МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРКАСЬКИЙ ДЕРЖАВНИЙ ТЕХНОЛОГІЧНИЙ УНІВЕРСИТЕТ

ВІСНИК
ЧЕРКАСЬКОГО ДЕРЖАВНОГО ТЕХНОЛОГІЧНОГО УНІВЕРСИТЕТУ

Науковий збірник

Том 30,
№ 2. 2025

ЧЕРКАСИ
2025

ISSN: 2306-4412
E-ISSN: 2708-6070

Засновник:
Черкаський державний технологічний університет

Рік заснування: 1996
Рекомендовано до друку та поширення через мережу Інтернет Вченою радою Черкаського державного технологічного університету (протокол № 12 від 16 червня 2025)

Державна реєстрація: Ідентифікатор медіа R30-04613.
Рішення Національної ради України з питань телебачення і радіомовлення № 1916, протокол № 17 (30.05.2024).

Науковий збірник входить до переліку наукових фахових видань України
Категорія «Б». Спеціальність:
технічні фізико-математичні (Наказ Міністерства освіти і науки України № 886 від 02.07.2020): 113 Прикладна математика, 121 Інженерія програмного забезпечення, 122 Комп'ютерні науки, 123 Комп'ютерна інженерія, 125 Кібербезпека та захист інформації, 126 Інформаційні системи та технології, 131 Прикладна механіка, 132 Матеріалознавство, 133 Галузеве машинобудування, 151 Автоматизація та комп'ютерно-інтегровані технології, 152 Метрологія та інформаційно-вимірювальна техніка, 172 Телекомунікації та радіотехніка; технічні (Наказ Міністерства освіти і науки України № 1188 від 24.09.2020): 101 Екологія, 161 Хімічні технології та інженерія

Науковий збірник представлено у міжнародних наукометричних базах даних, репозитаріях та пошукових системах: Bielefeld Academic Search Engine (BASE), Crossref, CiteFactor, Litmaps, Ulrich's Periodicals Directory, WorldCat, J-Gate, Open Ukrainian Citation Index (OUCI), Наукова періодика України, Національна бібліотека України імені В. І. Вернадського (НБУВ), Dimensions, UCSB Library, Google Академія, German Union Catalogue of Serials (ZDB), Бібліотека Університету Осло, Бібліотека Університету Халла, SOLO - Search Oxford Libraries Online, Інститут Європейського університету, Бібліотека Лейпцизького університету (UBL), Бібліотека Кембриджського університету

Адреса редакції:
Черкаський державний технологічний університет
18006, бульв. Шевченка, 460, м. Черкаси, Україна
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/uk

MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE
CHERKASY STATE TECHNOLOGICAL UNIVERSITY

BULLETIN
OF CHERKASY STATE TECHNOLOGICAL UNIVERSITY

Scientific Collection

Volume 30,
No. 2. 2025

CHERKASY
2025

ISSN: 2306-4412
E-ISSN: 2708-6070

Founder:
Cherkasy State Technological University

Year of foundation: 1996
Recommended for printing and distribution
via the Internet by the Academic Council
of Cherkasy State Technological University
(Minutes No. 12 of June 16, 2025)

State Registration: Media identifier R30-04613.
Decision of the National Council of Television and Radio Broadcasting of Ukraine
No. 1916, Minutes No. 17, dated 30.05.2024.

The scientific collection is included in the list of Professional Scientific Publications of Ukraine
Category "B". Specialties: technical, physical and mathematical (Order of the Ministry of Education and Science of Ukraine No. 886 of 02.07.2020): 0541 Mathematics,
0612 Database and network design and administration,
0613 Software and applications development and analysis, 0714 Electronics and automation,
0715 Mechanics and metal trades, 0716 Motor vehicles, ships and aircraft,
0788 Inter-disciplinary programmes and qualifications involving engineering, manufacturing and construction; technical (Order of the Ministry of Education and Science of Ukraine No. 1188 of 24.09.2020): 0521 Environmental sciences, 0522 Natural environments and wildlife,
0711 Chemical engineering and processes

The scientific collection is presented in the international scientometric databases, repositories and scientific systems: Bielefeld Academic Search Engine (BASE), Crossref, CiteFactor, Litmaps,
Ulrich's Periodicals Directory, WorldCat, J-Gate,
Open Ukrainian Citation Index (OUCI),
Scientific Periodicals of Ukraine, Vernadsky National Library of Ukraine (VNLU), Dimensions,
UCSB Library, Google Scholar, German Union Catalogue of Serials (ZDB), University of Oslo Library,
University of Hull Library, SOLO - Search Oxford Libraries Online, European University Institute,
Leipzig University Library (UBL), Cambridge University Library

