEE Access*, vol. 8, pp. 154437–154452, 2020.
12. Aqsa Rahim, Yawar Rasheed, Farooque Azam, Muhammad Waseem Anwar, Muhammad Abdul Rahim, and Abdul Wahab Muzaffar, "An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases," *IEEE Access*, vol. 9, pp. 106575–106588, 2021.
13. Bin Liu, Cheng Tan, Shuqin Li, Jinrong He, and Hongyan Wang, "A Data Augmentation Method Based on Generative Adversarial Networks for Grape Leaf Disease Identification," *IEEE Access*, vol. 8, pp. 102188–102198, 2020.
14. Rohit Bharti, Aditya Khamparia, Mohammad Shabaz, Gaurav Dhiman, Sagar Pande, and Parneet Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning," Computational Intelligence and Neuroscience, pp. 1–11, 2021.
15. Muhammad Imran Razzak, Muhammad Imran, and Guandong Xu, "Efficient Brain Tumor Segmentation with Multiscale Two-Pathway-Group Conventional Neural Networks," *IEEE Journal of Biomedical and Health Informatics*, pp.1–10, 2018.
16. Hania H. Faraga, Lamiaa A.A. Saidb, Mohamed R.M. Rizka, and Magdy Abd ElAzim Ahmedc, "Hyperparameters optimization for ResNet and Xception in the purpose of diagnosingCOVID-19," *Journal of Intelligent & Fuzzy Systems*, pp. 3555–3571, 2021.
17. Prabira Kumar Sethya, Santi Kumari Beherab, Komma Anithac, Chanki Pandeyd, and M.R. Khand, "Computer aid screening of COVID-19 usingX-ray and CT scan images:An inner comparison," *Journal of X-Ray Science and Technology*, pp. 197–210.
18. Cheng Wang et al., "Pulmonary Image classification based on Inception-V3 Transfer Learning Model," *IEEE Access*, vol. 7, pp. 146533–146541, 2019.
19. Qingmao Zeng, Xinhui Ma, Baoping Cheng, Erxun Zhou, and Wei Pang, "GANs-Based Data Augmentation for Citrus Disease Severity Detection Using Deep Learning," *IEEE Access*, vol. 8, pp. 172882–172891, 2020.
20. Gaobo Liang, Huichao Hong, Weifang Xie, and Lixin Zheng, "Combining Convolutional Neural Network With Recursive Neural Network for Blood Cell Image Classification," *IEEE Access*, vol. 6, pp. 36188–36197, 2018.
21. Chen Huang , Xiaochen Wang, Jiannong Cao, Shihui Wang, and Yan Zhang, "HCF: A Hybrid CNN Framework for Behavior Detection of Distracted Drivers," *IEEE Access*, vol. 8, pp. 109335–109349, 2020.
22. Joel Parente De Oliveira, Marly Guimarães Fernandes Costa, and Cícero Costa Filho, "Methodology of Data Fusion Using Deep Learning for Semantic Segmentation of Land Types in the Amazon," *IEEE Access*, vol. 8, pp. 187864–187875, 2020.

23. Andrés Anaya-Isaza, Leonel Mera-Jiménez, Johan Manuel Cabrera-Chavarro, Lorena Guachi-Guachi, Diego Peluffo-Ordóñez, and Jorge Ivan Rios-Patiño, “Comparison of Current Deep Convolutional Neural Networks for the Segmentation of Breast Masses in Mammograms,” *IEEE Access*, vol. 9, pp. 152206–152225, 2021.
24. Sin-Ae Lee, Hyun Chin Cho, and Hyun-Chong Cho, “A Novel Approach for Increased Convolutional Neural Network Performance in Gastric-Cancer Classification Using Endoscopic Images,” *IEEE Access*, vol. 9, pp. 51847–51854, 2021.
25. Yan Zhao, Baoqiang Ma, Pengbo Jiang, Debin Zeng, Xuetong Wang, and Shuyu Li, “Prediction of Alzheimer’s Disease Progression with Multi-Information Generative Adversarial Network,” *IEEE Journal of Biomedical and Health Informatics*,vol. 25, pp. 711–719, 2020.
26. Sudarshan Nandy, Mainak Adhikari, Venki Balasubramanian, Varun G. Menon, Xingwang Li, and Muhammad Zakarya, “An intelligent heart disease prediction system based on swarm artificial neural network,” *Neural Computing and Applications*, pp. 1–15, 2021.

Received 27.11.2022

INFORMATION ON THE ARTICLE

K. Tharageswari, Karpagam Academy of Higher Education, Coimbatore, India, e-mail: ktharageswari4@gmail.com

Dr. N. Mohana Sundaram, Karpagam Academy of Higher Education, Coimbatore, India

Dr. R. Santhosh, Karpagam Academy of Higher Education, Coimbatore, India, e-mail: santhoshrd@gmail.com

МОДЕЛЬ ПРОГНОЗУВАННЯ ЗАХВОРЮВАННЯ НА ОСНОВІ ПІДХОДУ КОНКАТЕНАЦІЇ ДЛЯ СТІЙКОЇ СИСТЕМИ ОХОРОНИ ЗДОРОВ’Я / К. Тарагесварі, Н. Мохана Сундарам, Р. Сантош

Анотація. У сучасному світі внаслідок багатьох факторів, таких як зміни навколишнього середовища, стилі харчування та життєві звички, на здоров’я людей постійно впливають різні захворювання, що призводить до того, що в системі охорони здоров’я потрібно керувати величезною кількістю даних. Деякі захворювання створюють небезпеку для життя, якщо їх не вилікувати на початковій стадії. Для системи охорони здоров’я це робить складним завданням розробити добре навчену модель прогнозування захворювань для точної їх ідентифікації. Моделі глибокого навчання найбільш широко використовуються в дослідженнях прогнозування захворювань, але їх продуктивність поступається звичайним моделям. Щоб вирішити цю проблему, у роботі подано конкатенацію моделей згорткових нейронних мереж глибокого навчання Inception V3 і Xception. Запропонована модель виділяє основні ознаки та створює результат прогнозу точніше, ніж інші традиційні моделі прогнозування. У роботі аналізується продуктивність запропонованої моделі з точки зору точності, прецизійності, запам’ятовування та F1-міра, порівнюються моделі з існуючими методами, такими як стековий автоматичний кодувальник (SDAE), логістична регресія (LR), MLP, MLP з механізмом уваги (MLP-A), опорна векторна машина (SVM), мультинейронна мережа (MNN), гібридна згорткова нейронна мережа (CNN), випадковий ліс (RF).

Ключові слова: вилучення функцій, прогнозування захворювань, глибоке навчання, Inception V3, Xception.

UDC 62-50
DOI: 10.20535/SRIT.2308-8893.2023.3.07

IDENTIFICATION OF LUNG DISEASE TYPES USING CONVOLUTIONAL NEURAL NETWORK AND VGG-16 ARCHITECTURE

S. BUKHORI, B.Y.N. VERDY, Y.R. WINDI EKA, A.P. JANUAR

Abstract. Pneumonia, tuberculosis, and Covid-19 are different lung diseases but have similar characteristics. One of the reasons for the worsening of disease in lung sufferers is a diagnosis that takes a long time. Another factor, the results of the X-ray photos look blurry and lack contracture, causing different diagnostic results of X-ray photos. This research classifies lung images into four categories: normal lungs, tuberculosis, pneumonia, and Covid-19 using the Convolutional Neural Network method and VGG-16 architecture. The results of the research with models and scenarios without pre-trained use data with a ratio of 9:1 at epoch 50, anaccuracy of 94%, while the lowest results are in scenarios using data with a ratio of 8:2 at epoch 50, non-pre-trained models, accuracy by 87%.

Keywords: tuberculosis, pneumonia, Covid-19, VGG-16, convolutional neural network.

INTRODUCTION

Human internal organs that are often associated with external environmental factors, one of which is the lungs [1; 2]. The human lungs consist of two organs or a pair, namely the right and left. The lungs are located in the thoracic region of the human body, one of two large respiratory organs located in the chest cavityand are responsible for adding oxygen and removing carbon dioxide from the blood. The lungs have a rubbery texture and are pinkish-gray in colour. The lungs consist of other tissues inside which function to exchange oxygen and carbon dioxide [3]. Because of the process of exchanging oxygen and carbon dioxide, the lungs are in contact with external environmental factors such as smoke, microbes, dust and also chemicals in the environment as pollutants. The relationship with these environmental factors increases the risk of lung disease [4].

Several diseases can attack the lungs. Common symptoms are shortness of breath and coughing. Lung disorders can be acute or chronic. Several diseases that can attack the lungs and related respiratory systems include bronchitis, pneumonia, asthma, tuberculosis and Covid-19 [5]. Bronchitis is a respiratory disease that occurs as a result of an upper respiratory infection and is usually caused by a virus [6]. Pneumonia is a respiratory disorder that causes inflammation of the smallest parts of the lungs, namely the bronchioles and alveolar tissue. Asthma is a disease caused by inflammation of the respiratory tract. This inflammation will cause swelling and narrowing of the airways. Air that should flow into the lungs becomes obstructed [7]. Tuberculosis is a bacterial infection caused by Mycobacterium tuberculosis which attacks and damages body tissues. Bacteria can be transmitted through the airways. Tuberculosis generally attacks the lungs, but also has the risk of spreading to the lymph nodes, bones, central nervous system, heart and other organs [8]. Covid-19 is an infectious disease caused by SARS-CoV-2, a type of coronavirus [9]. Typical symptoms are fever, cough, flu and shortness of breath. Covid-19 spreads from one person to another through droplets from the respiratory tract which are often produced when coughing or sneezing. Droplet range is usually up to 1 meter [10]. Droplets can stick to objects, but won't last long in the air. The time from transmission of the virus to the onset of clinical symptoms is between 1–14 days with an average of 5 days. Bronchitis, pneumonia, asthma, tuberculosis and Covid-19 if not handled properly in a short time can cause health complications [11].

Lung disease problems tend to increase due to delays in diagnosis. Diagnosis takes a long time because of the similarities in the symptoms of lung disease. According to WHO, various lung diseases including pneumonia, tuberculosis and Covid-19 have almost the same symptoms [12]. One of the main reasons for the increase in lung disease problems during the Covid-19 pandemic is the long process of diagnosis. Another factor is that X-Rays often appear blurry and have no contractures, leading to a different diagnosis [13]. Additional laboratory test results are needed to identify whether it is classified as tuberculosis, pneumonia, or Covid-19. One of the reasons for the unfavourable radiographic results is the difference in X-Ray intensity in photos of normal tissue and photos of glandular tissue affected by lung disease [14]. To overcome this problem, image processing is needed so that it can increase and improve image quality. A lung disease classification system is needed to help diagnose lung disease quickly. Alternative technologies that can be used to overcome this problem are the use of computer vision and deep learning.

Computer vision is one part of artificial intelligence [15]. Computer Vision is a technology that allows computers to recognize objects around them [16]. Science and technology are developing very fast, especially in the development of computer vision combined with deep learning. Deep learning can be used for decision making, detecting diseases based on their symptoms and early detection [17]. This research develops identification of lung disease types using Convolutional Neural Network and VGG-16 architecture.

Several research have identified lung disease from chest X-Rays using small volume datasets and applying machine learning [15; 18]. The results of applying this technology are quite important for medical progress. Research related to the early diagnosis and treatment of lung diseases using deep learning has also been researched and the results are quite important in clinical treatment [19; 20; 21]. This research develops lung disease identification using Convolutional Neural Network (CNN) and VGG-16 architecture. CNN with VGG-16 architecture has higher accuracy than other network architectures in processing ImageNet datasets [22]. In similar cases, classification of pneumonia from X-Ray images using VGG-16 results 97.93 % accuracy, while classification of pneumonia from X-Ray images using inception-V3 results 96.58% accuracy [23]. The inception-V3 is the latest version developed from the inception-V1 and V2 models. The inception-V3 model is a CNN trained directly on a low-configuration computer. The training is quite difficult, and takes much longer. This problem is solved through transfer learning which saves the last layer of the model for the new category. The parameters of the previous layer are stored, and the inception-V3 model is deconstructed when the last layer is removed using transfer learning techniques. CNNs inception is a network in the form of a repeated convolution design configuration pattern. The components in CNNs inception are input layer, 1×1 convolution layer, 3×3 convolution layer, 5×5 convolution layer, max pooling layer, and concatenation layer [24].

Test results on the diagnosis of pneumonia showed that VGG-16 architecture exceed Xception network at the accuracy with 87% and 82% respectively [25]. Xception is a development of inception. The inception model was developed with depth wise separable convolution. The number of parameters is almost the same as inception. Xception brings the inception hypothesis to eXtreme. First, the cross-feature map is captured by a 1×1 convolution. Consequently, the correlation of each channel is captured via a regular 3×3 or 5×5 convolution. This idea goes to the extreme of doing 1×1 to each channel, then doing 3×3 to each output [26]. This is identical to replacing the inception module with depth wise separable convolutions. With a higher level of accuracy, the VGG-16 architecture can increase the value of lung disease classification based on CT-scan images. This research also developed a classification of lung images into four classes, namely normal lungs, tuberculosis, pneumonia, and Covid-19.

The rest of this paper is organized as follows: The proposed model for identification of lung disease types using CNN and VGG-16 architecture is discussed in section 2. Section 3 provides a system design for identification of lung disease types using CNN and VGG-16 architecture. Section 4 discusses the results and analysis of CNN model and integrated system applications. Finally, conclusions are given in section 5.

RESEARCH PROPOSED

CNN architecture developed using VGG-16. The proposed architecture has 13 convolution layers and 3 fully connected layers so that a total of 16 layers are used, there are 5 max-polling layers which are forwarded to several convolution layers. Each layer has a different image size, on the first layer the image will be resized to 224×224, then on the next layer the image size will be reduced to 112×112, 56×56, 28×28, 14×14 as shown in Fig. 1.

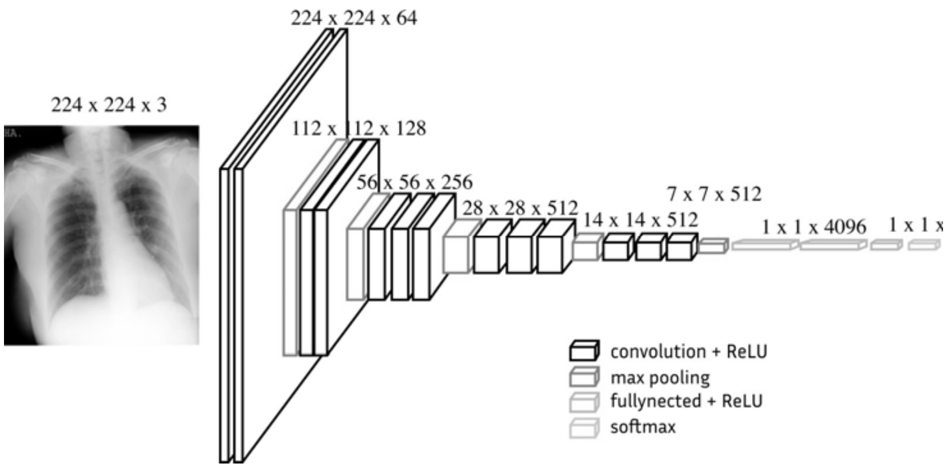


Fig. 1. The Proposed Model

Fig. 1 shows that the first layer is a fully connected layer which has 4096 neurons, the second layer has 4096 neurons, and the third layer has 4 neurons. In the last layer there are 4 layers according to the number of classifications, namely normal, pneumonia, tuberculosis, and Covid-19. There are several layers in the CNN architecture (Fig. 1) which are used in the training and testing process. This architecture is designed to produce models with better accuracy.

Input image is the process of entering an input file in the form of a chest X-Ray image. Before the data is entered into the training process, to prevent errors in classification, the dataset is adjusted specifically for biased data or truncated data. Inappropriate data will be discarded and appropriate data will be included in a dataset that is ready to be classified. This process includes 3 stages, namely changing the size of the photo with a target of 224×224 pixels, the photo data is separated into two parts, namely training data and testing data with a comparison ratio of 7 : 3, 8 : 2, and 9 : 1. Then the photo is converted from RGB to grayscale.

The convolutional layer includes filter, kernel, stride, Relu activation, and poolingoperations. The convolution layer used in this research uses a kernel with a size of 3×3 pixels and the number of strides is 1, and uses a padding configuration. There are thirteen convolution layers, the first convolution layer uses 64 filters,the kernel is 3×3 pixels. The size of the chest X-Ray is 224×224 pixels which then uses the ReLu activation function. The second convolution layer is similar to the first convolution process but continues with a pooling operation with a strides size of 2×2 and a pooling size of 2×2. The pooling operation will produce an image measuring 112×112 pixels and a total of 64 feature maps. This feature map will be included into the third convolution process. The third convlution layer has different parameters from the previous layer. The filter used in the third layer is 128 with a data input size of 112×112 pixels. With a total of 128 filters, the third convolution feature map obtained is 128. The fourth convolution layer is similar to the third convolution process, but the pooling operation is continued with a strides size of 2×2 and a pooling size of 2×2. The pooling operation will produce an image measuring 56 × 56 pixels and a total of 128 feature maps. This feature map will be included in the in the fifth convolution process. The fifth convolution layer has different parameters from the fourth layer. There are 256 filters used in this fifth layer with an input data size of 56 × 56 pixels. With a total of 256 filters, the convolution of the five feature maps obtained is 256. The sixth convolution layer is similar to the fifth convolution layer without changing any parameters. The filters used in the sixth layer are 256 with a data input size of 56 × 56 pixels. The seventh convolution layer has similarities with the fifth and sixth convolution layers without changing any parameters. The filters used in the seventh layer are 256 with a data input size of 56×56 pixels. In this layer, pooling operations are continued with a strides size of 2×2 and a pooling size of 2×2. The pooling operation will produce an image measuring 28×28 pixels anda total of 256 feature maps. The eighth convolution layer has different parameters from the previous layer. The filter used in this eighth layer is 512 with a data input size of 28×28 pixels. Because there are 512 filters, the eighth convolution produces 512 feature maps. The ninth convolution layer has the same parameters as the eighth convolution layer. The filter used in the ninth layer is 512 and the input data is 28×28 pixels. Filters with a total of 512 in the ninth convolution produce a feature map totaling 512. The tenth convolution layer has the same parameters as the eighth and ninth convolution layers. The filter used in the tenth layer is 512 and the input data is 28×28 pixels. This process is followed by a pooling operation with strides size of 2×2 and a pooling size of 2×2. Thepooling operation will produce an image size of 14×14 pixels and a total of 512 feature maps. The eleventh convolution layer has similarities with the tenth layer, but has a different input data size. The filter used in the eleventh layer is 512 with a data input size of 14×14 pixels. There are 512 filters and in the eleventh convolution, the resulting feature maps are 512. The twelfth convolution layer has the same parameters as the eleventh convolution layer. The filter used is 512 with a data input size of 14×14 pixels. These 512 filters will produce a feature map of 512. The thirteenth convolution layer has the same parameters as the eleventh and twelfth layers. The filter used is 512 with input data of 14×14 pixels. Then the process is continued with a pooling operation with a size of 2×2 strides and a pooling size of 2×2. The pooling operation will produce an image measuring c pixels and a total of 512 feature maps.

The flatten layer is the layer that converts the multidimensional array output in the feature extraction process into a one-dimensional matrix which is then followed by the fully connected layer process. While the fully connected layer will be used as many as three fully connected layers. The first layer is 4096, the second layer is 4096, and the last layer is 4 according to the classification designed.

SYSTEM DESIGN

The system is designed using the CNN architecture with 16 layers according to the VGG-16 architectural concept. The architectural design uses 13 convolution layers and 3 fully connected layers so that the total layers used are 16 layers. This research uses 5 layers of max-polling, which adjusts to several convolution layers. Each layer has a different image size. The first layer will resize the image to 224×224, then the next layer will reduce the size of the image to a configuration of 112×112, 56×56, 28×28, 14×14, and 7×7, as shown in Table 1.

Table 1. Configuration VGG-16

No.	Layer	Output Shape
1	Conv2D	224, 224, 64
2	Conv2D	224, 224, 64
3	MaxPooling2D	112, 112, 64
4	Conv2D	112, 112, 128
5	Conv2D	112, 112, 128
6	MaxPooling2D	56, 56, 128
7	Conv2D	56, 56, 256
8	Conv2D	56, 56, 256
9	Conv2D	56, 56, 256
10	MaxPooling2D	28, 28, 256
11	Conv2D	28, 28, 512
12	Conv2D	28, 28, 512
13	Conv2D	28, 28, 512
14	MaxPooling2D	14, 14, 512
15	Conv2D	14, 14, 512
16	Conv2D	14, 14, 512
17	Conv2D	14, 14, 512
18	MaxPooling2D	7, 7, 512
19	Flatten	25088
20	Dense	4096
21	Dense	4096
22	Dense	4

After modeling, it is continued with testing of the CNN model that has been designed to analyze the accuracy of the model by changing the data scenario, epoch and using pre-trained models. This process aims to select a model that has the highest accuracy so that it can be used for classification. Before comparing accuracy, the model will conduct data training and data testing with a ratio of 7 : 3, 8 : 2 and 9 : 1.

The test was carried out by changing the parameters to get the best image results for the classification of normal lungs, lungs with tuberculosis, lungs with pneumonia, and lungs with Covid-19. This research was conducted using three different epochs, namely: epoch 20, epoch 50, and epoch 100.While the data scenario uses three different data scenarios from each epoch that are used, namely scenario data 7 : 3, 8: 2 and 9: 1, and uses a comparison between using the pre- training and without using the pre-training model. Then the model will be compared for its level of accuracy and will be selected through the model with thehighest accuracy.