Editors office address:
Cherkasy State Technological University
18006, 460 Shevchenko Blvd., Cherkasy, Ukraine
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/en

Редакційна колегія

Editorial Board

Зміст / Contents

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 30 No. 2. 2025

UDC 004.31: 004.056
DOI: 10.62660/bcstu/2.2025.10

Hardware implementation of the HDG hash function

Volodymyr Luzhetskyi*
Doctor of Technical Sciences, Professor
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0000-0001-7466-7738

Vitalii Seleznov
Assistant
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0009-0004-0225-9697

Abstract. Given the increasing role of the Internet of Things and related low-resource devices, research into hash functions that provide a high level of cryptographic strength with minimal hardware costs is relevant. The aim of the study was to propose a hardware implementation of a new HDG hash function designed for use in small devices in the form of a specialised processor in order to reduce the hardware costs of implementation. The research methods included structural design of each functional block, digital modelling in the Logisim-evolution environment, and synthesis on an ASIC platform using 0.18 μm technology with the standard UMCL18G212T3 library, as well as calculation of hardware complexity in Gate Equivalents. HDG meets the requirements of low-resource cryptography thanks to its byte-oriented architecture, which allows data processing at the level of individual bytes, ensuring high efficiency with limited memory resources and computing capabilities of devices. The structure of a specialised processor for hashing was presented. The HDG specialised processor is decomposed into four functional blocks, each of which implements a corresponding function: a register block for storing intermediate hash values; a shift register with linear feedback, which provides the generation of a pseudorandom sequence; a block for addition modulo 256 and a control block. The simulation results confirmed the correctness of the structure of the specialised processor and the interaction of its components. The calculated complexity of the HDG processor hardware implementation is 1,683 GE for calculating a 256-bit hash value, which meets the requirements of the international standard ISO/IEC FDIS 29192 for low-resource cryptography. A comparison with hardware implementations of well-known low-resource hash functions PHOTON, SPONGENT, S-Quark, GLUON, and HVH showed a reduction in hardware costs of 15% or more. In some cases, the HDG processor demonstrated lower implementation complexity for 256-bit hash values values of 256 bits compared to hash functions that provide calculations of hash values of 224 or 160 bits, which indicates the effectiveness of the developed structure and the feasibility of using such a specialised processor for devices with limited hardware resources

Keywords: hardware device; specialised processor; low-resource cryptography; hashing method; hardware complexity

Article's History: Received: 08.01.2025; Revised: 19.05.2025; Accepted: 16.06.2025.

Suggested Citation:
Luzhetskyi, V., & Seleznov, V. (2025). Hardware implementation of the HDG hash function. *Bulletin of Cherkasy State Technological University*, 30(2), 10-21. doi: 10.62660/bcstu/2.2025.10.

*Corresponding author

INTRODUCTION

The rapid development of the Internet of Things (IoT) and the use of various mobile devices imposes a set of requirements on cryptographic information protection measures, relating to the level of cryptographic resistance, hardware implementation costs and energy consumption. To meet these requirements, a separate branch of cryptography is being developed, known as low-resource cryptography. Among cryptographic transformations, hash functions play a key role, as they ensure the integrity of information. Therefore, research aimed at implementing hardware data hashing tools is particularly relevant. The use of modern cryptographic hash functions faces problems of limited memory and low computing power. One approach to solving these problems is to create specialised hardware that implements the data hashing function. The ISO/IEC 29192-1:2012 (2012) standard sets out requirements for low-resource cryptography devices. According to the provisions of this standard, typical requirements include ensuring a sufficient level of cryptographic strength under conditions of limited chip area, low device power consumption, and limited time for cryptographic transformation. To compare the hardware complexity of implementations, the standard provides for the use of a metric based on the number of equivalent logic elements (Gate Equivalents, GE). This indicator is a unified assessment of the microchip area, which allows for a comparative analysis of implementations, regardless of the manufacturing technology used. Hardware complexity, expressed in GE, is a key parameter in the design of cryptographic primitives for embedded systems and IoT devices.

In recent studies on the hardware implementation of low-resource hash functions, there has been active development of both existing algorithms and new designs focused on use in resource-constrained environments. The works of D.N. Gupta & R. Kumar (2021) and S. Windarta *et al*. (2022) presented a comparative analysis of existing low-resource hash functions such as HASH-ONE, QUARK, PHOTON, SPONGENT, and GLUON, which are based on the "sponge" design. The comparison was made based on six parameters: security level, hardware complexity, hash value size, throughput, power consumption, and the number of cycles required for execution. SPONGENT-256 provides a hardware complexity of 1,950 GE and the highest collision resistance. The authors concluded that there is no universal solution among the existing algorithms and emphasised the need to develop a new hash function that would provide an adequate level of security at low hardware costs. In particular, the article by D.N. Gupta & R. Kumar (2023) presented a flexible hardware implementation of the DeeR-Hash function for IoT, based on a sponge structure with bitwise transformations. DeeR-Hash supports output lengths of 88 and 160 bits and has low hardware complexity – 984 GE for the 160-bit version. The authors noted its suitability for resource-constrained devices due to its high performance and small chip area.

The article by S. Windarta *et al*. (2023) proposed two new lightweight hash functions, ALIT-Hash and TJUILIK-Hash, based on the Saturnin block cipher with a modified S-block and an operation mode similar to the Beetle sponge structure. The authors investigated the hash functions for resistance to differential and linear cryptanalysis and performed statistical testing of the developed methods using NIST STS. Modelling performed using Contiki-NG and the Cooja simulator shows that the two developed hash functions demonstrate better performance compared to PHOTON and SPONGENT, measured by five indicators. A hardware implementation of TJUILIK-Hash on the ATmega2560 microcontroller was proposed. The authors B. Widhiara *et al*. (2023) also proposed a lightweight hash function RM70 as an alternative to SPONGENT-88, which uses an ARX-SPN type permutation and computes an 88-bit hash. The researchers claimed that RM70 demonstrates better processing time and lower memory consumption compared to SPONGENT-88. The hash function was implemented on the Arduino Uno R3 platform, without providing an assessment of hardware complexity or comparison with FPGA/ASIC implementations.

The work by Y. Huang *et al*. (2021) presented a family of lightweight HVH hash functions based on a "sponge" construction using the VH block cipher. The hash functions support five hash length options ranging from 88 to 256 bits for different security levels and resource constraints. The main focus of the study was to improve the hashing speed, which for an 8-bit microcontroller in RFID reaches 1.47 Mbit/s, which is 10 times higher than SPONGENT-88 hashing. The study also presented a hardware implementation for the TSMC 180 nm process, where the total hardware complexity of HVH with a hash length of 256 bits is 2,009 GE.

The study by P.R. Lawhale *et al*. (2025) presented a lightweight hash function for IoT and big data, based on bit permutations, linear transformations, and S-boxes. The hardware implementation on FPGA showed high results in terms of power consumption and throughput with a hardware complexity of 2,378 logic resources (LR). The authors emphasised the compliance of the proposed solution with security standards and its suitability for a wide range of IoT applications. The article by S. Khan *et al*. (2023) presented optimised hardware implementations of the four finalists of the third round of the NIST LWC competition: PHOTON-Beetle, Ascon, Xoodyak and SPARKLE. The results showed that the PHOTON-Beetle implementation reduces the complexity of the hardware implementation to 50% of the original value. For Ascon and XOODYAK, an increase in throughput was achieved. The SPARKLE hash function was implemented for the first time using FPGA technology.

In the work of V.I. Seleznov & V.A. Luzhetskyi (2023), the HDG (Hash Data Generator) hashing method was

proposed, which takes into account the requirement for low hardware costs for implementation while maintaining the necessary level of security. The peculiarity of this hashing method is that intermediate and final hash values were calculated by adding data bytes at positions determined by the symbol 1 in the code generated by a pseudorandom number generator, which is why the authors called this approach to hashing "data generator". This byte-oriented approach to data hashing does not require padding of the last data block when its length is less than the specified block length. However, the paper provided only a theoretical estimate of the complexity of hardware implementation based on the structure of the device without detailing the specifics of the construction of individual blocks.

Among the hash functions considered, there are hash functions that have a hash value length of less than 128 bits and a hardware complexity of less than 1,000 GE, which are only allowed for use in devices with low security requirements. However, most hash functions calculate hash values with a length of 128-256 bits. The complexity of the hardware implementation of such hash functions ranges from 2,000 to 3,000 GE. Therefore, there is a need for hash functions with a hardware complexity of less than 2,000 GE for hash values with a length of 256 bits. The aim of the study was to reduce the hardware costs of implementing hash functions by developing a specialised processor structure that implements the HDG hash function.