RESULTS AND ANALYSIS

The model that has been designed is tested to get the best scenario. Scenarios are made based on the number of epochs and the amount of data. The number of epochs tested were 20, 50, and 100, while the amount of data used was a ratio of 7 : 3, 8 : 2, and 9 : 1. Model testing was also carried out on models with pre-training and models without training.

Each scenario has a different value when using a different number of epochs, even though the model used is the same. Comparison of the performance of the CNN model as a whole and the test results are shown in Table 2 – Table 4 for the non-pretrained model test cases; Table 5 – Table 7 for the test cases with the pretrained model.

Table 2. Performance comparison with data ratio of 7 : 3 (non-pretrained)

Performance	***Epoch***		
	20	50	100
Data Scenario 7 : 3			
Accuracy	87.33%	89%	90%
Precision	89%	90%	90%
Recall	87%	89%	90%
F1 score	87%	89%	90%
Learning Curve	overfitting	overfitting	overfitting

The best performance on scenario data with a ratio of 7 : 3 (non–pretrained) is at epoch 100 with an accuracy value of 90%, a precision value of 90%, a recall valueof 90%, and a F1 score value of 90%. The results of the learning curves in this scenario have a high variance as shown in Fig. 2. The occurrence of a high variance indicates that there has been overfitting.

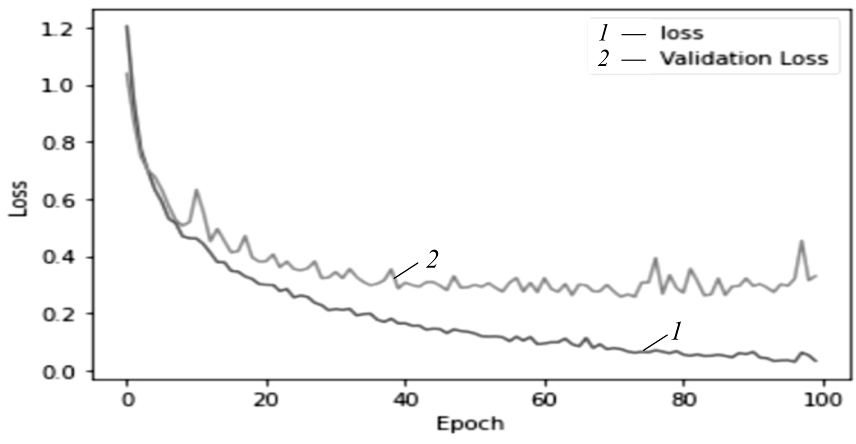


Fig. 2. Learning Curve Data Scenario 7 : 3 Epoch 100

Table 3. Performance comparison with data ratio of 8 : 2 (non-pretrained)

Performance	***Epoch***		
	20	50	100
Data Scenario 8 : 2			
Accuracy	87%	90.25%	92%
Precision	88%	90%	92%
Recall	87%	90%	92%
F1 score	87%	90%	92%
Learning Curve	overfitting	overfitting	overfitting

The best performance on scenario data with a ratio of 8 : 2 (non-pretrained) is at epoch 100 with an accuracy value of 92%, a precision value of 92%, a recall valueof 92%, and a F1 score value of 92%. The results of the learning curves in this scenario have a high variance as shown in Fig. 3. The occurrence of a high variance indicates that there has been overfitting.

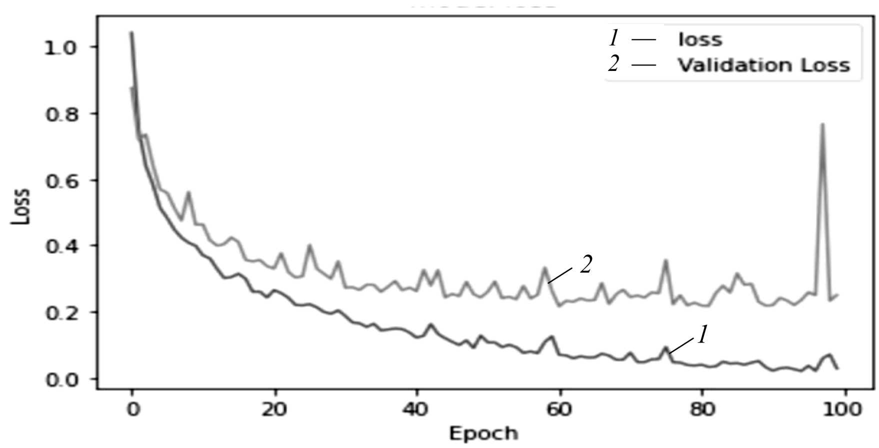


Fig. 3. Learning Curve Data Scenario 8 : 2 Epoch 100

Table 4. Performance comparison with data ratio of 9 : 1 (non-pretrained)

Performance	***Epoch***		
	20	50	100
Data Scenario 9 : 1			
Accuracy	89.50%	94%	92%
Precision	89%	94%	91%
Recall	89%	94%	93%
F1 score	90%	94%	92%
Learning Curve	overfitting	overfitting	overfitting

The best performance on scenario data with a ratio of 9 : 1 (non-pretrained) is at epoch 50 with an accuracy value of 94%, a precision value of 94%, a recall value of 94%, and a F1 score value of 94%. The results of the learning curves in this scenario have a high variance as shown in Fig. 4. The occurrence of a high variance indicates that there has been overfitting.

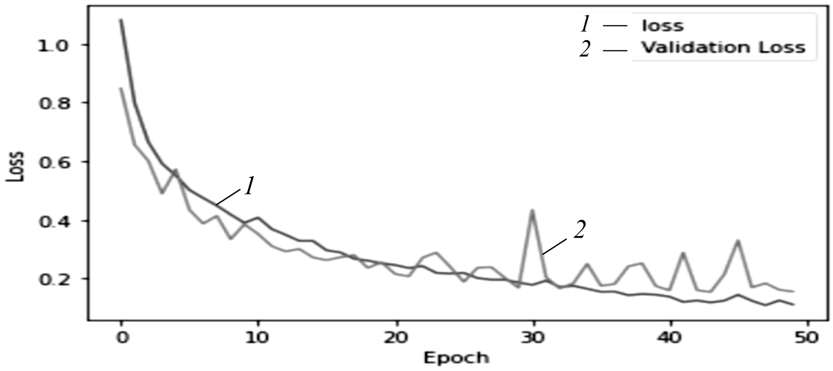


Fig. 4. Learning Curve Data Scenario 9 : 1 Epoch 50

Table 5. Performance comparison with data ratio of 7 : 3 (pretrained)

Performance	***Epoch***		
	20	50	100
Data Scenario 7 : 3			
Accuracy	91%	91.17%	91.33%
Precision	91%	92%	92%
Recall	91%	91%	91%
F1 score	91%	91%	91%
Learning Curve	overfitting	overfitting	overfitting

The best performance on scenario data with a ratio of 7 : 3 (pretrained) is at epoch 100 with an accuracy value of 91.33%, a precision value of 92%, a recall value of 91%, and a F1 score value of 91%. The results of the learning curves in this scenario have a high variance as shown in Fig. 5. The occurrence of a high variance indicates that there has been overfitting.

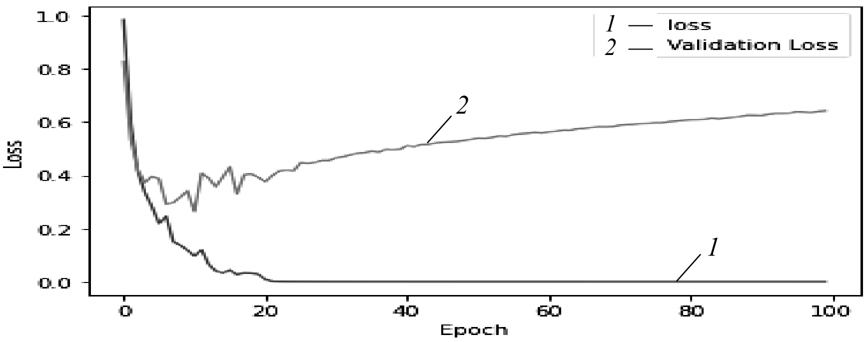


Fig. 5. Learning Curve Data Scenario 7 : 3 Epoch 100

Table 6. Performance comparison with data ratio of 8 : 2 (pretrained)

Performance	***Epoch***		
	20	50	100
Data Scenario 8 : 2			
Accuracy	91%	93.75%	92.50%
Precision	92%	94%	92%
Recall	91%	94%	92%
F1 score	91%	94%	93%
Learning Curve	overfitting	overfitting	overfitting

The best performance on scenario data with a ratio of 8 : 2 (pretrained) is at epoch 50 with an accuracy value of 93.75%, a precision value of 94%, a recall value of 94%, and a F1 score value of 94%. The results of the learning curves in this scenario have a high variance as shown in Fig. 6. The occurrence of a high variance indicates that there has been overfitting.

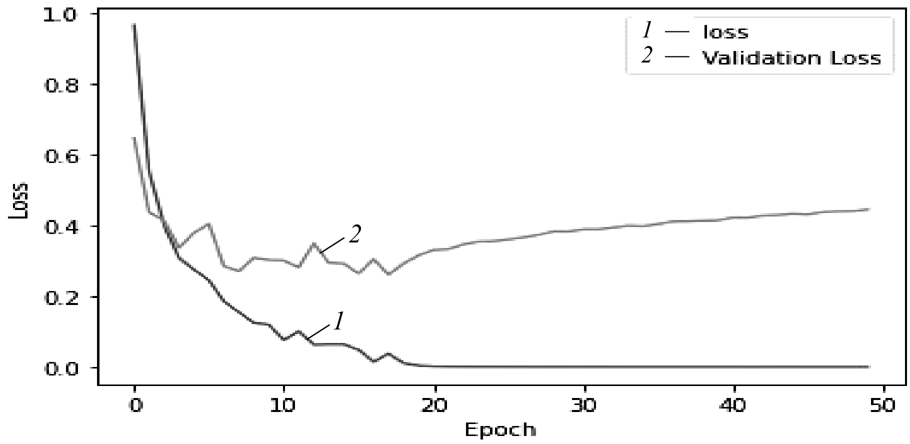


Fig. 6. Learning Curve Data Scenario 8 : 2 Epoch 50

Table 7. Performance comparison with data ratio of 9 : 1 (pretrained)

Performance	***Epoch***		
	20	50	100
Data Scenario 9 : 1			
Accuracy	94%	95%	94.50%
Precision	94%	96%	93%
Recall	94%	95%	94%
F1 score	94%	95%	94%
Learning Curve	overfitting	overfitting	overfitting

The best performance on scenario data with a ratio of 9 : 1 (pretrained) is at epoch 50 with an accuracy value of 95%, a precision value of 96%, a recall value of 95%, and a F1 score value of 95%. The results of the learning curves in this scenario have a high variance as shown in Fig. 7. The occurrence of a high variance indicates that there has been overfitting.

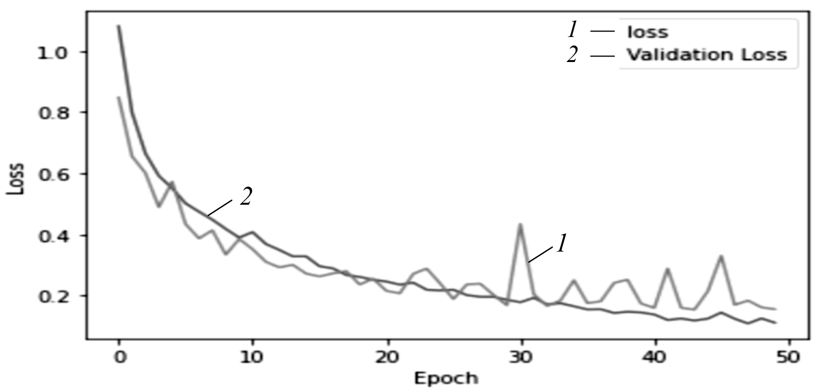


Fig. 7. Learning Curve Data Scenario 9 : 1 Epoch 50

CONCLUSIONS

Convolutional Neural Network classification model is tested by training the model using training data and testing the accuracy of the model with testing data. The better the model, the better the accuracy obtained. Testing results show that the CNN model has the highest accuracy in the scenario of non-pretrained data model 9 : 1 at epoch 50 which is 94%, while the lowest result is at 8 : 2 epoch 50 scenario testing of data without using a pre-trained model with an accuracy of 87%.

The model from this scenario will be used to be implemented into applications because it has the best accuracy among models with other scenarios. The pulmonary infectious disease classification system classifies pulmonary infectious diseases based on chest X-Rays uploaded by doctors/health workers with good accuracy. This system can classify into four classes: Normal, Covid-19, Pneumonia, and Tuberculosis. Pretraining models, epochs, and data scenarios can impact model performance. This research has conducted tests to determine the best performance achieved by using the VGG-16 architecture. The best performance is obtained in the 9 : 1 data scenario, epoch 50 on the non pre-trainedmodel, with an accuracy value of 94%, precision value of 94%, recall value of 94%, and F1-score value of 94%, while the lowest result is in the 8 : 2 data scenario test epoch 50 on the non-pretrained model with an accuracy value of 87%, precision value of 88%, recall value of 87%, and F1-score value of 87%.

REFERENCES

1. K.F. Budden et al., “Emerging pathogenic links between microbiota and the gut-lung axis,” *Nat. Rev. Microbiol.*, vol. 15, no. 1, pp. 55–63, 2017. doi: 10.1038/nrmicro.2016.142.
2. P.K. Wang L., F.H. Green, and S.M. Smiley-Jewell, “乳鼠心肌提取 HHS Public Access,”*Susceptibility aging lung to Environ. Inj.*, vol. 176, no. 1, pp. 539–553, 2010. doi: 10.1055/s-0030-1265895.
3. D.J. Randall, “Transport and exchange of respiratory gases in the blood | Carbon Dioxide Transport and Excretion,” *Encycl. Fish Physiol.*, vol. 2, pp. 909–915, Dec. 2011. doi: 10.1016/B978-0-12-374553-8.00027-7.
4. F. Yu, R. Xiao, X. Li, Z. Hu, L. Cai, and F. He, “Combined effects of lung disease history, environmental exposures, and family history of lung cancer to susceptibility of lung cancer in Chinese non-smokers,” *Respir. Res.*, vol. 22, no. 1, pp. 1–10, 2021. doi: 10.1186/s12931-021-01802-z.
5. N. Clementi et al., “Viral respiratory pathogens and lung injury,” *Clin. Microbiol. Rev.*, vol. 34, no. 3, pp. 1–45, 2021. doi: 10.1128/CMR.00103-20.
6. U. Grote, A. Fasse, T.T. Nguyen, and O. Erenstein, “Food Security and the Dynamics of Wheat and Maize Value Chains in Africa and Asia,” *Front. Sustain. Food Syst.*, vol. 4, no. February, pp. 1–17, 2021. doi: 10.3389/fsufs.2020.617009.
7. L.J. Quinton, A.J. Walkey, and J.P. Mizgerd, “Integrative physiology of pneumonia,” *Physiol. Rev.*, vol. 98, no. 3, pp. 1417–1464, 2018. doi: 10.1152/PHYSREV.00032.2017.
8. R. Gopalaswamy, S. Shanmugam, R. Mondal, and S. Subbian, “Of tuberculosis and non-tuberculous mycobacterial infections - A comparative analysis of epidemiology, diagnosis and treatment,” *J. Biomed. Sci.*, vol. 27, no. 1, pp. 1–17, 2020. doi: 10.1186/s12929-020-00667-6.
9. M.T. Adil et al., “SARS-CoV-2 and the pandemic of COVID-19,” *Postgrad. Med. J.*, vol. 97, no. 1144, pp. 110–116, 2021. doi: 10.1136/postgradmedj-2020-138386.

10. L. Morawska and J. Cao, "Airborne transmission of SARS-CoV-2: The world should face the reality," *Environ. Int.*, vol. 139, no. April, p. 105730, 2020. doi: 10.1016/j.envint.2020.105730.
11. M. Wei, Y. Zhao, Z. Qian, B. Yang, J. Xi, and J. Wei, "Pneumonia caused by Mycobacterium tuberculosis," *Microbes Infect.*, no. January, 2020.
12. "Clinical management of severe acute respiratory infection (SARI) when COVID-19 disease is suspected," *World Health Organization*, 2020. [Online]. Available: https://www.ptonline.com/articles/how-to-get-better-mfi-results
13. R. Yasin and W. Gouda, "Chest X-ray findings monitoring COVID-19 disease course and severity," *Egypt. J. Radiol. Nucl. Med.*, vol. 51, no. 1, 2020. doi: 10.1186/s43055-020-00296-x.
14. A.M. O'Hare et al., "Complexity and Challenges of the Clinical Diagnosis and Management of Long COVID," *JAMA Netw. Open*, vol. 5, no. 11, p. e2240332, 2022. doi: 10.1001/jamanetworkopen.2022.40332.
15. A.I. Khan and S. Al-Habsi, "Machine Learning in Computer Vision," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 1444–1451, 2020. doi: 10.1016/j.procs.2020.03.355.
16. S. Wael, A. Elshater, and S. Afifi, "Mapping User Experiences around Transit Stops Using Computer Vision Technology: Action Priorities from Cairo," *Sustain.*, vol. 14, no. 17, 2022. doi: 10.3390/su141711008.
17. S. Arooj, S.U. Rehman, A. Imran, A. Almuhaimeed, A.K. Alzahrani, and A. Alzahrani, "A Deep Convolutional Neural Network for the Early Detection of Heart Disease," *Biomedicines*, vol. 10, no. 11, p. 2796, 2022. doi: 10.3390/biomedicines10112796.
18. Y. Erdaw and E. Tachbele, "Machine learning model applied on chest X-ray images enables automatic detection of COVID-19 cases with high accuracy," *Int. J. Gen. Med.*, vol. 14, pp. 4923–4931, 2021, doi: 10.2147/IJGM.S325609.
19. T.N. Sree Ganesh, R. Satish, and R. Sridhar, "Learning effective embedding for automated COVID-19 prediction from chest X-ray images," *Multimed. Syst.*, no. 0123456789, 2022. doi: 10.1007/s00530-022-01015-4.
20. S.M. Fati, E.M. Senan, and N. ElHakim, "Deep and Hybrid Learning Technique for Early Detection of Tuberculosis Based on X-ray Images Using Feature Fusion," *Appl. Sci.*, vol. 12, no. 14, 2022. doi: 10.3390/app12147092.
21. V. Acharya et al., "AI-Assisted Tuberculosis Detection and Classification from Chest X-Rays Using a Deep Learning Normalization-Free Network Model," *Comput. Intell. Neurosci.*, vol. 2022, p. 2399428, 2022. doi: 10.1155/2022/2399428.
22. H. Handoko, F. Rozy, F.E. Gani, and A. Dharma, "Comparative Analysis of Convolutional Neural Network Methods in Detecting Mask Wear," *Budapest International Research and Critics Institute-Journal (BIRCI-Journal*), 5(2), 2022.
23. M. Mujahid, F. Rustam, R. Álvarez, J. Luis Vidal Mazón, I. de la T. Díez, and I. Ashraf, "Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network," *Diagnostics*, vol. 12, no. 5, pp. 1–16, 2022. doi: 10.3390/diagnostics12051280.
24. G. Natarajan and P. Dhanalakshmi, "Classification of pneumonia using pre-trained convolutional networks on chest X-Ray images," *Int. J. Health Sci. (Qassim)*, vol. 6, no. April, pp. 5378–5390, 2022. doi: 10.53730/ijhs.v6ns1.6097.
25. E. Ayan and H.M. Ünver, "Diagnosis of pneumonia from chest X-ray images using deep learning," *2019 Sci. Meet. Electr. Biomed. Eng. Comput. Sci. EBBT 2019*, pp. 10–13, 2019. doi: 10.1109/EBBT.2019.8741582.
26. Y.-J. Fan et al., "Machine Learning: Using Xception, a Deep Convolutional Neural Network Architecture, to Implement Pectus Excavatum Diagnostic Tool from Frontal-View Chest X-rays," *Biomedicines*, vol. 11, no. 3, p. 760, 2023. doi: 10.3390/biomedicines11030760.

Received 05.01.2023

INFORMATION ON THE ARTICLE

Saiful Bukhori, ORCID: 0000-0002-2527-1080, University of Jember, Indonesia, e-mail: saiful.ilkom@unej.ac.id

Verdy Bangkit Yudho Negoro, University of Jember, Indonesia, e-mail: verdybangkit@gmail.com

Windi Eka Yulia Retnani, ORCID: 0009-0001-7838-0205, University of Jember, Indonesia, e-mail: windi.ilkom@unej.ac.id

Januar Adi Putra, University of Jember, Indonesia, e-mail: januaradi.putra@unej.ac.id

ІДЕНТИФІКАЦІЯ ТИПІВ ЗАХВОРЮВАННЯ ЛЕГЕНЬ ЗА ДОПОМОГОЮ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ Й АРХІТЕКТУРИ VGG-16 / Сайфул Бухорі, Верді Бангкіт Юдхо Негоро, Вінді Ека Юліа Ретнані, Януар Аді Путра

Анотація. Пневмонія, туберкульоз і Covid-19 – різні захворювання легенів, але мають схожі характеристики. Однією з причин загострення захворювання легень є довготривала діагностика. Іншим фактором є те, що результати рентгенівських знімків виглядають розмитими і з відсутністю контрактури, що спричиняє різні результати діагностики рентгенівських знімків. Це дослідження класифікує зображення легенів на чотири категорії, а саме: нормальні легені, туберкульоз, пневмонія та Covid-19 за допомогою методу згорткової нейронної мережі та архітектури VGG-16. Результати дослідження з моделями та сценаріями без попередньої підготовки використовують дані зі співвідношенням 9:1 в епосі 50, точністю 94%, тоді як найнижчі результати в сценаріях з використанням даних зі співвідношенням 8:2 в епосі 50, моделі без попередньої підготовки, точність 87%.