MATERIALS AND METHODS

The HDG hashing function proposed in the work of V.I. Seleznov & V.A. Luzhetskyi (2023) implemented a byte-oriented model based on a modified Merkle-Demgard iterative structure. The peculiarity of the HDG function lies in the use of the "data-generator" hashing approach, according to which the hash value is calculated based on the input data, using the simplest arithmetic operations, a control pseudorandom sequence generated by a specialised generator. The message M to be hashed is presented as a sequence of L bytes:

$$M = \{m_1, m_2, \ldots, m_L\}. \quad (1)$$

The initial hash value h_0 is represented as a sequence of $k = l/8$ bytes, where l is the length of the hash value in bits:

$$h_0 = \{h_0, 0, h_{0,1}, \ldots, h_{0,k-1}\}. \quad (2)$$

The process of calculating intermediate hash values is provided by the compression function f. Each intermediate hash value $h_i = \{h_{i,0}, h_{i,1}, \ldots, h_{i,k-1}\}$ is calculated based on the previous h_{i-1} and the current value of the message byte m_i. For each value , the compression function implements an iterative process consisting of the following steps: generating pseudo-random bits $g_{i,j}$ for each j from 0 to; k-1 calculating the intermediate hash value h_i using the formula:

$$h_{i,j} = (h_{i-1,j} + m_i \cdot g_{i,(k-1-j)}) mod\ 256. \quad (3)$$

The final hash value of the message $h = \{h_{L,0}, h_{L,1}, \ldots, h_{L,k-1}\}$ is calculated as a result of executing the compression function f for the last byte of data M. To meet the requirements of low-resource cryptography, the described hash function is implemented in the form of a specialised processor (HDG processor). This processor is integrated into the computer system by connecting to the central processor. The system's central processor stores and forms the sequence of bytes of the input message for hashing. The HDG processor performs all arithmetic operations necessary for calculating hash values and stores intermediate hash values. The design of the specialised processor is based on the mathematical apparatus of digital automata theory. It consists of two key components: an operational automaton responsible for performing calculations and a control automaton that coordinates the sequence of operations performed (Harris & Harris, 2021). The complexity of implementing the hash function requires the decomposition of the operational automaton into blocks, each of which performs a specific set of functions, which avoids excessive complexity of the structure and increases the efficiency of data processing.

To verify the correct functioning of the specialised processor, the Logisim-evolution (n.d.) simulation environment was used, which allows the logical structure of the device to be created and verified at the level of the simplest elements and functional blocks, providing the ability to analyse in detail the operation of each block and their interaction. The use of modelling tools made it possible to evaluate the correctness of the device's operation and determine its hardware complexity even before the physical implementation stage. The hardware implementation of the developed HDG processor is planned on an ASIC platform using 0.18 μm technology and the UMCL18G212T3 library (Virtual Silicon Inc, 2004). The hardware complexity of the HDG processor was calculated in GE units. The research consisted of three main stages: development of structural components of the specialised HDG processor; modelling of individual components and the specialised processor as a whole in the Logisim-evolution environment; calculation of the hardware complexity of the HDG processor in conventional units (GE).

RESULTS AND DISCUSSION

Based on the characteristics of the HDG hash function, the structure of the specialised processor is represented as a set of four components: a block of registers (BRG) for storing intermediate hash values; a pseudorandom sequence generator implemented as a linear feedback shift register (LFSR); a modulo 256 addition block (ADDB); a control block (CB) that communicates with the central processor, implements the initialisation process and controls the data hashing device. The structural diagram of the HDG processor is shown in Figure 1.

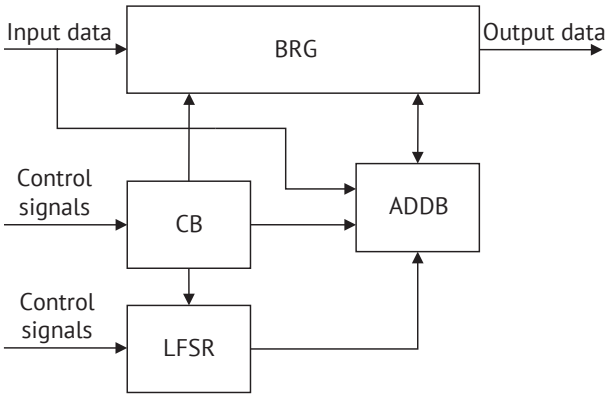

Figure 1. Block diagram of a specialised processor
Source: developed by the authors

The interaction between the HDG processor and the central processor takes place via a data exchange bus. The input bus provides byte-by-byte data input for hashing, while the output bus is used to transfer the hashing result of the processed data. The input signals of the HDG processor are clock signals and other control signals. The BRG block consists of 32 serially connected eight-bit registers that store intermediate hash values of 256 bits in length and are connected to data buses. By using a control input, the block provides the ability to set the initial state of the registers, which is the initial hash value. In the case of a keyless hash function, this initial state is known. If a keyed hash function is implemented, the initial state is a secret key.