Ключові слова: туберкульоз, пневмонія, Covid-19, VGG-16, згорткова нейронна мережа.

UDC 004.94
DOI: 10.20535/SRIT.2308-8893.2023.3.08

INTELLIGENT INFORMATION SYSTEM OF THE CITY'S SOCIO-ECONOMIC INFRASTRUCTURE

KHRYSTYNA LIPIANINA-HONCHARENKO, YEVGENIY BODYANSKIY, ANATOLIY SACHENKO

Abstract. Urban development is an important problem that can be solved with the help of intelligent information systems. Such systems ensure efficient management of the city's diverse infrastructure. The researchers developed a concept of such an information system based on a conceptual model and using data flow for intelligent decision-making. The system was tested for 1460 days in the city of Ternopil. The modelling results showed that the city's central area is stable, with 50% of enterprises in the "growing" state and 70% of people in the "satisfactory" state. People often move to the northeastern and western zones due to higher levels of comfort and more affordable housing. However, the total distance of car trips has increased by 249%, negatively impacting the environment. The condition of enterprises in other zones is less stable with lower "growth" indicators, but there are zones with "stable" and "satisfactory" conditions.

Keywords: modelling, information system, socioeconomic infrastructure, city.

INTRODUCTION

In recent years, studies of economic infrastructure at various levels of the country's socio-economic system have been actively conducted. These studies have led to the synthesis and systematization of knowledge about the economic infrastructure of states, regions and enterprises. This allowed us to reveal its essence and explain the patterns and cause-and-effect relationships at different levels of the economic system.

Recognition of the importance of the social component in the economic infrastructure of the state and regions led to the introduction of the category "socio-economic infrastructure". This is because the growth of activity and efficiency of economic entities is not the main goal, but should contribute to improving the welfare of the population through increased wages, social assistance and improved quality of social services.

Socio-economic infrastructure is a prerequisite for the stability and efficiency of the socio-economic system at any level. In particular, cities play an important role in the country's economy due to their industrial, scientific and technical potential, financial and commodity markets, and the formation of decisions that determine the vectors of socio-economic development of regions and the state.

Therefore, it is necessary to develop an intelligent information system of the city's socio-economic infrastructure to automate optimal decisions on the development of the city's socio-economic infrastructure. In addition, such a system will provide an opportunity to monitor the state of the city's infrastructure online, respond quickly to problems and ensure their effective resolution.

ANALYSIS OF LITERATURE SOURCES

Article [1] explains the lexical and economic meaning of the term economic security. To assess [2] the level of economic infrastructure, an integrated approach is proposed, which consists in combining a number of indicators into a single integral indicator that summarizes data on the level of sustainable development of an enterprise, which allows it to be used both in operational management and in strategy. A resource-functional model of security (consisting of partial indicators and components of economic security of business) was developed [3] and a resource-functional approach to calculations was also applied. A quantitative assessment of the level of financial solvency of countries based on the use of multidimensional methodological tools for assessing financial indicators of the country's development was carried out [4], which leads to the construction of appropriate integral security indices. Unlike other methods of assessing the level of security, the proposed approach makes it possible to determine not only the integrated level of the financial component of economic security, but also to calculate the quantitative thresholds of financial indicators aggregated in the integral index (foreign exchange reserves, external debt per capita, changes in the official exchange rate of the local currency, budget deficit/surplus to GDP); going beyond the thresholds is a signal of increased risk and lack of solvency.

The article [5] investigates the problems of organizing an intelligent system for managing complex socio-economic processes, defines its levels of intellectual development, proposes stages of intellectualization, and demonstrates the effectiveness of applying these solutions in practical tasks.

The article [6] assessed user satisfaction with the electronic social security system (SSES) as a widely used system in Iran. In [7], the Hans-Böckler-Stiftung and its research unit "Future Jobs" present a revised plan for Enzo Weber's DSS model. DSS addresses the problem of serious gaps in social security for platform workers. The model envisages that platforms around the world implement a digital mechanism to transfer a certain proportion of each agreed remuneration to a global DSS account for the platform worker. The DSS account collects the contributions generated globally and transfers them on a regular basis to the social security system of the country where the platform worker is located. Article [8] discusses the principles of building intelligent decision support systems of situational type for innovative development of megacities' infrastructure.

The article [9] provides new insights, develops a conceptual framework, and identifies promising research questions by putting local government AI systems under the microscope through the lens of responsible urban innovation.

The article [11] proposes a new conceptual framework for IDSS for disaster management, with a particular focus on forest fires and cold/heat waves. IDSS uses big data collected from APIs and AI to help decision makers make faster and more accurate decisions.

The article [12] explores the impact of governance on sustainability and reflects the impact of ICTs on decision-making by improving policy effectiveness, accountability and transparency in urban systems. The paper also presents conceptual system models of the cognitive city and energy behaviour, including three sub-levels: human-institutional, physical, and data. It proposes integrated conceptual models to improve the efficiency of energy systems in complex and uncertain environments, facilitate the resolution of energy consumption problems, and support capacity development at the individual, social and technical levels to improve future energy management.

The article [13] was based on a synthesized and aggregated literature review to build a new conceptual framework. The literature review revealed additional existing smart city frameworks, including city services (essential services, nonessential services, and complementary services); city resources (superstructure, infrastructure, infostructure); city architecture (enterprises); and city goals (livability, performance, and sustainability). This study contributes to a broader understanding of the smart city reference model for Indonesia and other developing countries.

The above-mentioned works mostly assess economic or social infrastructure as a separate system. Few works consider the socio-economic infrastructure as an information system (analogues).

In this regard, the purpose of this article is to develop the concept of an intelligent information system for the provision of socio-economic infrastructure of the city.

The developed intelligent information system for the provision of socio-economic infrastructure of the city differs from its analogues [5, 7, 8, 9, 12] in that it takes into account qualitative and quantitative indicators. This system can automate the distribution of powers between the state and regional governments in the development of master plans for the development of cities in the country, social programmers and other documents aimed at improving the quality of life of the population. In the future, this intelligent information system can be used as a basis for the development of other systems for similar purposes.

CONCEPT OF AN INTELLIGENT INFORMATION SYSTEM FOR THE CITY'S SOCIAL AND ECONOMIC INFRASTRUCTURE

To ensure the socio-economic infrastructure of a city, an intelligent information system (Fig. 1) should contain at least four main levels of infrastructure: economic, social, environmental, and socio-political. To achieve this goal, it is important to take into account both quantitative and qualitative indicators. Let us consider each level separately.

he social sphere includes people's attitudes towards culture, art, and tourism [18, 19], as well as an assessment of personal safety, the education system and personal education, the healthcare system and personal health, amenities and living conditions [16], and transport.

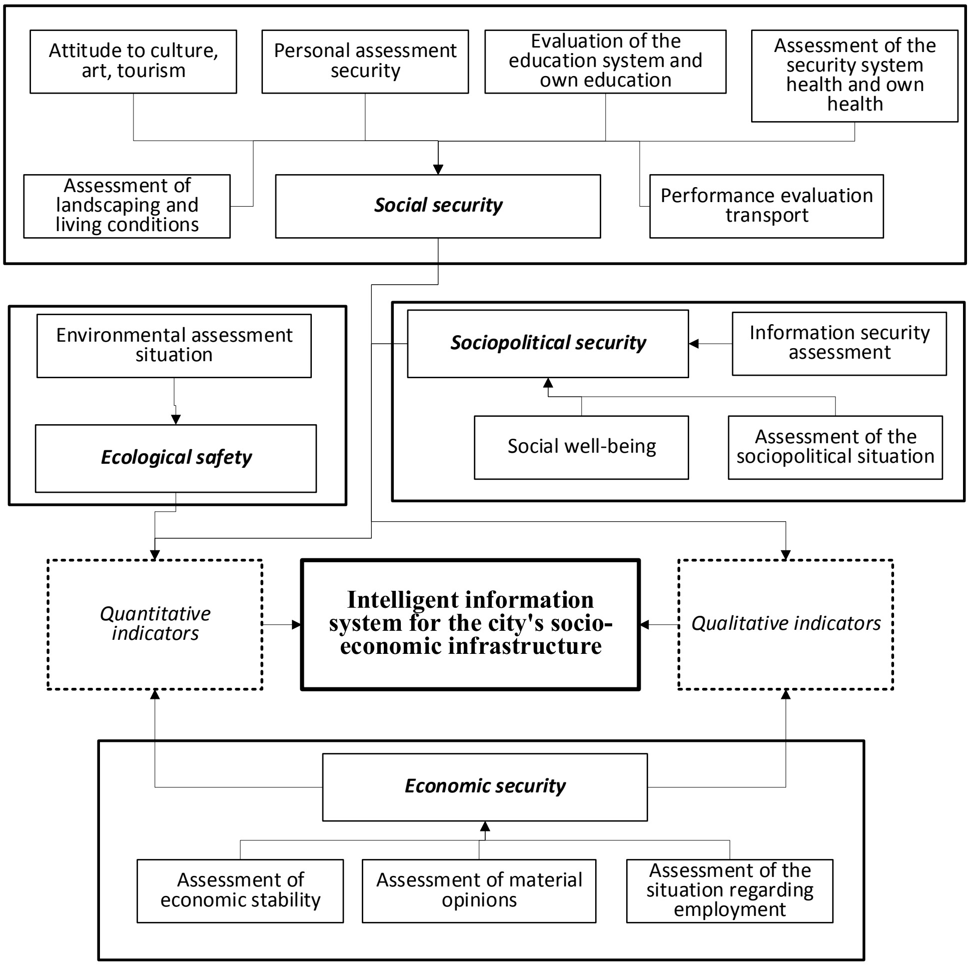


Fig. 1. The concept of an intelligent information system for the city's socio-economic infrastructure

The importance of the environmental sphere for a person is an assessment of the current environmental situation in the city.

As for the economic sphere, a city resident can assess economic stability [17], their financial situation and the employment situation.

Today, during the period of military aggression in Ukraine, socio-political security is becoming important for everyone, and every city resident can assess the socio-political situation, information security and personal social well-being.

Let's take a closer look at the structure of the data flow in the intelligent information system for the city's socio-economic infrastructure (Fig. 2).

Data will be collected based on surveys of residents, as well as on the basis of collected statistical indicators and sensor data. The latter is important for studying the city's environmental infrastructure.

All quantitative data will undergo preliminary processing and will be stored in the relevant databases. In the case of surveys, after collecting data into the database, it needs to be processed using intelligent methods to give it quantitative values.

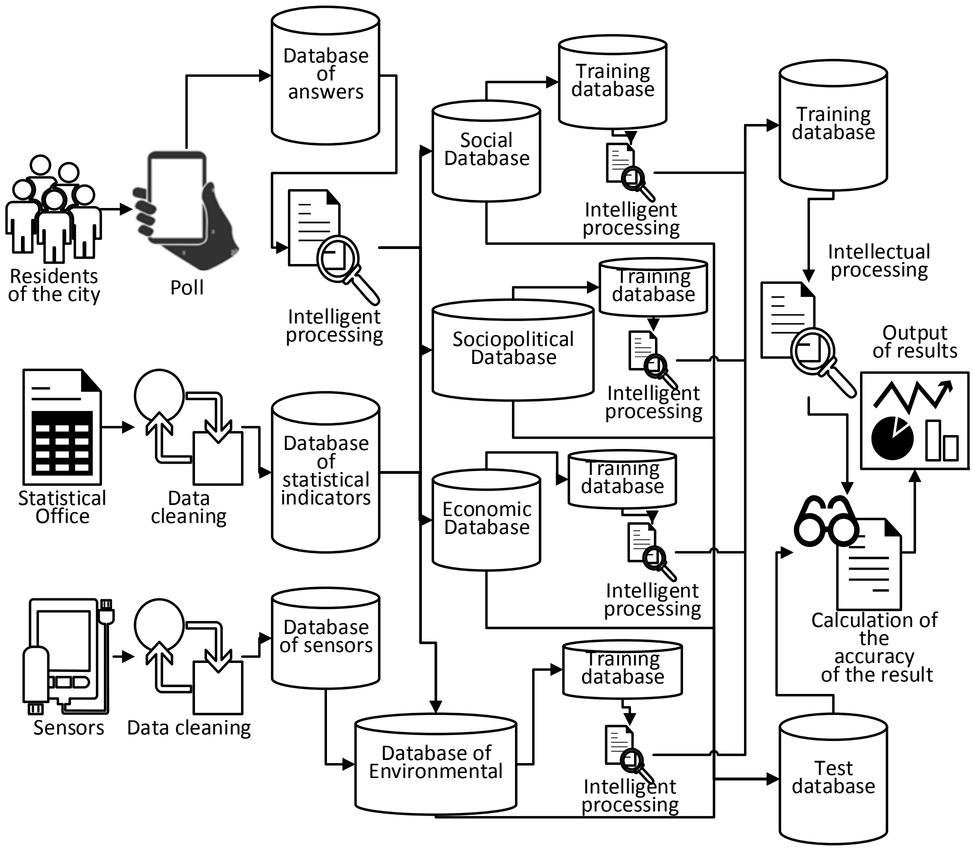


Fig. 2. Data flow in an intelligent information system for the provision of the city's socio-economic infrastructure

All data is then transferred to the databases of the four levels of socio-economic infrastructure. At each level, intelligent data processing is carried out, the results of which are transferred to a common training database. At the same time, data from the databases of the four levels of socio-economic infrastructure are transferred to a common test database.

The main intellectual processing is carried out on the common database, after which the accuracy is calculated relative to the common test database. Finally, the results are displayed.

The presented conceptual model of an intelligent information system for ensuring the socio-economic infrastructure of the city and the structure of the data flow in it makes it possible to monitor threats to society in the context of socio-economic infrastructure in real time. The developed conceptual model can be part of a smart city [10].

EXPERIMENTS AND RESULTS

In general, modelling such a system involves developing a mathematical model, programming and simulating the system on a computer. Next, we will simulate the system's operation and model the city's socio-economic infrastructure.

First of all, the objects to be modelled, the city's socio-economic infrastructure and its components, are identified. In this case, these may include city zones, residential and business units, enterprises and people, and CO_2 emissions.

Next, it is necessary to investigate how these objects will interact with each other, which is described using various algorithms and formulas. In this case, an agent-based model was created, where each agent has its own set of characteristics and interacts with other agents and the external environment.

Thus, the system uses an agent-based transport and dynamic model to simulate the movement of people and businesses in the city and their interaction. Each agent (person or enterprise) has its own properties and can change its state according to the interaction with other agents and changes in the environment. For example, a person can change his or her job, move to another area, or buy a car if he or she has sufficient funds. An enterprise can change its operation or production in response to changes in demand for its products or services.

Data is collected automatically in real time. Traffic information is collected using sensors on roads and vehicles and transmitted to the system for further processing and analysis. CO2 emissions are also collected and accounted for in the system. In addition, the system analyses data on the level of comfort of housing, the number of residential and commercial units in each zone, and other factors that affect the standard of living in the city.

Modelling based on data from a specific city, such as Ternopil, is an important stage in the development of an information system for managing socio-economic infrastructure. It allows to take into account specific features of the city, such as demographic, economic, transport and other characteristics.

Ternopil is a medium-sized city located in the western part of Ukraine. It has a rich history, as well as important economic and cultural significance for the region and the country as a whole. According to the Ukrainian State Statistics Committee, as of 1 January 2022, the population of Ternopil was over 219 thousand people [15]. The age group selected for modelling is 25-54 years, which is 44% of all citizens [14], as this category belongs to the active working age group, people who are potentially able-bodied citizens. This is an important factor for modelling the city's socio-economic infrastructure, as this category of citizens are the main users of transport and other public services, and the functioning of the city's infrastructure depends on their activity.

Based on the Ternopil city model, we will highlight some properties that may be characteristic of certain city zones:

1. The central area of Ternopil is the most commercial and business-oriented zone, with a high level of comfort, but also with high property values.
2. The South-Western zone of the city is a residential zone with a low level of comfort and average property prices.
3. The north-eastern area of the city is more industrial with a large number of factories and plants, and few residential areas.
4. The area on the western edge of the city is a residential area with a high level of comfort and high property prices.
5. The southern edge of the city is a more industrial area with few residential areas and average property prices.

Based on expert data (Table), the values of the system input parameters were formed.

An information system has been developed based on the model [15] and adapted for the city of Ternopil.

According to the results of the modelling carried out for a period of 1460 days, it was found that the largest number of enterprises (Fig. 3) is concentrated in the central, western and southern zones of the city. This is due to the fact that these zones are more developed in terms of economy and the location of the city centre. The largest concentration of population (Fig. 3) is also observed in the central and western zones of the city.

System parameters for the zones of Ternopil city

Zone	Housing capacity	Comfort level	Number of places for enterprises	Road capacity, % of total traffic
Central	600	1	35	20
North-western zone	500	0.7	25	55
North-eastern	1000	0.5	100	85
Western	700	0.9	20	30
South	500	0.6	70	90

a *b*

Fig. 3. Modelling results: *a* — workload of enterprises; *b* — po

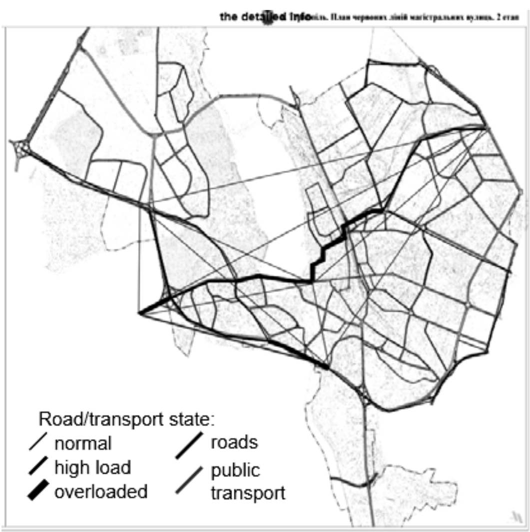


Fig. 4. Road load

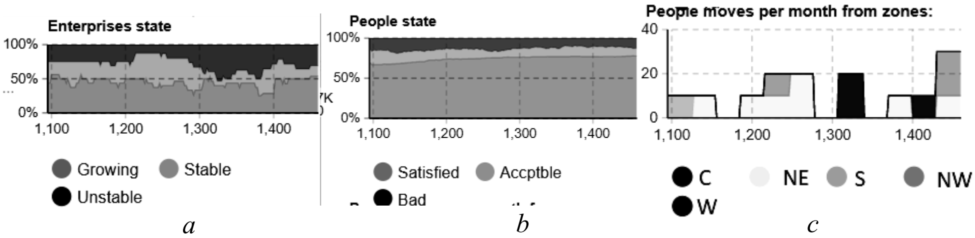
The modelling results showed that the highest road congestion (Fig. 4) in Ternopil is observed on the route from the western zone of the city to the north-eastern zone through the city center. This is because most of the city's residents live in the western and central parts of the city, but work in other parts of the city, particularly in the northeastern part. Thus, this route is key for transporting people and goods in the city.

The following is a more detailed analysis of the indicators for each zone.

First, the state of enterprises

will be analyzed: the data is presented as a time series of three variables for each point in time: growing, stable and unstable. Next, the data on the state of people, this data is a survey where people are asked to assess their condition in relation to their place of residence: satisfactory, acceptable or poor. And lastly, we will analyze the rate of moving from the respective area of the city to another.

So, let's first look at the central zone indicator (Fig. 5). The state of enterprises is characterized by "growth" at the level of 50%. This may mean that the overall level of economic development in the region is positive, or that there is a certain level of stability in this area. The state of people in a "satisfactory" condition is 70%. This indicates that the majority of the population feels satisfied with their lives, possibly due to economic achievements that allow people to meet their needs. People most often move to the north-eastern zone. This may be due to certain factors, such as job opportunities, infrastructure development, better living conditions or other factors.



a *b* *c*

Fig. 5. Central zone: *a* — the state of enterprises; *b* — human condition; *c* — people move to other zones every month

According to the data, the state of enterprises in the Northwest zone (Fig. 6) is stable on average with a business growth rate of 50%, but there is a certain probability that the state will become "unstable". The situation of people in this zone is generally satisfactory with a comfort level of 70%. People tend to move to the Western zone, probably because of the higher levels of comfort in this zone. Therefore, the North-Western zone can be an attractive place to live and develop business, provided that the businesses remain stable.

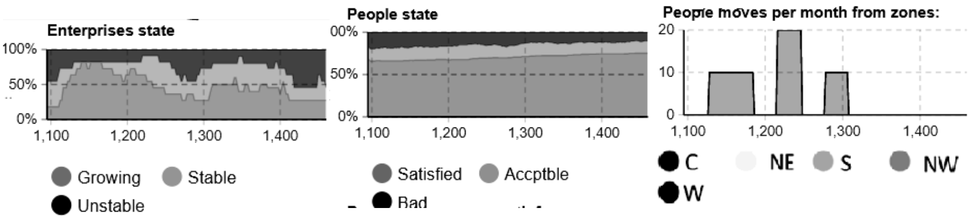
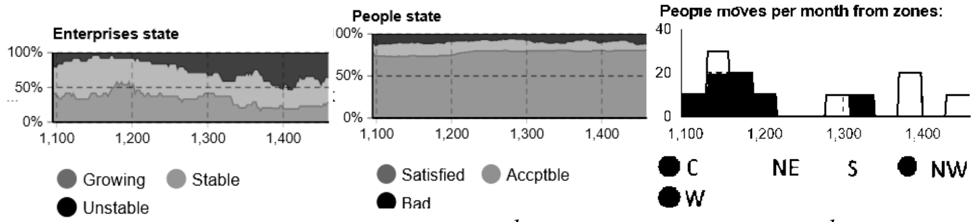


Fig. 6. North-Western zone: *a* — the state of enterprises; *b* — human condition; *c* — people move to other zones every month

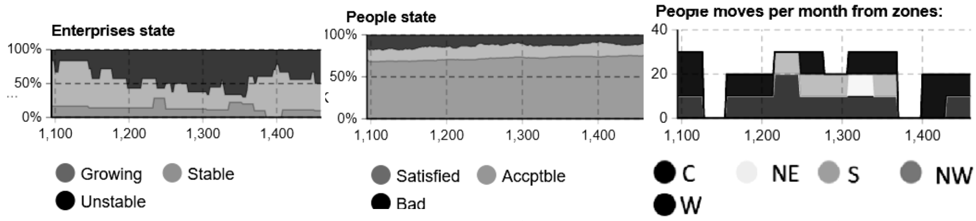
The North-Eastern zone is characterized (Fig. 7) by a less stable state of enterprises compared to other zones. The growth rate is less than 30%, although sometimes it can become unstable and reach 50%. In such periods, enterprises may have problems with maintaining and developing their business. The state of people in this zone is relatively satisfactory, as most people are currently satisfied with their situation. This may be due to the high level of employment in the area, or possibly other social factors that keep people comfortable. Nevertheless, people in the northeastern zone are more likely to move to the western zone, where they have a higher level of comfort. This may be due to fewer career opportunities in the northeastern zone, or to a lower quality of life due to less stable businesses. It is also possible that people are moving to the central zone, where there are more employment and career opportunities.



a *b* *d*

Fig. 7. North-eastern zone: *a* — the state of enterprises; *b* — human condition; *c* — people move to other zones every month

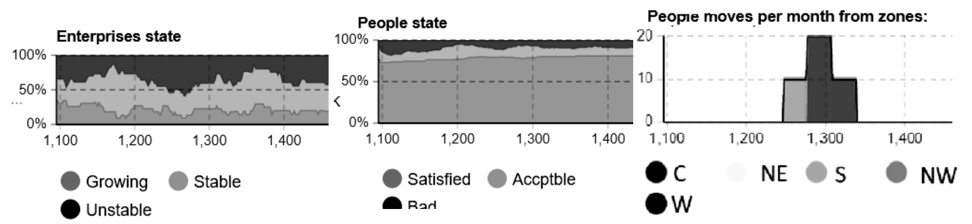
The Western zone is characterized (Fig. 8) by a lower level of economic growth compared to other zones. Most of the enterprises in the zone are in an "unstable" state, which can create difficulties for businesses and investors. However, in most cases, the condition of enterprises is "stable". People's perception of the area's quality of life is fairly high, with a "satisfactory" rating of 85%. People tend to move to the north-western zone, where housing is cheaper, and to the central zone, where there is a higher level of comfort and opportunities for career and business development.



a *b* *d*

Fig. 8. Western zone: *a* — the state of enterprises; *b* — human condition; *c* —people move to other zones every month

In the Southern zone (Fig. 9), the situation of enterprises can be characterized as stable in most cases, although some enterprises may experience instability. In general, the state of enterprises in this zone can be described as "growing" at the level of 25–30%. As for the state of people, they are in a "satisfactory" state at 85%. People are most often moving to the central zone, as it offers a higher level of comfort.



a *b* *d*

Fig. 9. Southern zone: *a* — the state of enterprises; *b* — human condition; *c* — people move to other zones every month

Also, the system has statistics (Fig. 10) that show that the total distance of car trips has increased by 249% or more than three times (the initial distance was 83.661 km and now it is 208.414 km). This has also resulted in an increase in CO_2 emissions by the same amount – 249% of the initial emissions level. This is an indicator of increasing air pollution and can have a negative impact on human health and the environment. This data can be used to evaluate the effectiveness of programmers and policies aimed at reducing car traffic and air pollution.

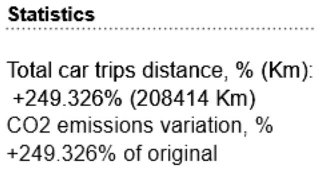


Fig. 10. Statistics on CO_2 emissions from motor vehicles

It is also possible to display information on each individual enterprise (Fig. 11) and person (Fig. 12).

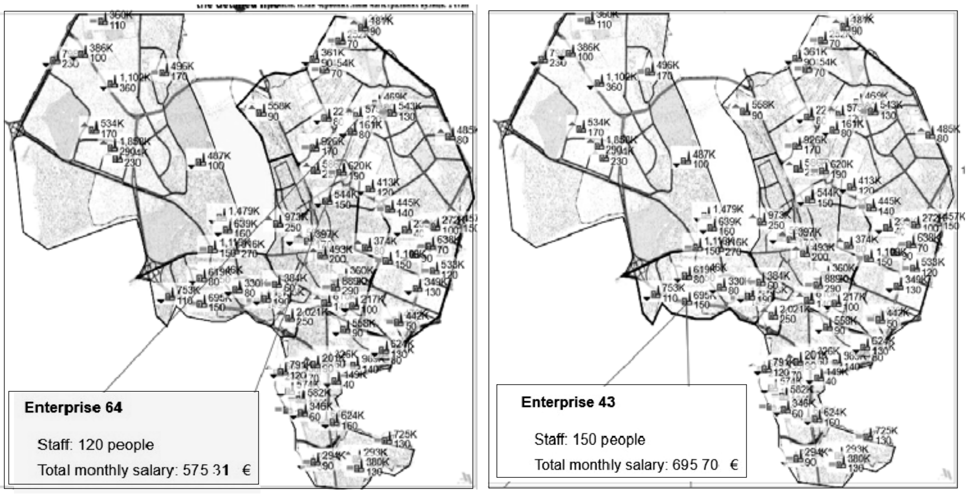


Fig. 11. Display of information on the enterprise

Fig. 12. Displaying information on people

For example, the first 64 enterprises have 120 employees and a total monthly production capacity of 57.5 thousand euros. The state of the enterprise is "growing", which may indicate a successful team and effective business management. The second enterprise 43 has a larger number of employees – 150 people. The total monthly production capacity of this enterprise is 69.5 thousand euros. The company's condition is "stable", which may indicate its long-term sustainability and reliability. Overall, both enterprises have significant potential and can be successful in the market for their respective goods and services. However, their future development depends on how efficiently they use their resources and how well they can adapt to changes in the current business environment.

The first person, 917 (Fig. 12), is 24 years old, lives in the north-western zone and also works in the central zone. According to the status, she also considers her condition to be satisfactory, with an income of €1058 per month. The second person 381 is 39 years old, lives in the southern zone and works in the central zone. According to the status, she considers her situation to be satisfactory. Her income is 892 euros per month.

Thus, during the 1460 days of modelling the socioeconomic infrastructure of Ternopil, it was found that the state of enterprises in different zones has different levels of growth, and the state of people is mostly satisfactory. People tend to move to areas with higher levels of comfort or cheaper housing. The total distance of car journeys has increased by 249%, resulting in CO_2 emissions in line with this increase. The model also provides information about two people – one living in the southern zone and the other in the north-western zone, both working in the center and in good health.

CONCLUSIONS

The concept of an intelligent information system for the provision of the city's socioeconomic infrastructure and the structure of the data flow in it have been developed. This system will automate the distribution of powers between the state and regional governments in the development of master plans for the development of cities in the country, social programmers and other documents aimed at improving the quality of life of the population.

Experimental results of the 1460-day simulation of the intelligent information system for the city's socioeconomic infrastructure show a 50% increase in businesses in the central zone and a 70% increase in people's satisfaction, as well as a move of residents to the north-eastern and western zones due to a higher level of comfort. The total distance of car journeys increased by 249%, and the change in CO_2 emissions also increased by the same percentage, which is negative for the environment. The condition of enterprises in other zones is less stable, with lower "growth" scores, but there are also zones with a stable and "satisfactory" condition.

Further research may include the development of new methods and algorithms for using data from the city's socioeconomic infrastructure in intelligent information systems, analysis of the impact of intelligent information systems on the economic and social development of the city, development of new information technologies for creating intelligent information systems, and research on the impact of an intelligent information system on the environmental sustainability of the city. In addition, the interaction of intelligent information systems with local governments, business and the public, as well as the impact of intelligent information systems on ensuring accessibility and equality of use of the city's socioeconomic infrastructure for all its residents and visitors can be studied.

REFERENCES

1. A.U.B Usmonjon o'g'li and I.B. Botirali o'g'li, "The Essence and Normative Indicators of Economic Security," *Academic Journal of Digital Economics and Stability*, vol. 15, pp. 1–15, 2022.
2. S.I. Sotskova and I.V. Kalashnikova, "Comprehensive Assessment of Enterprise's Economic Security System in the Digital Economy Conditions," in Ashmarina S.I., Mantulenko V.V. (eds) *Digital Technologies in the New Socioeconomic Reality. ISCDTE 2021. Lecture Notes in Networks and Systems*, vol. 304. Springer, Cham, 2022. doi: 10.1007/978-3-030-83175-2_64.
3. L. Lelyk, V. Olikhovskyi, N. Mahas, and M. Olikhovska, "An integrated analysis of enterprise economy security," *Decision Science Letters*, 11(3), pp. 299–310, 2022.
4. O. Hrybinenko, O. Bulatova, and O. Zakharova, "Financial indicators in the system of economic security of the world countries," *11th International Scientific Conference "Business and Management 2020"*, pp. 273–281.
5. T. Obelets, "Data mining tools for complex socioeconomic processes and systems," *System Research and Information Technologies*, no. 4, pp. 68–78, 2022.
6. R. Khajouei and F. Farahani, "The evaluation of users' satisfaction with the Social Security Electronic System in Iran," *Health Technol.*, 9(1), pp. 797–804, 2019. doi: 10.1007/s12553-019-00347-y.
7. E. Weber, *Digital Social Security: Outline of a concept for the 21st century, Working Paper Forschungsförderung, No. 138*. Hans-Böckler-Stiftung, Düsseldorf, 2019.
8. O. Trofymchuk, A. Stenin, M. Soldatova, and I. Drozdovich, "Intelligent decision support systems in the development of megalopolis infrastructure," *System Research and Information Technologies*, no. 2, pp. 61–74, 2022. doi: 10.20535/SRIT.2308-8893.2022.2.04.
9. T. Yigitcanlar, J.M. Corchado, R. Mehmood, R.Y.M. Li, K. Mossberger, and K. Desouza, "Responsible Urban Innovation with Local Government Artificial Intelligence (AI): A Conceptual Framework and Research Agenda," *Journal of Open Innovation: Technology, Market, and Complexity*, 7(1):71, 2021. doi: 10.3390/joitmc7010071.
10. J. Wang, D.Q. Nguyen, T. Bonkalo, and O. Grebennikov, "Smart governance of urban data," in *E3S Web of Conferences*, vol. 301, pp. 1–8, 2021.
11. D. Jung, V. Tran Tuan, D. Quoc Tran, M. Park, and S. Park, "Conceptual Framework of an Intelligent Decision Support System for Smart City Disaster Management," *Applied Sciences*, 10(2):666, 2020. doi: 10.3390/app10020666.
12. M. Mansouri and N. Khansari, "A Conceptual Model for Intelligent Urban Governance: Influencing Energy Behaviour in Cognitive Cities," in Portmann E., Tabacchi M., Seising R., Habenstein A. (eds) *Designing Cognitive Cities. Studies in Systems, Decision and Control*, vol 176. Springer, Cham, 2019. doi: 10.1007/978-3-030-00317-3_8.
13. K.A. Achmad, L.E. Nugroho, and A. Djunaedi, "Smart city for development: towards a conceptual framework," in *2018 4th International Conference on Science and Technology (ICST)*, pp. 1–6.
14. *State Statistics Service of Ukraine*. Available: https://ukrstat.gov.ua/
15. *AnyLogic Cloud: Online Simulation Tools. AnyLogic Cloud: Online Simulation Tools*. Available: https://cloud.anylogic.com/model/dc9cf4ee-0463-4324-92be-f13f394fe840?mode=SETTINGS&tab=GENERAL

16. K. Lipianina-Honcharenko, C. Wolff, Z. Chyzhovska, A. Sachenko, T. Lendiuk, and S. Grodskyi, "Intelligent method for forming the consumer basket," *Information and Software Technologies*, 2022. doi: 10.1007/978-3-031-16302-9_17.
17. H. Lipyanina, S. Sachenko, T. Lendyuk, V. Brych, V. Yatskiv, and O. Osolinskiy, "Method of detecting a fictitious company on the machine learning base," *Advances in Computer Science for Engineering and Education*, 2021. doi: 10.1007/978-3-030-80472-5_12.
18. K. Lipianina-Honcharenko, R. Savchyshyn, A. Sachenko, A. Chaban, I. Kit, and T. Lendiuk, "Concept of the intelligent guide with AR support," *International Journal of Computing*, 21(2), pp. 271–277, 2022. doi: 10.47839/ijc.21.2.2596
19. V. Krylov et al., "Multiple regression method for analyzing the tourist demand considering the influence factors," Paper presented at the *Proceedings of the 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2019*, pp. 974–979. doi: 10.1109/IDAACS.2019.8924461.

Received 15.03.2023

INFORMATION ON THE ARTICLE

Khrystyna V. Lipianina-Honcharenko, ORCID: 0000-0002-2441-6292, West Ukrainian National University, Ukraine, e-mail: xrustya.com@gmail.com, kh.lipianina@wunu.edu.ua

Yevgeniy V. Bodyanskiy, ORCID: 0000-0001-5418-2143, Kharkiv National University of Radio Electronics, Ukraine, e-mail: yevgeniy.bodyanskiy@nure.ua

Anatoliy O. Sachenko, ORCID: 0000-0002-0907-3682, West Ukrainian National University, Ukraine, e-mail: as@wunu.edu.ua

ІНТЕЛЕКТУАЛЬНО ІНФОРМАЦІЙНА СИСТЕМА СОЦІАЛЬНО-ЕКОНОМІЧНОЇ ІНФРАСТРУКТУРИ МІСТА / Х.В. Ліп'яніна-Гончаренко, Є.В. Бодянський, А.О. Саченко

Анотація. Розвиток міст є важливою проблемою, яку можна вирішити за допомогою інтелектуальних інформаційних систем. Такі системи забезпечують ефективне управління різноманітною інфраструктурою міста. Дослідники розробили концепцію такої інформаційної системи, яка базується на концептуальній моделі та використовує потік даних для розумного прийняття рішень. Систему протестовано на період 1460 днів у місті Тернопіль. Результати моделювання показали, що центральна зона міста є стабільною зі станом підприємств «зростаючий» на рівні 50% та станом людей у стані «задовільно» на рівні 70%. Люди найчастіше переїжджають у північно-східну та західну зони через вищий рівень комфорту та більш доступне житло, проте загальна відстань автомобільних поїздок збільшилась на 249%, що має негативний вплив на довкілля. Стан підприємств у інших зонах є менш стабільним з нижчими показниками «зростання», але є зони зі «стабільним» станом і станом «задовільно».

Ключові слова: моделювання, інформаційна система, соціально-економічна інфраструктура, місто.

UDC 519.004.942
DOI: 10.20535/SRIT.2308-8893.2023.3.09

ALGORITHM FOR SIMULATING MELTING POLAR ICE, EARTH INTERNAL MOVEMENT AND VOLCANO ERUPTION WITH 3-DIMENSIONAL INERTIA TENSOR

YOSHIO MATSUKI, PETRO BIDYUK

Abstract. This paper reports the result of an investigation of a hypothesis that the melting polar ice of Earth flowing down to the equatorial region causes volcano eruptions. We assumed a cube inside the spherical body of Earth, formulated a 3-dimensional inertia tensor of the cube, and then simulated the redistribution of the mass that is to be caused by the movement of melted ice on the Earth's surface. Such mass distribution changes the inertia tensor of the cube. Then, the cube's rotation inside Earth was simulated by multiplying the Euler angle matrix by the inertia tensor. Then, changes in the energy intensity and the angular momentum of the cube were calculated as coefficients of Hamiltonian equations of motion, which are made of the inertia tensor and sine and cosine curves of the rotation angles. The calculations show that the melted ice increases Earth's internal energy intensity and angular momentum, possibly increasing volcano eruptions.

Keywords: inertia tensor, volcano eruption, mass distribution, Hamiltonian equation of motion.

INTRODUCTION

The polar ice melts and flows to the Equatorial region by Earth's centrifugal force by its rotation (from Fig. 1 to Fig. 2), then Earth swells along the Equatorial region (Fig. 3). The mass balance of Earth changes, then its inertia tensor changes. We assume that the change of the inertia tensor, which is caused by the redistribution of mass from the polar region to the equatorial region, excites the Earth's internal energy and the angular momentum, causing volcano eruptions.

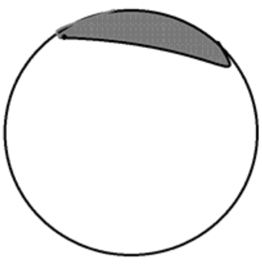


Fig. 1. Polar ice

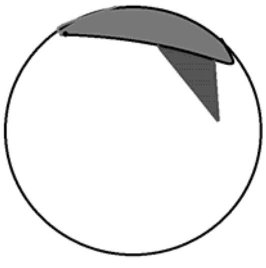


Fig. 2. Melting ice

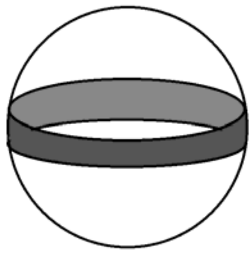


Fig. 3. Swell around Equator

ALGORITHM OF THE SIMULATION

Equation (1) shows how to calculate inertia tensor of a solid body, r is a distance between the centre of mass and each point of the body, V is the volume of the body, α and β are suffixes that represent coordinates' axis, and δ is Kronecker's delta:

$$I_{\alpha\beta} = \int_V \rho(r)(\delta_{\alpha\beta} r^2 - r_\alpha r_\beta) dV . \quad (1)$$

Fig. 4 shows a spherical body in a 3-dimensional flat space. Inside of the sphere, we put a cube, 3 sides of which are on x, y and z axis of a flat 3-dimensional coordinate system. The cube's mass is M, each side's length is a, and we set:

$$b = Ma^2 .$$

Applying (1) to this cube, the inertia tensor becomes:

$$I = \begin{bmatrix} \frac{2}{3}b & -\frac{1}{4}b & -\frac{1}{4}b \\ -\frac{1}{4}b & -\frac{2}{3}b & -\frac{1}{4}b \\ -\frac{1}{4}b- & \frac{1}{4}b- & \frac{2}{3}b \end{bmatrix} . \quad (2)$$

The inertia tensor (2) of this cube is taken from the example shown at end of the chapter 5.3 "The inertia tensor and the moment of inertia" in page 94 of [1].

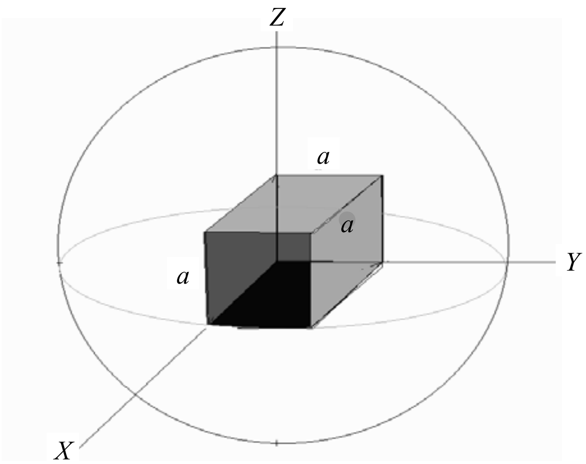


Fig. 4. A spherical body in a 3-dimensional flat space

By rotating the spherical body around z-axis, (2) is multiplied by Euler angle D:

$$D = \begin{bmatrix} \cos & \sin & 0 \\ -\sin & \cos & 0 \\ 0 & 0 & 1 \end{bmatrix} .$$

Then we get:

$$ID = \begin{bmatrix} \frac{2}{3}b\cos & -\frac{1}{4}b\sin & \frac{2}{3}b\sin & -\frac{1}{4}b\cos & -\frac{1}{4}b \\ -\frac{1}{4}b\cos & -\frac{2}{3}b\sin & -\frac{1}{4}b\sin & +\frac{2}{3}b\cos & -\frac{1}{4}b \\ -\frac{1}{4}b\cos & +\frac{1}{4}b\sin & -\frac{1}{4}b\sin & -\frac{1}{4}b\cos & \frac{2}{3}b \end{bmatrix}. \quad (3)$$

For calculating the energy intensity of the rotation, we take the diagonal components of (3) to make a vector

$$X = \begin{bmatrix} (2/3)b\cos \quad +(1/4)b\sin \\ -(1/4)b\sin \quad +(2/3)b\cos \\ (2/3)b \end{bmatrix}, \quad (4)$$

while, for calculating the angular momentum of the rotation, we take non-diagonal (y, x) and (x, y) components of (3) to make a vector

$$X = \begin{bmatrix} -\frac{1}{4}\cos & +\frac{2}{3}b\sin \\ -\frac{1}{4}\cos & -\frac{2}{3}b\sin \end{bmatrix}. \quad (5)$$

This re-formulation of the matrix to the vectors for energy intensity and for angular momentum is explained from page 14 to page 21 of [2].