The LFSR block is responsible for generating pseudo-random values. The LFSR is based on a 32-bit shift register with linear feedback of maximum period. The output signal of the LFSR is a generated pseudo-random bit *g*, which is then fed to the ADDB block. The LFSR also uses a separate control signal to initialise the initial state of the register. The ADDB block performs the addition operation modulo 256 of two 8-bit numbers. ADDB consists of eight 1-bit adders and eight AND logic elements, which together implement the operation described by formula (1). Unlike other blocks, ADDB does not require clock signals to operate. The CB block provides initialisation and synchronisation of other blocks. It consists of a five-bit counter and an RS flip-flop. Based on the control signal, the CB block sets the counter to zero and, during the next 32 cycles, controls the recording of the initial states of the BRG block registers and the LFSR register. After initialisation is complete, the device enters the hash value calculation mode. The HDG processor and its components are presented in the form of models:

$$HDG = \{D, R, S, Y, X\}, \quad (4)$$

where D is the set of input signals received by the device; R is the set of output signals resulting from hashing; S is the set of internal states used to store intermediate results during the execution of operations; Y is the set of operations performed by the device; X is the set of logical conditions. The set of input signals D includes: data bytes m_i to be hashed; bits S_0 that make up the initial state of the LFSR register. The set of output signals R includes the hashing result H. Among the internal states S of the device are: S_{BRG} – the state values of the BRG block registers; S_{LFSR} – the state of the LFSR; S_{CB} – the state of the CB block components. The set Y includes the following operations: synchronisation (clock signal clk0); initialisation of the initial state of the HDG processor and its components (signal rst_0); cumulative addition *ADD*. The model of the hardware device that implements the HDG hash function is shown in Figure 2.

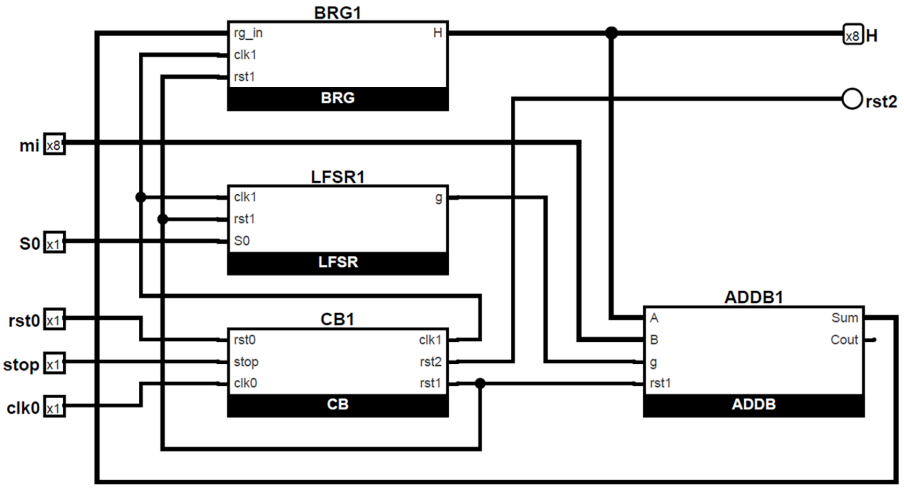

Figure 2. HDG processor model
Source: developed by the authors

Thus, the HDG processor model is described by the following set of sets.

1. Set of input signals $D = \{ m_i, S_0\}$.
2. Set of output signals $R = \{H\}$.

3. Set of internal states $S = \{ S_{BRG}, S_{LFSR}, S_{CB}\}$.
4. Set of operations $Y = \{clk_0, rst_0, ADD\}$.
5. Set of logical conditions $X = \{ x_0, x_1\}$, where: x_0 – condition that $rst_2 = 1$; x_1 – condition that $stop = 1$.

The LFSR block is a 32-bit shift register with linear feedback, the diagram of which is shown in Figure 3, and the algorithm of its operation is shown in Figure 4.