Then we make the Hamiltonian equation of motion:

$$H = kT - Xc. \quad (6)$$

Here, kT is a stress energy that reflects the energy intensity and angular momentum of the rotating body. In this simulation, we set it as a unity vector. And c is a coefficient vector, which is to be calculated as energy intensity of the rotating body or the angular momentum; Xc for energy intensity is

$$Xc = C_0\left(\frac{2}{3}b\cos \quad +\frac{1}{4}\sin\right) + C_1\left(-\frac{1}{4}b\sin \quad +\frac{2}{3}b\cos\right) + C_2\left(\frac{2}{3}b\right),$$

and for angular momentum it is

$$Xc = C_0\left(-\frac{1}{4}b\cos \quad +\frac{2}{3}b\sin\right) + C_1\left(-\frac{1}{4}b\cos \quad -\frac{2}{3}b\sin\right).$$

Then, X' is multiplied from the left side of (6), and we set it to be zero as the boundary condition to make $X'H = X'(kT - Xc) = 0$. Then, c will be calculated by transforming $X'H = X'(kT - Xc) = 0$ to $X'Xc = X'kt$, and then to $X'Xc = X'kt$.

Here, X' is a transposed vector of X. $(X'X)^{-1}$ is an inverse matrix of $(X'X)$.

INPUT DATA FOR THE NUMERIC SIMULATION

We assign unity for b, therefore M and a become unity in (2) in order to simulate the relative values and their changes of the energy intensity and angular momentum, not the absolute values. Then we deduct dx, dy and dz from M in each of 3 directions in x, y and z-axis, as shown in (7) for the energy level and (8) for angular momentum:

$$X = \begin{bmatrix} \frac{2}{3}b\cos \quad +\frac{1}{4}b\sin \quad +dx \\ -\frac{1}{4}b\sin \quad +\frac{2}{3}b\cos \quad +dy \\ \frac{2}{3}b - dz \end{bmatrix}; \qquad (7)$$

$$X = \begin{bmatrix} -\frac{1}{4}\cos \quad +\frac{2}{3}b\sin \quad +dx \\ -\frac{1}{4}\cos \quad -\frac{2}{3}b\sin \quad +dy \end{bmatrix}. \qquad (8)$$

First, we assign the value for d_z, then calculate d_x and d_y by $dx = dy = \sqrt{\frac{1}{a-dz}} - 1 = \sqrt{\frac{1}{1-dz}} - 1$, so that these can make the volume of the cube to be unity. Then we simulate 4 cases by assigning 4 different sets of the values of d_x, d_y and d_z, which are shown in Table 1:

$$dx = dy = \left(\sqrt{\frac{1}{a-dz}} - 1 = \sqrt{\frac{1}{1-dz}} - 1\right).$$

Table 1. Marginal changes, d_x, d_y and d_z, of Earth's mass, which are to be reflected to the cubic

Marginal mass change	Case 1	Case 2	Case 3	Case 4
d_x	0	0.00503	0.0541	0.118
d_y	0	0.00503	0.0541	0.118
d_z	0	0.01	0.1	0.2

Here, the image of Case 1 is shown in Fig. 1, and Cases 2, 3 and 4 are in Fig. 3.

For (4), (5), (7) and (8), we use sine and cosine curves shown in Fig. 5.

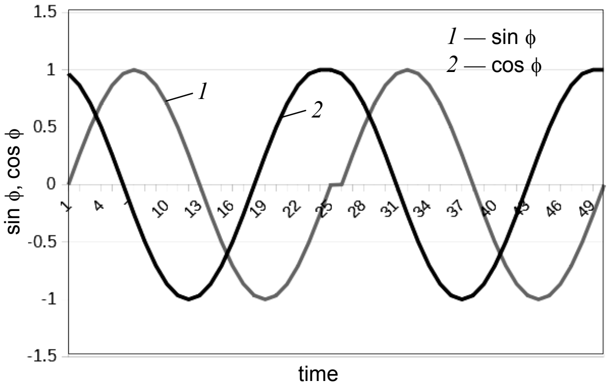


Fig. 5. Sin ϕ and cos ϕ

RESULT

The result of the calculation of the coefficient vectors are shown in Table 2 and Fig. 6 for the energy intensity and in Table 3 and Fig. 7 for the angular momentum. As the mass moves from the North Pole to the Equatorial region the energy intensity becomes larger on z-axis and the angular momentum also becomes larger on x-y plane. Fig. 8 shows the caricatured images of these calculated results.

Table 2. Calculated coefficient vector for energy intensity

Energy intensity in	Case 1, d_z=0	Case 2, d_z=0.01	Case 3, d_z=0.1	Case 4, d_z=0.2
x	$3.053 \cdot 10^{-16}$	$-6.94 \cdot 10^{-17}$	$-2.08 \cdot 10^{-16}$	$-7.49 \cdot 10^{-16}$
y	$-2.220 \cdot 10^{-16}$	$-5.03 \cdot 10^{-17}$	$-2.03 \cdot 10^{-16}$	$-8.67 \cdot 10^{-18}$
z	1.50	1.52	1.76	2.14

Table 3. Calculated coefficient vector for angular momentum

Angular momentum in	Case 1, d_z=0	Case 2, d_z=0.01	Case 3, d_z=0.1	Case 4, d_z=0.2
x	- 0.146	$-7.19 \cdot 10^{-2}$	0.617	1.16
y	0.178	$-8.72 \cdot 10^{-2}$	0.750	1.41

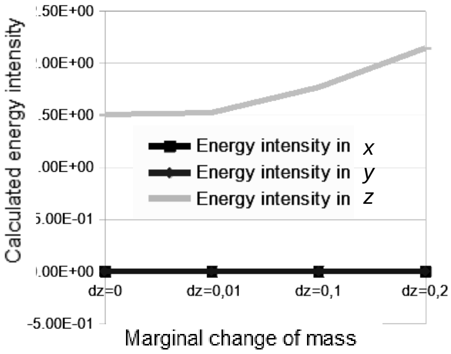


Fig. 6. Calculated energy intensity

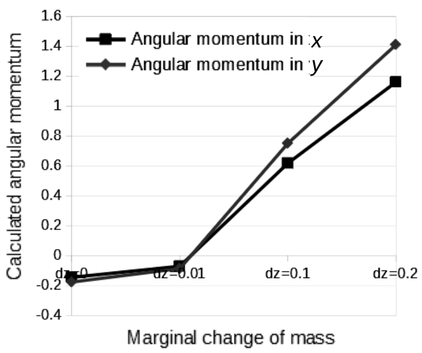


Fig. 7. Calculated angular momentum

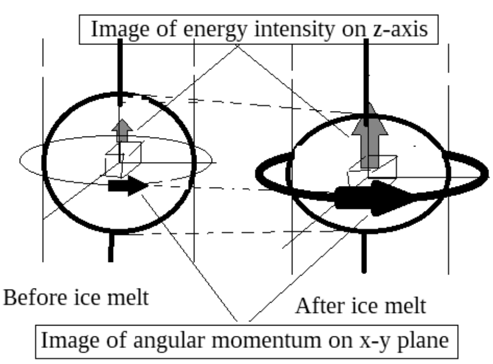


Fig. 8. Images of the simulation results

CONCLUSIONS AND RECOMMENDATIONS

After the polar ice melting, the internal energy intensity of Earth becomes larger on *z*-axis, and the angular momentum also becomes larger on *x*-*y* plane. This result suggests that the polar ice melting influences the Earth's internal energy intensity as well as the internal angular momentum, and larger. This result means a possibility of volcano eruptions.

The simulated result should be compared to the observations on Earth, and appropriate methodologies need to be developed for the comparison.

REFERENCES

1. Goldstein, C.P. Poole, and J.L. Safko, *Classical Mechanics*, 3rd Edition. Pearson Education, Inc., 2002, pp. 646
2. Y. Matsuki and P.I. Bidyuk, *A Course in Black Hole Simulation*, LAP Lambert Academic Publishing, 2021, pp. 69.

Received 08.01.2023

INFORMATION ON THE ARTICLE

Petro I. Bidyuk, ORCID: 0000-0002-7421-3565, Educational and Research Institute for Applied System Analysis of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: pbidyuke_00@ukr.net

Yoshio Matsuki, ORCID: 0000-0002-5917-8263, National University of Kyiv-Mohyla Academy, Ukraine, e-mail: matsuki@wdc.org.ua

АЛГОРИТМ МОДЕЛЮВАННЯ ТАНЕННЯ ПОЛЯРНОГО ЛЬОДУ, ВНУТРІШНЬОГО РУХУ ЗЕМЛІ ТА ВИВЕРЖЕННЯ ВУЛКАНУ З 3-ВИМІРНИМ ТЕНЗОРОМ ІНЕРЦІЇ / Й. Мацукі, П.І. Бідюк

Анотація. Подано результати дослідження гіпотези про те, що танення полярного льоду Землі, що стікає до екваторіальної ділянки, викликає виверження вулканів. Припустили, що всередині сферичного тіла Землі є куб, сформовано тривимірний тензор інерції куба, потім змодельовано перерозподіл маси, який буде спричинений рухом талого льоду на поверхні Землі. Такий розподіл мас змінює тензор інерції куба. Змодельовано обертання куба всередині Землі шляхом множення матриці кута Ейлера на тензор інерції. Зміни енергоємності та моменту імпульсу куба розраховано як коефіцієнти гамільтонових рівнянь руху, які складаються з тензора інерції та кривих синусів і косинусів кутів повороту. Результати розрахунків показують, що талий лід збільшує інтенсивність внутрішньої енергії Землі та кутовий момент, що означає можливе збільшення вивержень вулканів.

Ключові слова: тензор інерції, виверження вулкана, розподіл маси, гамільтонове рівняння руху.

UDC 004.942:537.525
DOI: 10.20535/SRIT.2308-8893.2023.3.10

BASIC ALGORITHM FOR APPROXIMATION OF THE BOUNDARY TRAJECTORY OF SHORT-FOCUS ELECTRON BEAM USING THE ROOT-POLYNOMIAL FUNCTIONS OF THE FOURTH AND FIFTH ORDER

I. MELNYK, A. POCHYNOK

Abstract. The new iterative method of approximating the boundary trajectory of a short-focus electron beam propagating in a free drift mode in a low-pressure ionized gas under the condition of compensation of the space charge of electrons is considered and discussed in the article. To solve the given approximation task, the root-polynomial functions of the fourth and fifth order were applied, the main features of which are the ravine character and the presence of one global minimum. As an initial approach to solving the approximation problem, the values of the polynomial coefficients are calculated by solving the interpolation problem. After this, the approximation task is solved iteratively. All necessary polynomial coefficients are calculated multiple times, taking into account the values of the function and its derivative at the reference points. The final values of polynomial coefficients of high-order root-polynomial functions are calculated using the dichotomy method. The article also provides examples of the applying fourth-order and fifth-order root-polynomial functions to approximate sets of numerical data that correspond to the description of ravine functions. The obtained theoretical results are interesting and important for the experts who study the physics of electron beams and design modern industrial electron beam technological equipment.

Keywords: approximation, interpolation, root-polynomial function, ravine function, least-square method, discrepancy, approximation error, electron beam, electron-beam technologies.

INTRODUCTION

Today, an important task regarding the further development and industrial application of electron beam technologies is the preliminary estimate of the boundary trajectory of the electron beam using different suitable approaches. Therefore, in addition to development the basic theory of electron beam optics and obtaining necessary analytical ratios and corresponded numerical methods for solving differential equations, methods of interpolation and approximation are widely used also [1–3].

A separate issue in this aspect is the evaluation of electron beam trajectories and finding the focal parameters of beams in high-voltage glow discharge (HVGD) electron guns [1; 4–7]. Main singularities of such kind of beams, at the physical point of view, is its propagation in the soft vacuum in the medium of residual gas with compensation the space charge of electrons. In additional, usually such beams are formed by the cathodes with large emission surface, therefore the convergence angle of beam is generally large and its focal diameter is not so small, range of few millimeters. Just today HVGD electron guns widely used in various branches of industry, in particular in the electronic industry, instrument building, mechanical engineering, metallurgy, automobile and aerospace industry [5–9]. The main advantages of these types of guns, regarding the possibility of their industrial application, are operation in a soft vacuum in the medium of various technological gases, including noble and active gases, high stability and reliability of operation of the HVGD electron, the relative simplicity of the design and the cheap of HVGD electron guns, as well as stability and reliability of its operation [1–3]. Ease of control the power of the electron beam both by gas dynamic lows and changing the operation pressure in discharge chamber and by the lighting of additional discharges is also possible [8; 9].

Among the advanced application of HVGD electron guns in the modern electronic production most important are follows.

1. Welding of contacts and casualization of crystals. For example, such application is very advanced in the experimental production of cryogenic low-temperature devices [10; 11].

2. Production of high-quality capacitors with the small value of current losses on the base of ceramic films [12–14].

3. Production of communication devices as receivers and transmitters of microwaves antennas on the base of high-quality ceramic films [12–14].

4. Refining of silicon ingot for obtaining the pure material for electronic industry [15–18].

Main problems of HVGD optics and energetics are well-known and have been complexly analyzed in papers [1; 19–21]. The problems of guiding short-focus electron beam in ionized gas also have been studied carefully both theoretically and experimentally, corresponded mathematical model was presented in the paper [1]. However, mathematical methods of interpolation and approximation of electron beam boundary trajectories in the medium of ionized gas still wasn't developed up to the necessary stage, corresponded mathematical function also haven't considered complexly. This shortcoming largely hinders the introduction into the industry of advanced electron-beam technologies.

In the papers [22; 23] the root-polynomial function was considered as the suitable mathematical tools to interpolation the boundary trajectories of shorty-focus electron beams in the case of its propagation in the medium of ionized gases with compensation of the space charge of electrons. Root-polynomial functions from second to fifth order and corresponded interpolation results were presented and analyzed in papers [22; 23]. The interpolation results have been compared with the accurate solution of differential equation of electron beam propagation, and corresponded interpolation error usually was smaller, than 5% [22; 23]. Therefore, the aim of investigations, which are described in this article, is forming the algorithm of approximation of boundary trajectories of short-focus electron beam, propagated in the ionized gas with compensation the space charge of electrons. Testing examples of using such approximation for the root-polynomial functions of fourth and fifth order are also considered and obtained results of numerical simulation are analyzed.

THE PREVIOUS RESEARCHES AND THEORETICAL FUNDAMENTALS OF PROPOSED APPROACH

The basic theory of polynomials interpolation and approximation is considered generally in the manual books [24; 25]. In the papers [22; 23] was considered the task of interpolation the ravine functions, which corresponded to the boundary trajectories of electron beam, propagated in the medium of ionized gas, by the root-polynomial functions, which in the general form are written as:

$$r(z) = \sqrt[n]{C_n z^n + C_{n-1} z^{n-1} + \ldots + C_1 z + C_0}, \tag{1}$$

where z is the longitudinal coordinate, r is the radius of the boundary trajectory of the electron beam, n is the degree of the polynomial, as well as the order of the root function, $C_0 - C_n$ are the polynomial coefficients.

The analytical relations for coefficients of forth order root-polynomial function C_0, C_1, C_2, C_3 and C_4, which, in general form, corresponding to relation (1), is written as follows [22; 23]:

$$r(z) = \sqrt[4]{C_4 z^4 + C_3 z^3 + C_2 z^2 + C_1 z + C_0}, \tag{2}$$

are also was obtained and analyzed in the papers [22; 23].

Clear, that for 5 unknown polynomial coefficients of function (2) C_0, C_1, C_2, C_3 and C_4, with defined basic values of the spatial coordinates r_1, r_2, r_3, r_4, r_5, z_1, z_2, z_3, z_4 and z_5, corresponded set of 5 linear equation for calculation the polynomial coefficients is written as follows [22; 23]:

$$\begin{cases} C_4 z_1^4 + C_3 z_1^3 + C_2 z_1^2 + C_1 z_1 + C_0 = r_1^4; \\ C_4 z_2^4 + C_3 z_2^3 + C_2 z_2^2 + C_1 z_2 + C_0 = r_2^4; \\ C_4 z_3^4 + C_3 z_3^3 + C_2 z_3^2 + C_1 z_3 + C_0 = r_3^4; \\ C_4 z_4^4 + C_3 z_4^3 + C_2 z_4^2 + C_1 z_4 + C_0 = r_4^4; \\ C_4 z_5^4 + C_3 z_5^3 + C_2 z_5^2 + C_1 z_5 + C_0 = r_5^4. \end{cases} \tag{3}$$

For solving the set of equations (3) firstly considered the coefficients basic intermediate variables $a_{k,l}$, where k — number of iterations for solving set of equation (2) and l — number of equation i in the set (3) [22; 23]. Corresponded analytical relations are look as follows:

$$a_{1,2} = \frac{r_2^4 - r_1^4}{z_2 - z_1}; \;\; a_{1,3} = \frac{r_3^4 - r_1^4}{z_3 - z_1}; \;\; a_{1,4} = \frac{r_4^4 - r_1^4}{z_4 - z_1}; \;\; a_{1,5} = \frac{r_5^4 - r_1^4}{z_5 - z_1};$$

$$a_{2,3} = \frac{a_{1,3} - a_{1,2}}{z_3 - z_2}; \;\; a_{2,4} = \frac{a_{1,4} - a_{1,2}}{z_4 - z_2}; \;\; a_{2,5} = \frac{a_{1,5} - a_{1,2}}{z_5 - z_2}. \tag{4}$$

After that, considering the second set of additional variables $b_{k,m,l}$, where parameter m is the power of variable z in the set of equations (3). Corresponded analytical relations are written as follows:

$$b_{2,3,3} = \frac{z_3^3 - z_2^3 + z_3^2 z_1 - z_2^2 z_1 + z_1^2 z_3 - z_1^2 z_2}{z_3 - z_2};$$

$$b_{2,3,3} = \frac{z_3^2 - z_2^2 + z_1 z_3 - z_1 z_2}{z_3 - z_2};$$

$$b_{2,3,3} = \frac{z_4^3 - z_2^3 + z_4^2 z_1 - z_2^2 z_1 + z_1^2 z_4 - z_1^2 z_2}{z_4 - z_2};$$

$$b_{2,2,4} = \frac{z_4^2 - z_2^2 + z_1 z_4 - z_1 z_2}{z_4 - z_2};$$

$$b_{2,3,5} = \frac{z_5^3 - z_2^3 + z_5^2 z_1 - z_2^2 z_1 + z_1^2 z_5 - z_1^2 z_2}{z_5 - z_2};$$

$$b_{2,2,5} = \frac{z_5^2 - z_2^2 + z_1 z_5 - z_1 z_2}{z_5 - z_2}; \tag{5}$$

$$b_{3,3,4} = \frac{b_{2,3,4} - b_{2,3,3}}{b_{2,2,4} - b_{2,2,3}}; \quad b_{3,3,5} = \frac{b_{2,3,5} - b_{2,3,3}}{b_{2,2,5} - b_{2,2,3}}.$$

After that, with known values of the coefficients $b_{2,2,4}$, $b_{2,2,3}$ and $b_{2,2,5}$, five additional variables from the first data set $a_{4,5}$, $a_{3,4}$, $a_{3,5}$, $a_{3,2}$, and $a_{3,1}$ as well as two new coefficients from the second set of variables $b_{3,1}$ and $b_{3,2}$, arecalculated by using such analytical relations:

$$a_{4,5} = \frac{a_{2,4} - a_{2,3}}{b_{2,2,4} - b_{2,2,3}}; \quad a_{3,4} = \frac{a_{2,4} - a_{2,3}}{b_{2,2,4} - b_{2,2,3}}; \quad a_{3,5} = \frac{a_{2,5} - a_{2,3}}{b_{2,2,5} - b_{2,2,3}}; \tag{6}$$

$$a_{3,2} = \frac{a_{2,5} - a_{2,3}}{b_{2,2,5} - b_{2,2,3}}; \quad a_{3,1} = \frac{a_{2,4} - a_{2,3}}{b_{2,2,4} - b_{2,2,3}};$$

$$b_{3,1} = \frac{b_{2,3,4} - b_{2,3,3}}{b_{2,2,4} - b_{2,2,3}}; \quad b_{3,2} = \frac{b_{2,3,5} - b_{2,3,3}}{b_{2,2,5} - b_{2,2,3}}.$$

And finally, taking into account relations (4)–(6) and the first equation of the set (3), all polynomial coefficients of the set of equation (3) are defined with applying the following relations:

$$C_4 = \frac{a_{3,2} - a_{3,1}}{b_{3,2} - b_{3,1}}; \; C_3 = \frac{a_{2,4} - a_{2,3}}{b_{2,2,4} - b_{2,2,3}} - \frac{a_{3,2} - a_{3,1}}{b_{3,2} - b_{3,1}} \cdot \frac{b_{2,3,4} - b_{2,3,3}}{b_{2,2,4} - b_{2,2,3}};$$

$$C_2 = \frac{a_{1,3} - a_{1,2}}{z_3 - z_2} - \left(\frac{a_{3,2} - a_{3,1}}{b_{3,2} - b_{3,1}} - \frac{a_{2,4} - a_{2,3}}{b_{2,2,4} - b_{2,2,3}} - \frac{a_{3,2} - a_{3,1}}{b_{3,2} - b_{3,1}} \cdot \frac{b_{2,3,4} - b_{2,3,3}}{b_{2,2,4} - b_{2,2,3}} \right) \times \tag{7}$$

$$\times \frac{z_4^3 - z_2^3 + z_4^2 z_1 - z_2^2 z_1 + z_1^2 z_4 - z_1^2 z_2}{z_4 - z_2};$$

$$C_1 = a_{1,2} - C_4 (z_2^3 + z_2^2 z_1 + z_1^2 z_2 + z_1^3) - C_3 (z_2^2 + z_1 z_2 + z_1^2) - C_2 (z_1 + z_2);$$

$$C_0 = r_1^4 - C_4 z_1^4 - C_3 z_1^3 - C_2 z_1^2 - C_1 z_1.$$

The analytical relations for coefficients of fifth order root-polynomial function C_0, C_1, C_2, C_3, C_4 and C_5, corresponding to relation (1), is written as follows [22; 23]:

$$r(z)=\sqrt[5]{C_5z^4+C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\,. \qquad (8)$$

Therefore, the set of equation for defining the polynomial coefficient including 6 equations and generally it writing as follows [22; 23]:

$$\begin{cases} C_5z_1^5+C_4z_1^4+C_3z_1^3+C_2z_1^2+C_1z_1+C_0=r_1^5; \\ C_5z_2^5+C_4z_2^4+C_3z_2^3+C_2z_2^2+C_1z_2+C_0=r_2^5; \\ C_5z_3^5+C_4z_3^4+C_3z_3^3+C_2z_3^2+C_1z_3+C_0=r_3^5; \\ C_5z_4^5+C_4z_4^4+C_3z_4^3+C_2z_4^2+C_1z_4+C_0=r_4^5; \\ C_5z_5^5+C_4z_5^4+C_3z_5^3+C_2z_5^2+C_1z_5+C_0=r_5^5; \\ C_5z_6^5+C_4z_6^4+C_3z_6^3+C_2z_6^2+C_1z_6+C_0=r_6^5. \end{cases} \qquad (9)$$

But the advance of proposed method of calculation the polynomial coefficients is that with using the set of coefficients for four-order function, defined by relations (4)–(6), some of that relations are also correct for defining the coefficients of fifth-order polynomial. For example, among the first set of the coefficient a only the values $a_{1,l}$ are different form relations (4), since they are including fifth order of beam radius r. Corresponded relations for defining the coefficients $a_{1,l}$ are written as follows [22; 23]:

$$a_{1,2}=\frac{r_2^5-r_1^5}{z_2-z_1};\; a_{1,3}=\frac{r_3^5-r_1^5}{z_3-z_1};\; a_{1,4}=\frac{r_4^5-r_1^5}{z_4-z_1};\; a_{1,5}=\frac{r_5^5-r_1^5}{z_5-z_1}. \qquad (10)$$

Other two coefficients from the first set a_{26} and a_{36} are defined by the following analytical relations:

$$a_{2,6}=\frac{a_{1,6}-a_{1,2}}{z_6-z_2};\quad a_{3,6}=\frac{a_{2,6}-a_{2,3}}{b_{2,2,6}-b_{2,2,3}}. \qquad (11)$$

The corresponded coefficients b from the second set of additional variables are calculated for five order root-polynomial functions by analytical solving the set of linear equations (9) by the following relations:

$$b_{2,2,6}=\frac{z_6^2-z_2^2+z_6z_1-z_2z_1}{z_6-z_2};$$

$$b_{2,4,3}=\frac{z_3^4-z_2^2+z_3^3z_1-z_2^3z_1+z_3^2z_1^2-z_2^2z_1^2+z_1^3z_3-z_1^3z_2}{z_3-z_2};$$

$$b_{2,4,4}=\frac{z_4^4-z_2^4+z_4^3z_1-z_2^3z_1+z_4^2z_1^2-z_2^2z_1^2+z_1^3z_4-z_1^3z_2}{z_4-z_2};$$

$$b_{2,4,5}=\frac{z_5^4-z_2^4+z_5^3z_1-z_2^3z_1+z_5^2z_1^2-z_2^2z_1^2+z_1^3z_5-z_1^3z_2}{z_5-z_2};$$

$$b_{2,4,6}=\frac{z_6^4-z_2^3+z_6^2z_1-z_2^3z_1+z_1^2z_6-z_1^2z_2}{z_6-z_2}; \qquad (12)$$

$$b_{2,3,6}=\frac{z_6^3-z_2^3+z_6^2z_1-z_2^2z_1+z_1^2z_6-z_1^2z_2}{z_6-z_2};$$

$$b_{3,3,4}=\frac{b_{2,4,4}-b_{2,4,3}}{b_{2,2,4}-b_{2,2,3}};\quad b_{3,4,6}=\frac{b_{2,4,6}-b_{2,4,3}}{b_{2,2,6}-b_{2,2,3}};\quad b_{3,3,6}=\frac{b_{2,3,6}-b_{2,3,3}}{b_{2,2,6}-b_{2,2,3}}.$$

With known additional variables a and b, defined by the relations (4), (10)–(12), the polynomial coefficients of root-polynomial function (8) are calculated with using following relations:

$$C_5=\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}};\quad C_4=b_{4,4,5}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-a_{4,5};$$

$$C_3=a_{3,4}-b_{3,4,4}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-b_{3,4,4}\left(b_{4,4,5}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-a_{4,5}\right);$$

$$C_2=a_{2,3}-b_{2,4,3}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-b_{2,3,3}\left(b_{4,4,5}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-a_{4,5}\right)-$$

$$-b_{2,2,3}\left(a_{3,4}-b_{3,4,4}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-b_{3,4,4}\left(b_{4,4,5}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-a_{4,5}\right)\right);\quad (13)$$

$$C_1=a_{1,2}-b_{2,4,3}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}(z_2^4+z_2^3z_1+z_2^2z_1^2+z_1^3z_2+z_1^4)-$$

$$-\left(b_{4,4,5}\frac{a_{4,6}-a_{4,5}}{b_{4,4,6}-b_{4,4,5}}-a_{4,5}\right)(z_2^4+z_2^3z_1+z_2^2z_1^2+z_1^3z_2+z_1^4)-$$

$$-C_3(z_2^2+z_2z_1+z_1^2)-C_2(z_2+z_1);$$

$$C_0=r_1^5-C_5z_1^5-C_4z_1^4-C_3z_1^3-C_2z_1^2-C_1z_1.$$

Relations (4)–(7) have been used in this work for calculation the coefficients of forth order root-polynomial function (2), and relations (10)–(13) — for calculation the corresponded coefficients of fifth order root-polynomial function (8). Such kind of ravine functions are generally characterized by one minimum, as well as by quasi-linear dependence outside the region of local minimum. In any case, such functional dependences are very suitable for approximation the trajectories of electron beam, propagated in the medium of ionized gas with compensation the space charge of electrons, because, as it was proved theoretically, the behavior of electron beams in such physical conditions is exactly the same [1; 20–23; 26–29]. An effective and simple method of calculation the optimal values of polynomial coefficients for function (3) and (8), have been used in this work for solving the task of approximation the suitable numerical data. Describing of this method, as well as corresponded examples of approximation for some of ravine functions, will be considered in the next parts of this article.

STATEMENT OF APPROXIMATION PROBLEM

In general, the approximation task is that for given approximation basis points and a given approximation function $r(z)$, for example, for function (1) with unknown coefficients C_n, C_{n-1}, ... C_1, C_0, write an analytical expression based on the method of least squares [24; 25; 30; 31].

For example, for fourth order root-polynomial function:

$$S(C_4,C_3,C_2,C_1,C_0)=\sum_{i=1}^{m}(r_i^2-r^2(z_i,C_4,C_3,C_2,C_1,C_0))\Big|_{\min}= \qquad (14)$$
$$=\sum_{i=1}^{m}\left(r_i^2-\sqrt{C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\Bigg|_{\min},$$

and for fifth-order function, correspondently,

$$S(C_5,C_4,C_3,C_2,C_1,C_0)=\sum_{i=1}^{m}(r_i^2-r^2(z_i,C_5,C_4,C_3,C_2,C_1,C_0))\Big|_{\min}= \qquad (15)$$
$$=\sum_{i=1}^{m}\left(r_i-\sqrt[5]{(C_5z^5+C_4z^4+C_3z^3+C_2z^2+C_1z+C_0)^2}\right)\Bigg|_{\min},$$

where n is the degree of the root-polynomial function, m is the number of reference values.

Applying known methods of solving extremal problems through the search for partial derivatives of a function of many variables [24; 25], the generalized relation (14) can be rewritten in the form of a system of algebraic differential equations as follows [24; 25; 30; 31]:

$$\begin{cases}\sum_{i=1}^{m}\left(r_i-\sqrt[4]{C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_4,C_3,C_2,C_1,C_0)}{\partial C_0};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[4]{C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_4,C_3,C_2,C_1,C_0)}{\partial C_1};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[4]{C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_4,C_3,C_2,C_1,C_0)}{\partial C_2};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[4]{C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_4,C_3,C_2,C_1,C_0)}{\partial C_3};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[4]{C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_4,C_3,C_2,C_1,C_0)}{\partial C_4}.\end{cases} \qquad (16)$$

Correspondently, relation (15) for fifth order root-polynomial function is re-written as follows:

$$\begin{cases}\sum_{i=1}^{m}\left(r_i-\sqrt[5]{C_5z^5+C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_5,C_4,C_3,C_2,C_1,C_0)}{\partial C_0};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[5]{C_5z^5+C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_5,C_4,C_3,C_2,C_1,C_0)}{\partial C_1};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[5]{C_5z^5+C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_5,C_4,C_3,C_2,C_1,C_0)}{\partial C_2};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[5]{C_5z^5+C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_5,C_4,C_3,C_2,C_1,C_0)}{\partial C_3};\\ \sum_{i=1}^{m}\left(r_i-\sqrt[5]{C_5z^5+C_4z^4+C_3z^3+C_2z^2+C_1z+C_0}\right)\dfrac{\partial r(z_i,C_5,C_4,C_3,C_2,C_1,C_0)}{\partial C_4}.\end{cases} \qquad (17)$$

The problem is that the solution of the set of equations (16), (17) in the case of a nonlinear function $r(z)$ of many variable parameters, is extremely difficult. Methods of analytical solution of some simpler approximation problems for linear, quadratic, polynomial and one-parameter functions $f(z)$, as well as for the sum of arbitrarily specified functions $\varphi_1(z)$, $\varphi_2(z)$, $\varphi_n(z)$ with unknown numerical coefficients a_0, a_1, ..., a_n are described in the textbook [30; 31], and the methods of numerical solution of systems of nonlinear equations, similar to (4), are considered in textbooks [24; 25; 32–34]. But generally, in the theory of approximation is assumed, that with increasing the number of varied variables up to 5 and more the applying methods of multicriterial analyze aren't suitable and lead to obtaining the wrong results. Usually in mathematical software tools the gradient methods, the Nelder–Mead method, the Broyden–Fletcher–Goldfarb–Shanno algorithm and others are used for solving multi-criteria optimization tasks [32–35].

Let's we will the approximation task for root-polynomial functions (2), (8) by the other approach. As an initial approximation we will choose the result of interpolation for four base points using, to calculate the polynomial coefficients of the root-polynomial function (2), (8) by applying the analytical relations (4–7; 10–13).

Regarding hat the root-polynomial function of the fourth and fifth order (2), (8) is symmetric about the axis $z = z_{\min}$, considering now the linear approximation for the second and third branches of ravine function and find the corresponding angles of inclination of the tangents $k_{\mathrm{int}\,2}$ and $k_{\mathrm{int}\,3}$. The solution of the linear approximation problem is simple and well-known, the corresponding analytical relations are given in textbooks [30; 31]. For the second branch of interpolation, they are written as follows:

$$r_{B2}(x) = m_{r2}^* + \frac{\sum_{i=N_{B2}}^{N_{B3}} (z_i - m_{z2}^*)(r_i - m_{r2}^*)}{\sum_{i=N_{B2}}^{N_{B3}} (z - m_{z2}^*)^2}(z - m_{z2}^*)\,; \quad m_{z2}^* = \frac{\sum_{i=N_{B2}}^{N_{B3}} z_i}{N_{B3} - N_{B2} + 1}\,;$$

$$m_{r2}^* = \frac{\sum_{i=N_{B2}}^{N_{B3}} r_i}{N_{B3} - N_{B2} + 1}; \quad k_{\mathrm{int}\,2} = \frac{\sum_{i=N_{B2}}^{N_{B3}} (z_i - m_{z2}^*)(r_i - m_{r2}^*)}{\sum_{i=N_{B2}}^{N_{B3}} (z - m_{z2}^*)^2}\,, \tag{18}$$

and for the third branch:

$$r_{B3}(x) = m_{r3}^* + \frac{\sum_{i=N_{B3}}^{N_{End}} (z_i - m_{z3}^*)(r_i - m_{r3}^*)}{\sum_{i=N_{B3}}^{N_{End}} (z - m_{z3}^*)^2}(z - m_{z3}^*)\,; \quad m_{z3}^* = \frac{\sum_{i=N_{B3}}^{N_{End}} z_i}{N_{End} - N_{B3} + 1}\,;$$

$$m_{r3}^* = \frac{\sum_{i=N_{B3}}^{N_{End}} r_i}{N_{End} - N_{B3} + 1}; \quad k_{\mathrm{int}\,3} = \frac{\sum_{i=N_{B3}}^{N_{End}} (z_i - m_{z3}^*)(r_i - m_{r3}^*)}{\sum_{i=N_{B3}}^{N_{End}} (z - m_{z3}^*)^2}\,, \tag{19}$$

where N_{B2} the starting point of the second approximation branch, N_{B3} is the starting point of the third approximation branch, N_{End} is the end point of the data set for approximation region.

Taking into account equations (2), (18), (19), let's we rewrite the set of equations (16) to find the minimum of the regression function (2) as follows:

$$\begin{cases} \dfrac{dr(z_1)}{dz} = k_{\text{int}\,2}; & \dfrac{dr(z_2)}{dz} = k_{\text{int}\,3}; \\ r(z_3) = r_3; & r(z_4) = r_4; \\ r(z_5) = r_5. \end{cases} \tag{20}$$

Correspondently, to fifth order root polynomial function (8), one can rewrite the set of equations (17) as follows:

$$\begin{cases} \dfrac{dr(z_1)}{dz} = k_{\text{int}\,2}; & \dfrac{dr(z_2)}{dz} = k_{\text{int}\,3}; \\ r(z_3) = r_3; & r(z_4) = r_4; \\ r(z_5) = r_5; & r(z_5) = r_5. \end{cases} \tag{21}$$

The separate problem is finding the derivations for root-polynomial functions (2), (8) in the form of suitable polynomials for providing further iterative calculations. This task was solved in provided researches with applying the tools of symbolic calculation of the MatLab scientific and technical software [32]. Corresponded obtained results for taking a derivative of the function (2) is follows:

$$\begin{gathered} R_4(Cf_0, Cf_1, Cf_2, Cf_3, Cf_4, kf, zf) = Cf_1^4 zf^4 + 8Cf_1^3 Cf_2 zf^5 + 12Cf_1^3 Cf_3 zf^6 + \\ + 16Cf_1^3 Cf_4 zf^7 - Cf_1^3 kf^4 zf^3 + 24Cf_1^2 Cf_2^2 zf^6 + 72Cf_1^2 Cf_2 Cf_3 zf^7 + \\ + 96Cf_1^2 Cf_2 Cf_4 zf^8 - 3Cf_1^2 Cf_2 kf^4 zf^4 + 54Cf_1^2 Cf_3 zf^8 + 144Cf_1^2 Cf_3 Cf_4 zf^9 - \\ - 3Cf_1^2 Cf_3 kf^4 zf^5 + 96Cf_1^2 Cf_4^2 zf^{10} - 3Cf_1^2 Cf_4 kf^4 zf^6 + \\ + 222{,}9124Cf_1^2 kf^4 zf^2 + 32Cf_1 Cf_2^3 zf^7 + 144Cf_1 Cf_2^2 Cf_3 zf^8 + \\ + 192Cf_1 Cf_2^2 Cf_4 zf^9 - 3Cf_1 Cf_2^2 kf^4 zf^5 + 216Cf_1 Cf_2 Cf_3^2 zf^9 + \\ + 576Cf_1 Cf_2 Cf_3 Cf_4 zf^{10} - 6Cf_1 Cf_2 Cf_3 kf^4 zf^6 + 384Cf_1 Cf_2 Cf_2^4 zf^{11} - \\ - 6Cf_1 Cf_2 Cf_4 kf^4 zf^7 + 445{,}8249Cf_1 Cf_2 kf^4 zf^3 + 108Cf_1 Cf_3^3 zf^{10} + \\ + 432Cf_1 Cf_3^2 Cf_4 zf^{11} - 3Cf_1 Cf_3^2 kf^4 zf^7 + 576Cf_1 Cf_3 Cf_4^2 zf^{12} - \\ - 6Cf_1 Cf_3 Cf_4 kf^4 zf^8 + 445{,}8249Cf_1 Cf_3 kf^4 zf^4 + 256Cf_1 Cf_4^3 zf^{13} - \\ - 3Cf_1 Cf_4^2 kf^4 zf^9 + 445{,}8249Cf_1 Cf_4 kf^4 zf^5 + 1{,}6563 \cdot 10^4 Cf_1 kf^4 zf + \end{gathered} \tag{22}$$

$$+16Cf_2^4zf^8+96Cf_2^3Cf_3zf^9+128Cf_2^3Cf_4zf^{10}-Cf_2^3kf^4zf^6+$$
$$+216Cf_2^2Cf_3^2zf^{10}+576Cf_2^2Cf_3Cf_4zf^{11}-3Cf_2^2Cf_4kf^4zf^8+$$
$$+384Cf_2^2Cf_4^2zf^{12}+3Cf_2^2Cf_4kf^4zf^4+222{,}9124Cf_2^2kf^4zf^4+$$
$$+216Cf_2Cf_3^3zf^{11}+864Cf_2Cf_3^2Cf_4zf^{12}-3Cf_2Cf_3^2kf^4zf^8+$$
$$+1152Cf_2Cf_3Cf_4^2zf^{13}-6Cf_2Cf_3Cf_4kf^4zf^9+445{,}8249Cf_2Cf_3kf^4zf^5+$$
$$+512Cf_2Cf_4^3zf^{14}-3Cf_2Cf_4^2kf^4zf^{10}+445{,}8249Cf_2Cf_4kf^4zf^6+$$
$$+1{,}6563\cdot10^4Cf_2kf^4zf^2+81Cf_3^4zf^{12}+432Cf_3^3Cf_4zf^{13}+Cf_3^3kf^4zf^9+$$
$$+864Cf_3^2Cf_4^2zf^{14}-3Cf_3^2Cf_4^2kf^4zf^{10}+222{,}9124Cf_3^2kf^4zf^6+$$
$$+768Cf_3Cf_4^3zf^{15}-3Cf_3Cf_4^2kf^4zf^{11}+445{,}8249Cf_3Cf_4kf^4zf^7-$$
$$-1{,}6563\cdot10^4Cf_3kf^4zf^3+256Cf_4^4zf^{16}-Cf_4^3kf^4zf^{12}+$$
$$+222{,}9124Cf_4^2kf^4zf^8+1{,}6563\cdot10^4Cf_4kf^4zf^4+4{,}1024\cdot10^5kf^4.$$

For the derivative of fifth order root-polynomial ravine function (8) with using MatLab symbolic processor such polynomial expression have been obtained:

$$R_5(Cf_0,Cf_1,Cf_2,Cf_3,Cf_4,Cf_5,kf,zf)=Cf_1^5+10Cf_1^4Cf_2zf+15Cf_1^4Cf_3zf^2+$$
$$+20Cf_1^4Cf_4zf^3+25Cf_1^4Cf_5zf^4+40Cf_1^3Cf_2^2zf^2+120Cf_1^3Cf_2Cf_3zf^3+$$
$$+160Cf_1^3Cf_2Cf_4zf^4+200Cf_1^3Cf_2Cf_5zf^5+90Cf_1^3Cf_3^2zf^4+$$
$$+240Cf_1^3Cf_3Cf_4zf^5+300Cf_1^3Cf_3Cf_5zf^6+160Cf_1^3Cf_4^2zf^6+$$
$$+400Cf_1^3Cf_4Cf_5zf^7+250Cf_1^3Cf_5^2zf^8+80Cf_1^2Cf_2^3zf^3+$$
$$+360Cf_1^2Cf_2^3Cf_3zf^4+480Cf_1^2Cf_2^3Cf_4zf^5+600Cf_1^2Cf_2^3Cf_5zf^6+$$
$$+540Cf_1^2Cf_2^2Cf_3^2zf^5+1440Cf_1^2Cf_2Cf_3Cf_4zf^6+1800Cf_1^2Cf_2Cf_3Cf_5zf^7+$$
$$+960Cf_1^2Cf_2Cf_4^2zf^7+2400Cf_1^2Cf_2Cf_4Cf_5zf^8+1500Cf_1^2Cf_2Cf_5^2zf^9+$$
$$+270Cf_1^2Cf_3^3zf^6+1080Cf_1^2Cf_3^2Cf_4zf^7+1350Cf_1^2Cf_3^2Cf_5zf^8+$$
$$+1440Cf_1^2Cf_3Cf_4^2zf^8+3600Cf_1^2Cf_3Cf_4Cf_5zf^9+2250Cf_1^2Cf_3Cf_5^2zf^{10}+$$
$$+640Cf_1^2Cf_4^3zf^9+2400Cf_1^2Cf_4^2Cf_5zf^{10}+3000Cf_1^2Cf_4Cf_5^2zf^{11}+$$
$$+1250Cf_1^2Cf_5^3zf^{12}+80Cf_1Cf_2^4zf^4+480Cf_1Cf_2^3Cf_3zf^5+$$
$$+640Cf_1Cf_2^3Cf_4zf^6+800Cf_1Cf_2^3Cf_5zf^7+1080Cf_1Cf_2^2Cf_2^3zf^6+\qquad(23)$$
$$+2280Cf_1Cf_2^2Cf_3Cf_4zf^7+3600Cf_1Cf_2^2Cf_3Cf_5zf^8+1920Cf_1Cf_2^2Cf_4^2zf^8+$$
$$+4800Cf_1Cf_2^2Cf_4Cf_5zf^9+3000Cf_1Cf_2^2Cf_5^2zf^{10}+1080Cf_1Cf_2Cf_3^3zf^7+$$

$$+4320Cf_1Cf_2Cf_3^2Cf_4zf^8+5400Cf_1Cf_2Cf_3^2Cf_5zf^9+5760Cf_1Cf_2Cf_3Cf_4^2zf^9+$$
$$+14400Cf_1Cf_2Cf_3Cf_4Cf_5zf^{10}+9000Cf_1Cf_2Cf_3Cf_5^2zf^{11}+$$
$$+2560Cf_1Cf_2Cf_4^3zf^{10}+9600Cf_1Cf_2Cf_4^2Cf_5zf^{11}+12000Cf_1Cf_2Cf_4Cf_5^2zf^{12}+$$
$$+5000Cf_1Cf_2Cf_4Cf_5^3zf^{13}+405Cf_1Cf_2Cf_3^4zf^8+2160Cf_1Cf_3^3Cf_4zf^9+$$
$$+2700Cf_1Cf_3^3Cf_5zf^{10}+4320Cf_1Cf_3^2Cf_4^2zf^{10}+10800Cf_1Cf_3^2Cf_4Cf_5zf^{11}+$$
$$+6750Cf_1Cf_3^2Cf_5^2zf^{12}+3840Cf_1Cf_3Cf_4^3zf^{11}+14400Cf_1Cf_3Cf_4^2Cf_5zf^{12}+$$
$$+18000Cf_1Cf_3Cf_4Cf_5^2zf^{13}+7500Cf_1Cf_3Cf_5^3zf^{14}+1280Cf_1Cf_3Cf_4^4zf^{12}+$$
$$+6400Cf_1Cf_4^3Cf_5zf^{13}+1{,}2\cdot 10^4Cf_1Cf_4^2Cf_5^2zf^{14}+10^5Cf_1Cf_4Cf_5^3zf^{15}+$$
$$+3125Cf_1Cf_5^4zf^{16}-Cf_1kf^5zf+32Cf_2^5zf^5+240Cf_2^4Cf_3zf^6+$$
$$+320Cf_2^4Cf_4zf^7+400Cf_2^4Cf_5zf^8+720Cf_2^3Cf_3^2zf^7+1920Cf_2^3Cf_3Cf_4zf^8+$$
$$+2400Cf_2^3Cf_3Cf_5zf^9+1280Cf_2^3Cf_4^2zf^9+3200Cf_2^3Cf_4Cf_5zf^{10}+$$
$$+2000Cf_2^3Cf_5^2zf^{11}+1080Cf_2^2Cf_3^3zf^8+4320Cf_2^2Cf_3^2Cf_4zf^9+$$
$$+5400Cf_2^2Cf_3^2Cf_5zf^{10}+5760Cf_2^2Cf_3Cf_4zf^{10}+14400Cf_2^2Cf_3Cf_4Cf_5zf^{11}+$$
$$+9000Cf_2^2Cf_3Cf_5^2zf^{12}+2560Cf_2^2Cf_4^3zf^{11}+9600Cf_2^2Cf_4^2Cf_5zf^{12}+$$
$$+1{,}2\cdot 10^4Cf_2^2Cf_4Cf_5^2zf^{13}+5\cdot 10^3Cf_2^2Cf_5^3zf^{14}+810Cf_2Cf_3^4zf^9+$$
$$+4320Cf_2Cf_3^3Cf_2zf^{10}+5400Cf_2Cf_3^3Cf_5zf^{11}+8640Cf_2Cf_3^2Cf_4^2zf^{11}+$$
$$+21660Cf_2Cf_3^2Cf_4Cf_5zf^{12}+13500Cf_2Cf_3^2Cf_5^2zf^{13}+7680Cf_2Cf_3Cf_4^3zf^{12}+$$
$$+28800Cf_2Cf_3Cf_4^2Cf_5zf^{13}+36000Cf_2Cf_3Cf_4Cf_5^2zf^{14}+$$
$$+1{,}5\cdot 10^4Cf_2Cf_3Cf_5^3zf^{15}+2560Cf_2Cf_4^4zf^{13}+12{,}8\cdot 10^4Cf_2Cf_4^3Cf_5zf^{14}+$$
$$+2{,}4\cdot 10^4Cf_2Cf_4^2Cf_5^2zf^{15}+2\cdot 10^4Cf_2Cf_4Cf_5^3zf^{16}+6250Cf_2Cf_5^4zf^{17}-$$
$$-Cf_2kf^5zf^2+243Cf_3^5zf^{10}+1620Cf_3^4Cf_4zf^{11}+2025Cf_3^4Cf_5zf^{12}+$$
$$+4320Cf_3^3Cf_4^2zf^{12}+10800Cf_3^3Cf_4Cf_5zf^{13}+6750Cf_3^3Cf_5^2zf^{14}+$$
$$+5760Cf_3^2Cf_4^3zf^{13}+2{,}16\cdot 10^4Cf_3^2Cf_4^2Cf_5zf^{14}+2{,}7\cdot 10^4Cf_3^2Cf_4Cf_5zf^{15}+$$
$$+11250Cf_3^2Cf_5^3zf^{16}+3840Cf_3Cf_4^4zf^{14}+1{,}92\cdot 10^4Cf_3Cf_4^3Cf_5zf^{15}+$$
$$+3{,}6\cdot 10^4Cf_3Cf_4^2Cf_5^2zf^{16}+3\cdot 10^4Cf_3Cf_4Cf_5^3zf^{17}+9375Cf_3Cf_5^4zf^{18}-$$
$$-Cf_3kf^5zf^3+1024Cf_4^5zf^{15}+6400Cf_4^4Cf_5zf^{16}+1{,}6\cdot 10^4Cf_4^3Cf_5^2zf^{17}+$$
$$+2\cdot 10^4Cf_4^2Cf_5^3zf^{18}+12{,}5\cdot 10^4Cf_4Cf_5^4zf^{19}-Cf_4kf^5zf^4+$$
$$+3125Cf_5^5zf^{20}-Cf_5kf^5zf^5-Cf_0kf^5.$$

Obtained polynomial relations (22), (23) have been applied in provided investigation for iterative calculation the values of polynomial coefficient with using the method of consistent upper relaxation [33–36]. These relations included the values of the coefficients root-polynomial function (2), (8), as well as zf coordinate and slope angles of ravine function kf. The particularities of proposed algorithm, as well as some examples of solving approximation task, will be considered in the next sections of the article.

PARTICULARITIES OF DEVELOPED ITERATIVE ALGORITHM FOR SOLVING THE APPROXIMATION TASK

Iterative algorithm for approximation the ravine sets of numerical data by using root-polynomial dependences four and fifth order (2), (8) generally including the follow necessary steps.

1. As basic approach the interpolation task is solved and the basic values of polynomial coefficients are calculated withusing relations (4)–(7) for fourth order function (2), and by the relations (10)–(13) for fifth order function (8).

2. Solving the relaxation task for the finding values of polynomial coefficient with using iterative algorithm. Corresponded iterative relations for function (2), taking into account (22), are the follows:

$$\begin{gathered}
C_1^i = -R_4(C_0^{i-1}, C_1^{i-1}, C_2^{i-1}, C_3^{i-1}, C_4^{i-1}, kf_1, zf_3) w_{C_1}; \\
C_3^i = R_4(C_0^{i-1}, C_1^i, C_2^{i-1}, C_3^{i-1}, C_4^{i-1}, kf_2, zf_4) w_{C_3}; \\
C_4^i = ((r_2^4 - C_3^i zf_2^3 - C_2^{i-1} zf_2^2 - C_1^i zf_2 - C_0^{i-1}) w_{C_4}) \big/ zf_2^4; \\
C_2^i = ((r_1^4 - C_4^i zf_1^4 - C_2^{i-1} zf_1^2 - C_1^i zf_1 - C_0^{i-1}) w_{C_2}) \big/ zf_1^2; \\
C_0^i = (r_1^4 - C_4^i zf_1^4 - C_2^i zf_1^2 - C_1^i zf_1 - C_0^{i-1}) w_{C_0}.
\end{gathered} \tag{24}$$

For fifth order root-polynomial function (8), taking into account (23), the iterative relations, in the general form, are rewritten as follows:

$$\begin{gathered}
C_1^i = -R_5(C_0^{i-1}, C_1^{i-1}, C_2^{i-1}, C_3^{i-1}, C_4^{i-1}, C_5^{i-1}, kf_1, zf_3) w_{C_1}; \\
C_3^i = R_5(C_0^{i-1}, C_1^i, C_2^{i-1}, C_3^{i-1}, C_4^{i-1}, C_5^{i-1}, kf_2, zf_4) w_{C_3}; \\
C_5^i = ((r_6^5 - C_4^{i-1} zf_6^4 - C_3^i zf_6^3 - C_2^{i-1} zf_6^2 - C_1^i zf_6 - C_0^{i-1}) w_{C_5}) \big/ zf_6^5; \\
C_4^i = ((r_2^5 - C_5^i zf_2^5 - C_3^i zf_2^3 - C_2^{i-1} zf_2^2 - C_1^i zf_6 - C_0^{i-1}) w_{C_4}) \big/ zf_2^4; \\
C_2^i = ((r_1^5 - C_5^i zf_1^5 - C_4^i zf_1^4 - C_3^i zf_1^3 - C_1^i zf_1 - C_0^{i-1}) w_{C_2}) \big/ zf_1^2; \\
C_0^i = (r_5^5 - C_5^i zf_5^5 - C_4^i zf_5^4 - C_3^i zf_5^3 - C_2^i zf_5^3 - C_1^i zf_5) w_{C_0}.
\end{gathered} \tag{25}$$

3. Finding the best function of approximation by calculation the error with using relations (14), (15) for least-square method.

4. Finding the new optimal values of polynomial coefficients wit using dichotomy calculations, namely:

$$C_l^i = \frac{C_l^{i-1} - C_l^{i-2}}{2}, \qquad (26)$$

where l — the number of polynomial coefficient, defined in the basic relations (4)–(6), (10)–(13).

As shown the tests calculation experiments, the convergence of proposed algorithm is provided by 4–6 iterations. It also must be taking into account, that first iteration for high order root-polynomial function, which have been obtained by interpolation through $n + 1$ points, where n is the polynomial order, is really usually close to optimal. The best solution has been chosen by the smallest value of sum in relation (14), (15), therefore the well-known least-square method is considered in this research as the criterium of optimization task [30–39]. Choosing of basic points is also very important, therefore considering the different sets of its for finding the best approximative root-polynomial function is also have been provided. Analyzing the interpolation error of the different order root-polynomial functions in dependence on choosing the set of basic points have been provided generally in the papers [22; 23].

And finally, choose the correct values of $w_{C_0} - w_{C_5}$ relaxation coefficient is very important problem for providing the stability of convergence of proposed iteration algorithm, because the value problem of coefficients $C_0 - C_5$ in the relations (24)–(26) are usually, for the practice engineering tasks, can be both extra low or extra high.

Some examples of solving approximation for the ravine data sets with using root-polynomial functions (2), (8) will be considered in the next part of the article.

SOME TESTING EXAMPLES OF SOLVING THE APPROXIMATION TASK

Example 1. Find the coefficients of fourth order approximative root polynomial function for ravine function data set with one global minimum. Corresponded digital data are presented at Table 1.

Table 1. The first set of numerical data of the ravine function, have been used for testing the software tools for solving the approximation problem

z, mm	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
r, mm	2.5	2.4	2.3	2.1	1.8	1.75	2.11	2.3	2.4

The obtained graphic results of solving the approximation task of this example with using relations (24), (26) are presented at Fig 1.

Corresponded values of relaxation coefficient for providing the convergence of iterative relations (24) have been defined as follows:

$$w_{C_1} = -1.7995 \cdot 10^6; \; w_{C_3} = 1.9163 \cdot 10^{-5}; \; w_{C_4} = 1; \; w_{C_2} = -1.131; \; w_{C_0} = 0.945.$$

The values of calculated polynomial coefficients are presented at Table 2. The error of approximation δ, defined by the equation (14) as sum of squares of differences between basic points and value4 of approximative function (2), also noted in this table. The relative error of approximation is calculated as $\delta[\%] = r_{max}^2 (S/()) \cdot 100$, де r_{max} — the maximum value of electron beam radius.

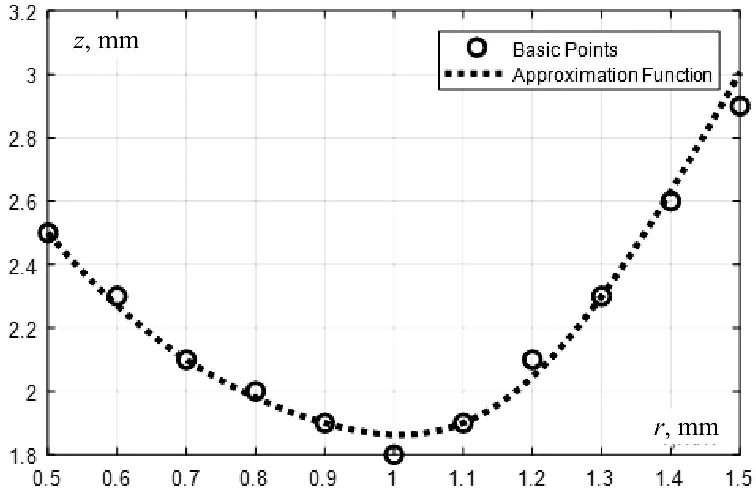


Fig. 1. The results of solving the approximation task with using fourth order root-polynomial function (2) for data set, presented at Table 1. Corresponded values of polynomial coefficients and the approximation error are given at Table 2

Table 2. The results of solving the approximation task for numerical data, presented in the Table 1

Nubmer of iteration	Values of polynomial coefficients					δ, %
	C_0	C_1	C_2	C_3	C_4	
Basic Approach	–493.16	2750.4	4976.8	3690.6	–961.7	8
First iteration	–493.179	2750.41	4976.81	3686.9	–956.2	5
Second iteration	–493.172	2750.408	4976.81	3688.7	–958.961	4
Third iteration	–493.17	2750.408	4976.7	3689.68	–960.33	3

In this example first iteration by the polynomial coefficients have been realized with applying relations (24), (25), and the second and third iterations, in which have been obtained the best solution of approximation task, have been provided with using relation (26).

Example 2. Find the coefficients of fourth order approximative root-polynomial function for ravine function data set with one global minimum. Corresponded digital data are presented in Table 3.

Table 3. The second set of numerical data of the ravine function, have been used for testing the software tools for solving the approximation problem

z, mm	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
r, mm	2.5	2.3	2.1	2.0	1.9	1.8	1.9	2.1	2.3	2.6	2.9

The obtained graphic results of solving the approximation task of this example with using one iteration step are presented at Fig 2. Really the task of interpolation by choose 5 basis points among the 11 given have been solved in this case, and the interpolation error was smaller, than 2%. The calculated values of polynomial coefficients for this test example are presented in Table 4.

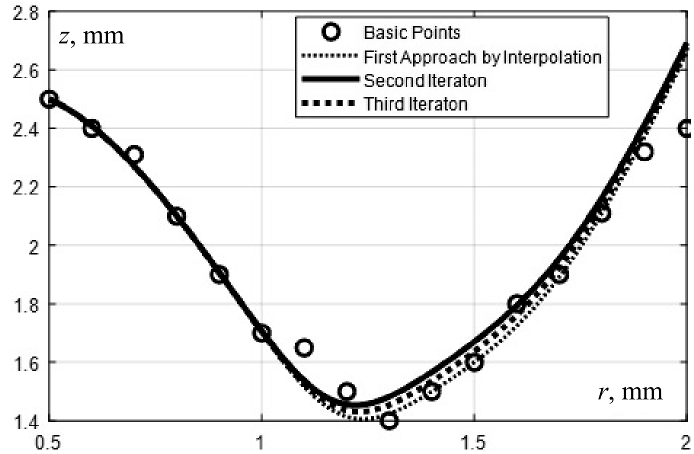


Fig. 2. The results of solving the approximation task with using fourth order root-polynomial function (2) for data set, presented at Table 3. Corresponded values of polynomial coefficients and the approximation error are given at Table 4

Table 4. The results of solving the approximation task for numerical data, presented in the Table 3

Nubmer of iteration	**Values of polynomial coefficients**					δ, %
	C_0	C_1	C_2	C_3	C_4	
Basic Approach	328.6	–1251.26	1953.66	–1417.8	398.9	1.5

Example 3. Find the coefficients of fifth order approximative root polynomial function for ravine function data set with one global minimum. Corresponded digital data are presented at Table 5.

Table 5. The third set of numerical data of the ravine function, have been used for testing the software tools for solving the approximation problem

z, mm	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2
r, mm	2.5	2.4	2.31	2.1	1.9	1.7	1.65	1.63
z, mm	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
r, mm	1.6	1.55	1.7	1.8	1.9	2.11	2.32	2.4

In this and in the next example the corresponded values of relaxation coefficient for providing the consvergence of iterative relations (25) have been defined as follows:

$$w_{C_1} = -5{,}53 \cdot 10^{-7}; \; w_{C_3} = 9{,}5 \cdot 10^{-7}; \; w_{C_5} = 1; \; w_{C_4} = 1; \; w_{C_2} = 0{,}39; \; w_{C_0} = 0{,}145.$$

The obtained graphic results of solving the approximation task of this example with using one iteration step are presented at Fig 3. It is clear from this example, that considered type of the root-polynomial functions is not very suitable for approximation the ravine data sets with large area of minimum numerical values. Corresponded approximation error was greater than 12 %. But for the left and right branches of considered data set the error of approximation is much smaller, nearly few percents.

Example 4. Find the coefficients of fifth order approximative root polynomial function for ravine function data set with one global minimum. Corresponded digital data are presented at Table 6.

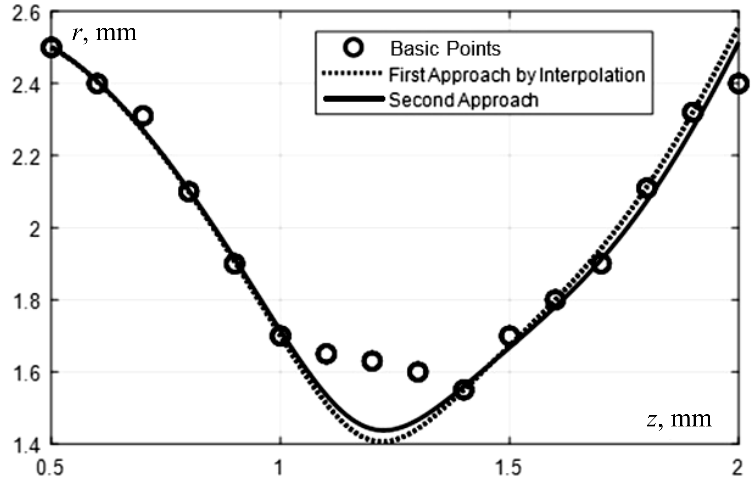


Fig. 3. The results of solving the approximation task with using fifth order root-polynomial function (8) for data set, presented at Table 5. Corresponded approximation error for the left and right branches of considered ravine function was smaller, than few percent

Table 6. The fourth set of numerical data of the ravine function, have been used for testing the software tools for solving the approximation problem

z, mm	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2
r, mm	2.5	2.4	2.31	2.1	1.9	1.7	1.65	1.5
z, mm	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
r, mm	1.4	1.5	1.6	1.8	1.9	2.11	2.32	2.4

The obtained graphic results of solving the approximation task of this example with using 3 iteration step are presented at Fig 4.

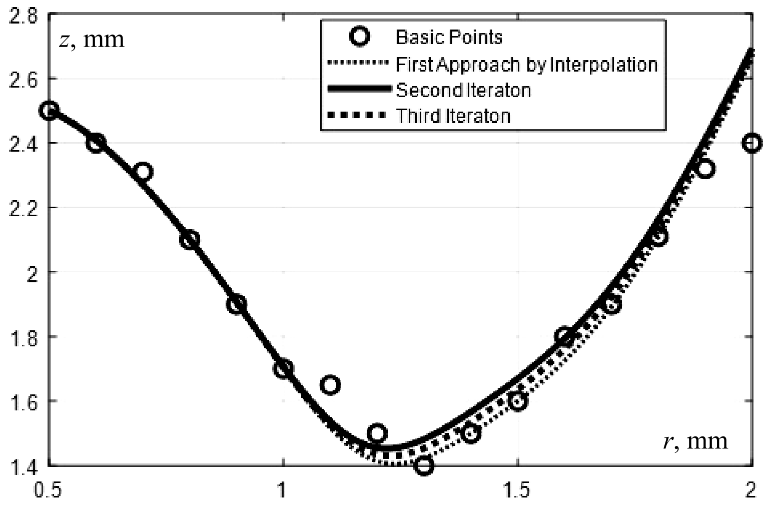


Fig. 4. The results of solving the approximation task with using fifth order root-polynomial function (8) for data set, presented at Table 6. Corresponded values of the polynomial coefficients and approximation error are given in Table 7

T a b l e 7 . The results of solving the approximation task for numerical data, presented in the Table 6

Nubmer of iteration	Values of polynomial coefficients						δ, %
	C_0	C_1	C_2	C_3	C_4	C_5	
Basic Approach	–231.51	2340.2	–5259.9	5083.15	–2278.7	392.01	5
First iteration	–39.32	2340.1	–5260.0	5082.7	–2276.1	392.1	4
Second iteration	–151.0	2340.0	–5259.9	5083.0	–2278.2	392.1	3

In this example, the iterative process was carried out in the same way as in example 1. The first values of polynomial coefficients were obtained by solving interpolation task. Therefore, for defining the coefficients of the root-polynomial function (8) relations (10)–(13) was applied. In this step of solving approximation task 5 basis points have been choose among the 16 given. By the such way the basic the approach for calculation the polynomial coefficients have been provided. At the second iteration the values of polynomial coefficient have been calculated iteratively with using relations (25), and in the third iteration — with using relation (26). Corresponded results of for the coefficients of fifth order root-polynomial function, as well as the estimated value of approximation error, have been presented in the Table 7.

ANALYZING OF OBTAINED RESULTS AND ITS' DISCUSSION

The provided numerical experiments for solving the approximation task for ravine data set with using high order root-polynomial functions (2), (8) shown, that generally by providing simple iterative process with relaxation by using relations (24), (25) is really possible to find the optimal correct values of polynomial coefficient. Estimated theoretically approximation error in most cases was in range of few percent, and, taking into account, that usually approximated experimental data for the trajectories of electron beams are included the large amount instrumental errors, such estimations, in the most cases, are generally useful from the practical point of view.

The main distinguishing feature of proposed algorithm is possibilities of obtaining the correct results for approximation task just on the first step of calculations by solving the simpler interpolation task. Such approach is very effective on the practical point of view for developing corresponded software for simulation the boundary trajectories of electron beams.

The provided testing experiment also shown, that root-polynomial functions of high order, like (8), (12), are generally not suitable for approximation the ravine data sets with the large area near the minimum, but such kind of functions isn't corresponded to the trajectories of electron beam in standard physical conditions. Usually, the region of beam focus is relatively small, and, as shown the provided numerical experiments, such data sets can be approximated by the high order root-polynomial functions with small error.

The provided theoretical analyze also shown, that for obtaining the convergence of proposed iterative approximation algorithm correct choosing of relaxation coefficients $w_{C_0} - w_{C_5}$ is very important numerical factor.

Analyzing of another advanced possibilities of applying high order root-polynomial functions for accurate approximation of ravine data sets and forming the corresponded algorithms is the subject of further theoretical researches.

But, generally, the complicity of proposed iterative algorithm is connected only with solving the strong non-linear task for derivatives of ravine functions (2), (8) and to numerical solving the complex polynomial relations (22), (23). The provided numerical experiments shown, that the polynomial coefficient, which are calculated throw the derivatives by numerical solving the relations (22), (23), are strongly depended on relaxation parameters and the values of these parameters are magnificently grater, than for the polynomial coefficients, which are calculated independently by using simple relations (2), (8). But it is also clear from the provided testing numerical experiments, that the approximation error is mostly defined by the polynomial coefficients, which values are calculated through the derivations. And the values of coefficients, calculated directly by the relations (2), (8), are generally stable and corresponded relaxation parameter is in the range of medium value, it is not very large and not so small. But since usually ravine functions in the region of a local minimum are non-linear, it is often complex non-linear equations for derivatives, like (22), (23), are solved. Its makes possible to ensure, that the smallest error and the optimal values of the polynomial coefficients for solving the approximation problem have been choose.

Choosing the optimal values of relaxation parameters for calculation the polynomial coefficients throw the derivatives is also the subject of further theoretical researches.

CONCLUSION

The theoretical researches, have been provided and described in this work, as well as realized test numerical experiments, given in the corresponded examples, have shown, that applying of high order root-polynomial functions, which have been determinate by the analytical relations (2), (8), always leads to solving the task of approximation the numerical sets of ravine functions data with the small value of error. Therefore, such functions can be used successfully for approximation the boundary trajectories of electron beams, propagated in the ionized gas with compensation the space charge of beam electrons. Such approximation can also be applied to the noisy experimental data, which included experimental errors. Therefore, studied and proposed in these researches the theoretical approach and the numerical algorithm of the approximation of digital data are very important from the practical point of view for further development of modern industrial electron beam equipment with applying HVGD electron gun as the source of intensive beams. For solving this task proposed iterative algorithm has to be integrated with the software tools for treatment of experimental photographs, where the beam propagates through ionized gas. Analyzing of the brightness of a burning discharge on such photographs gives the important experimental information about the boundary trajectories of propagated electron beam in the real physical conditions. [1].

Advanced singularities of proposed iterative algorithm, including the possibilities of approximation the ravine data sets with a large area of the global minimum, as well as studying of influence of the derivative of root-polynomial functions on the values of relaxation coefficients, are the subject of further theoretical researches. But, in any case, the results of provided theoretical researches, presented in this article, are enough for estimation the boundary trajectories of electron beams, propagated in ionizing gas. Therefore, theoretical results, which have already been obtained, are very interesting and important for experts in the branch of elaboration of modern electron-beam equipment and its industrial application.

REFERENCES

1. I. Melnyk, S. Tuhai, and A. Pochynok, "Universal Complex Model for Estimation the Beam Current Density of High Voltage Glow Discharge Electron Guns," *Lecture Notes in Networks and Systems*, vol. 152, pp. 319–341, 2021. Available: https://www.springer.com/gp/book/9783030583583
2. S.V. Denbnovetsky, V.G. Melnyk, and I.V. Melnyk, "High voltage glow discharge electron sources and possibilities of its application in industry for realising of different technological operations," *IEEE Transactions on Plasma Science*, vol. 31, no. 5, pp. 987–993, 2003. Available: https://ieeexplore.ieee.org/document/1240048
3. S. Denbnovetskiy et al., "Principles of operation of high voltage glow discharge electron guns and particularities of its technological application," *Proceedings of SPIE, The International Society of Optical Engineering*, pp. 10445–10455, 2017. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10445/104455R/Principles-of-operation-of-high-voltage-glow-discharge-electron-guns/10.1117/12.2280736.short
4. M. Grechanyuk, A. Melnyk, I. Grechanyuk, V. Melnyk, and D. Kovalchuk, "Modern electron beam technologies and equipment for melting and physical vapor deposition of different materials," *Electrotechnics and Electronics* (E+E), vol. 49, no. 5–6, pp. 115–121, 2014.
5. T. Hassel, N. Beniyash, N. Murray, R. Konya, and Fr.-W. Bach, "Non-vacuum electron beam for cutting application," *Electrotechnics and Electronics (E+E)*, vol. 47, no. 5–6, pp. 146–151, 2012.
6. V. Vassilieva, K. Vutova, and V. Donchev, "Recycling of alloy steel by electron beam melting," *Electrotechnics and Electronics (E+E)*, vol. 47, no. 5–6, pp. 142–145, 2012.
7. G. Mattausch et al., "Gas discharge electron sources – proven and novel tools for thin-film technologies," *Electrotechnics and Electronics (E+E)*, vol. 49, no. 5–6, pp. 183–195, 2014.
8. S.V. Denbnovetsky, V.I. Melnyk, I.V. Melnyk, and B.A. Tugay, "Model of control of glow discharge electron gun current for microelectronics production applications," *Proceedings of SPIE. Sixth International Conference on "Material Science and Material Properties for Infrared Optoelectronics"*, vol. 5065, pp. 64–76, 2003. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/5065/0000/ Model-of-control-of-glow-discharge-electron-gun-current-for/10.1117/12.502174.short

9. I.V. Melnyk, "Estimating of current rise time of glow discharge in triode electrode system in case of control pulsing," *Radioelectronic and Communication Systems*, vol. 56, no. 12, pp. 51–61, 2017. Available: http://radioelektronika.org/article/view/S0735272713120066
10. A.A. Druzhinin, I.P. Ostrovskii, Y.N. Khoverko, N.S. Liakh-Kaguy, and A.M. Vuytsyk, "Low temperature characteristics of germanium whiskers," *Functional Materials*, no. 21(2), pp. 130–136, 2014. Available: https://nanoscalereslett.springeropen.com/articles/10.1186/s11671-017-1923-1
11. A.A. Druzhinin, I.A. Bolshakova, I.P. Ostrovskii, Y.N. Khoverko, and N.S. Liakh-Kaguy, "Low temperature magnetoresistance of InSb whiskers," *Materials Science in Semiconductor Processing*, no. 40, pp. 550–555, 2015. Available: https://academic-accelerator.com/search?Journal=Druzhinin
12. A. Zakharov, S. Rozenko, S. Litvintsev, and M. Ilchenko, "Trisection Bandpass Filter with Mixed Cross-Coupling and Different Paths for Signal Propagation," *IEEE Microwave Wireless Component Letters*, vol. 30, no. 1, pp. 12–15, 2020.
13. A. Zakharov, S. Litvintsev, and M. Ilchenko, "Trisection Bandpass Filters with All Mixed Couplings," *IEEE Microwave Wireless Components Letter*, vol. 29, no. 9, pp. 592–594, 2019. Available: https://ieeexplore.ieee.org/abstract/document/8782802
14. A. Zakharov, S. Rozenko, and M. Ilchenko, "Varactor-tuned microstrip bandpass filter with loop hairpin and combline resonators," *IEEE Transactions on Circuits Systems. II. Experimental Briefs*, vol. 66, no. 6, pp. 953–957, 2019. Available: https://ieeexplore.ieee.org/document/8477112
15. T. Kemmotsu, T. Nagai, and M. Maeda, "Removal Rate of Phosphorous form Melting Silicon," *High Temperature Materials and Processes*, vol. 30, no. 1–2, pp. 17–22, 2011. Available: https://www.degruyter.com/journal/key/htmp/30/1-2/html
16. J.C.S. Pires, A.F.B. Barga, and P.R. May, "The purification of metallurgically grade silicon by electron beam melting," *Journal of Materials Processing Technology*, vol. 169, no. 1, pp. 347–355, 2005. Available: https://www.academia.edu/9442020/The_purification_of_metallurgical_grade_silicon_by_electron_beam_melting
17. D. Luo, N. Liu, Y. Lu, G. Zhang, and T. Li, "Removal of impurities from metallurgically grade silicon by electron beam melting," *Journal of Semiconductors*, vol. 32, no. 3, Article ID 033003, 2011. Available: http://www.jos.ac.cn/en/article/doi/10.1088/ 1674-4926/ 32/3/033003
18. D. Jiang, Y. Tan, S. Shi, W. Dong, Z. Gu, and R. Zou, "Removal of phosphorous in molten silicon by electron beam candle melting," *Materials Letters*, vol. 78, pp. 4–7, 2012.
19. I.V. Melnyk, "Numerical simulation of distribution of electric field and particle trajectories in electron sources based on high-voltage glow discharge," *Radioelectronic and Communication Systems*, vol. 48, no. 6, pp. 61–71, 2005. Available: http://radioelektronika.org/article/view/S0735272705060087
20. S.V. Denbnovetsky, J. Felba, V.I. Melnik, and I.V. Melnik, "Model of Beam Formation in a Glow Discharge Electron Gun With a Cold Cathode," *Applied Surface Science*, vol. 111, pp. 288–294, 1997. Available: https://www.sciencedirect.com/ science/article/pii/S0169433296007611?via%3Dihub
21. J.I. Etcheverry, N. Mingolo, J.J. Rocca, and O.E. Martınez, "A Simple Model of a Glow Discharge Electron Beam for Materials Processing," *IEEE Transactions on Plasma Science*, vol. 25, no. 3, pp. 427–432, 1997.
22. I. Melnyk, S. Tuhai, and A. Pochynok, "Interpolation of the Boundary Trajectories of Electron Beams by the Roots from Polynomic Functions of Corresponded Order," *2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO). Conference Proceedings*, pp. 28–33. Available: https://ieeexplore.ieee.org/ servlet/opac?punumber=9085228
23. I. Melnik, S. Tugay, and A. Pochynok, "Interpolation Functions for Describing the Boundary Trajectories of Electron Beams Propagated in Ionised Gas," *15-th Interna-*

tional Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET – 2020) Conference Proceedings, pp. 79–83, 2020. Available: https://ieeexplore.ieee.org/document/9088609

24. G.M. Phillips, *Interpolation and Approximation by Polynomials*. Springer, 2023, 312 p. Available: http://bayanbox.ir/view/2518803974255898294/George-M.-Phillips-Interpolation-and-Approximation-by-Polynomials-Springer-2003.pdf
25. N. Draper and H. Smith, *Applied Regression Analysis*; 3rd Edition. Wiley Series, 1998, 706 p. Available: https://www.wiley.com/en-us/Applied+Regression+ Analysis,+3rd+Edition-p-9780471170822
26. J.D. Lawson, *The Physics of Charged-Particle Beams*. Oxford: Clarendon Press, 1977, 446 p. Available: https://www.semanticscholar.org/paper/The-Physics-of-Charged-Particle-Beams-Stringer/80b5ee5289d5efd8f480b516ec4bade0aa529ea6
27. M. Reiser, *Theory and Design of Charged Particle Beams*. John Wiley & Sons, 2008, 634 p. Available: https://www.wiley.com/en-us/Theory+and+Design+of+Charged +Particle +Beams-p-9783527617630
28. M. Szilagyi, *Electron and Ion Optics*. Springer Science & Business Media, 2015, 539 p. Available: https://www.amazon.com/Electron-Optics-Microdevices-Miklos-Szilagyi/dp/1461282470
29. S.J.R. Humphries, *Charged Particle Beams*. Courier Corporation, 2013, 834 p. Available: https://library.uoh.edu.iq/admin/ebooks/76728-charged-particle-beams---s.-humphries.pdf
30. E. Wentzel and L. Ovcharov, *Applied Problems of Probability Theory*. Mir, 1998, 432 p. Available: https://mirtitles.org/2022/06/03/applied-problems-in-probability-theory-wentzel-ovcharov/
31. J.A. Gubner, *Probability and random processes for electrical and computer engineers*. Cambridge, UK: Cambridge University Press, 2006. Available: http://www.cambridge.org/gb/academic/subjects/engineering/communications-and-signal-processing/probability-and-random-processes-electrical-and-computer-engineers
32. J.H. Mathews and K.D. Fink, *Numerical Methods Using MATLAB*; 3rd Edition. Amazon, 1998, 720 p. Available: https://www.abebooks.com/book-search/title/numerical-methods-using-matlab/author/john-mathews-kurtis-fink/?cm_mmc=ggl-_-COMUS_ETA_DSA-_-naa-_-naa&gclid=CjwKCAiAh9qdBhAOEiwAvxIok6hZ7XHTvi420qugGwqNZ20QF4PyaaJai-74Z0EK2c3dbVRqo1P17hoCP2wQAvD_BwE
33. C. Mohan and K. Deep, *Optimization Techniques*. New Age Science, 2009, 628 p. Available: https://www.amazon.com/Optimization-Techniques-C-Mohan/dp/ 1906574219
34. M.K. Jain, S.R.K. Iengar, and R.K. Jain, *Numerical Methods for Scientific & Engineering Computation*. New Age International Pvt. Ltd., 2010, 733 p. Available: https://www.google.com.ua/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwippcuT7rX8AhUhlYsKHRfBCG0QFnoECEsQAQ&url=https%3A%2F%2Fwww.researchgate.net%2Fprofile%2FAbiodun_Opanuga%2Fpost%2Fhow_can_solve_a_non_linear_PDE_using_numerical_method%2Fattachment%2F59d61f7279197b807797de30%2FAS%253A284742038638596%25401444899200343%2Fdownload%2FNumerical%2BMethods.pdf&usg=AOvVaw0MjNl3K877lVWUWw-FPwmV
35. S.C. Chapra and R.P. Canale, *Numerical Methods for Engineers*; 7th Edition. McGraw Hill, 2014, 992 p. Available: https://www.amazon.com/Numerical-Methods-Engineers-Steven-Chapra/dp/007339792X
36. I.N. Bronshtein, K.A. Semendyayev, G. Musiol, and H. Mühlig, *Handbook of Mathematics*; 5th Edition. Springer, 2007, 1164 p.

Received 02.02.2023

INFORMATION ON THE ARTICLE

Igor V. Melnyk, ORCID: 0000-0003-0220-0615, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: imelnik@phbme.kpi.ua

Alina V. Pochynok, ORCID: 0000-0001-9531-7593, Research Institute of Electronics and Microsystem Technology of the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Ukraine, e-mail: alina_pochynok@yahoo.com

БАЗОВИЙ АЛГОРИТМ АПРОКСИМАЦІЇ ГРАНИЧНОЇ ТРАЄКТОРІЇ КОРОТКОФОКУСНОГО ЕЛЕКТРОННОГО ПУЧКА ЗА ДОПОМОГОЮ КОРЕНЕВО-ПОЛІНОМІАЛЬНИХ ФУНКЦІЙ ЧЕТВЕРТОГО ТА П'ЯТОГО ПОРЯДКІВ / І.В. Мельник, А.В. Починок

Анотація. Розглянуто новий ітераційний метод апроксимації граничної траєкторії короткофокусного електронного пучка, який поширюється в режимі вільного дрейфу в іонізованому газі низького тиску за умови компенсації просторового заряду електронів. Використано коренево-поліноміальні функції четвертого та п'ятого порядків, головними особливостями яких є яружний характер та наявність одного глобального мінімуму. Як початкове наближення для розв'язування апроксимаційної задачі розраховано значення поліноміальних коефіцієнтів через розв'язання задачі інтерполяції. Задачу апроксимації розв'язано ітераційно. Для цього поліноміальні коефіцієнти обчислено багаторазово з урахуванням значень функції та її похідної у відлікових точках. Остаточні значення поліноміальних коефіцієнтів коренево-поліноміальних функцій високого порядку розраховано з використанням методу дихотомії. Наведено приклади використання коренево-поліноміальних функцій четвертого та п'ятого порядків для апроксимації наборів числових даних, які відповідають опису яружних функцій. Отримані теоретичні результати є цікавими та корисними для спеціалістів, які вивчають фізику електронних пучків та займаються проектуванням сучасного промислового електронно-променевого технологічного обладнання.

Ключові слова: апроксимація, інтерполяція, кореневo-поліноміальна функція, яружна функція, метод найменших квадратів, нев'язка, похибка апроксимації, електронний пучок, електронно-променеві технології.

ВІДОМОСТІ ПРО АВТОРІВ

Бідюк Петро Іванович,

професор, доктор технічних наук, професор кафедри математичних методів системного аналізу ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Бодянський Євгеній Володимирович,

професор, доктор технічних наук, професор кафедри штучного інтелекту Харківського національного університету радіоелектроніки, Україна, Харків

Гаврилович Марія Павлівна,

аспірантка ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Гуськова Віра Геннадіївна,

доктор філософії з комп'ютерних наук, асистент кафедри математичних методів системного аналізу ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Данилов Валерій Якович,

професор, доктор технічних наук, професор кафедри математичних методів системного аналізу ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Зайченко Олена Юріївна,

професор, доктор технічних наук, професор кафедри математичних методів системного аналізу ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Зайченко Юрій Петрович,

професор, доктор технічних наук, професор кафедри математичних методів системного аналізу ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Згуровський Михайло Захарович,

академік НАН України, професор, доктор технічних наук, ректор КПІ ім. Ігоря Сікорського, Україна, Київ

Кузьменко Олексій Віталійович,

аспірант ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Левенчук Людмила Борисівна,

асистент кафедри математичних методів системного аналізу ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Ліп'яніна-Гончаренко Христина Володимирівна,

доцент, кандидат технічних наук, доцент кафедри інформаційно-обчислювальних систем і управління Західноукраїнського національного університету, Україна, Тернопіль

Мацукі Йошіо,

професор, доктор технічних наук, професор Національного університету «Києво-Могилянська академія», Україна, Київ

Мельник Ігор Віталійович,

професор, доктор технічних наук, професор кафедри електронних пристроїв та систем факультету електроніки КПІ ім. Ігоря Сікорського, Україна, Київ

Починок Аліна Володимирівна,

кандидат технічних наук, старший науковий співробітник Науково-дослідного інституту електроніки та мікросистемної техніки КПІ ім. Ігоря Сікорського, Україна, Київ

Саченко Анатолій Олексійович,

професор, доктор технічних наук, директор НДІ інтелектуальних комп'ютерних систем Західноукраїнського національного університету, Україна, Тернопіль

Тимощук Оксана Леонідівна,

доцент, кандидат технічних наук, завідувачка кафедри математичних методів системного аналізу ІПСА КПІ ім. Ігоря Сікорського, Україна, Київ

Dipa D. Dharmadhikari,

Research Scholar, Department of CSIT Dr. Babasaheb Ambedkar Marathwada University, India, Aurangabad

Dr. N. Mohana Sundaram,

Professor, Department of Computer Science and Engineering of Faculty of Engineering of Karpagam Academy of Higher Education, India, Coimbatore

Dr. R. Santhosh,

Professor, Department of Computer Science and Engineering of Faculty of Engineering of Karpagam Academy of Higher Education, India, Coimbatore

Dr. Sharvari Chandrashekhar Tamane,

Professor and Head, Information Technology Department, MGM's Jawaharlal Nehru Engineering College, MGM University, India, Aurangabad

Januar Adi Putra,

SKom., MKom., Department of Computer Science, University of Jember, Indonesia

K. Tharageswari,

Research Scholar, Department of Computer Science and Engineering of Faculty of Engineering of Karpagam Academy of Higher Education, India, Coimbatore

Dr. Saiful Bukhori,

ST., MKom., Professor, Department of Computer Science, University of Jember, Indonesia

Verdy Bangkit Yudho Negoro,

Skom., Department of Computer Science, University of Jember, Indonesia

Windi Eka Yulia Retnani,

SKom., MT, Department of Computer Science, University of Jember, Indonesia