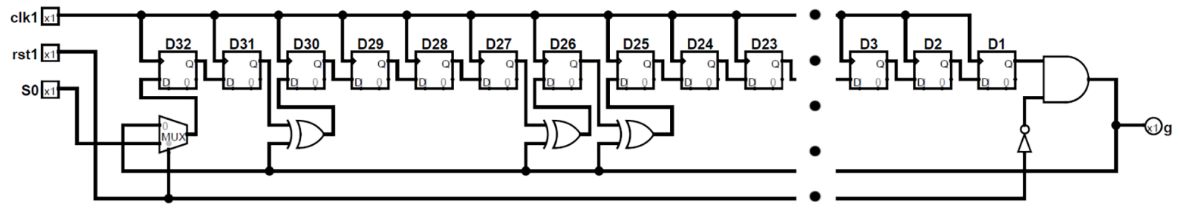

Figure 3. Fragment of the LFSR block model

Source: developed by the authors

Thus, the LFSR block model is described by the following set of sets.

1. Set of input signals $D = \{S_0\}$.
2. Set of output signals $R = \{g\}$.
3. Set of internal states $S = \{D_1 \dots D_{32}\}$.
4. Set of operations $Y = \{clk_1, rst_1, y_1, y_2\}$.

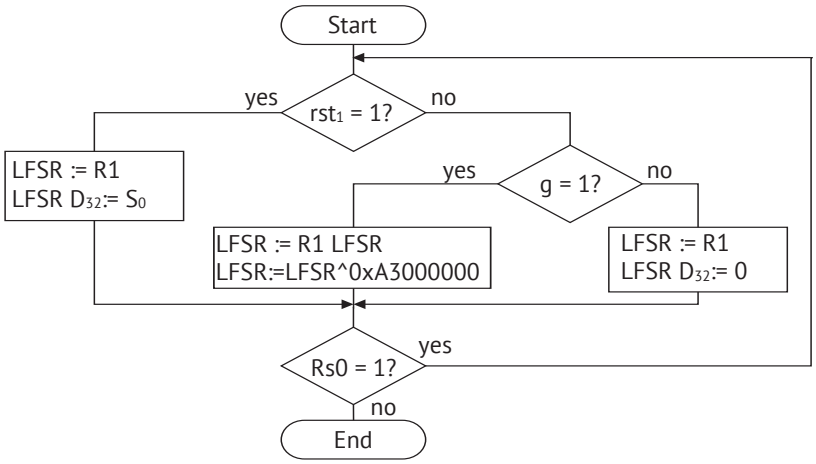

Figure 4. LFSR block operation algorithm

Source: developed by the authors

It consists of 32 D-triggers D_{32} – D_1, where each trigger is responsible for storing one bit of state. A 2-in-1 multiplexer (MUX) is connected to the input of trigger D_{32}, and an AND logic element is connected to the output of D_1, to which the inverted value of the rst_1 signal is supplied through another input. When the rst_1 signal is 1, the multiplexer transmits S_0 signals to the input of trigger D_{32} to initialise the initial state of the LFSR, and the AND element breaks the feedback from the output of trigger D_1 by forming $g = 0$. If the rst_1 signal is 0, the multiplexer restores the connection from the output of the AND element to trigger D_{32}, and the AND element restores the feedback to the XOR elements connected to the inputs of the D-triggers, the numbers of which are determined by a primitive polynomial:

$$P(x) = x^{32} + x^{30} + x^{26} + x^{25} + 1. \quad (5)$$

The presented configuration provides minimal hardware costs for implementing a shift register with linear feedback and a maximum period of the pseudorandom sequence $2^{32}–1$ (Ward & Molteno, 2012). The input signals D of the LFSR block include: bit S_0 (32 bits S_0 constitute the initial state of the LFSR register). The initial state of the LFSR is the secret key code in the case of a key hash function or a cryptographic salt, which provides an additional level of cryptographic resistance in the case of a keyless hash function. The output signal of the block is a pseudorandom bit g. Among the internal states S of the LFSR block are the values of the D-flip-flops ($D_1 \dots D_{32}$). The set Y includes the following operations: synchronisation (clock signal clk_1); initialisation of the initial state of the LFSR HDG processor and its components (signal rst_1); writing to a register with feedback (operation y_1) and writing without feedback (operation y_2).

The BRG block consists of 32 eight-bit registers Rg_1 – Rg_{32} connected in series, where each register is responsible for storing eight bits of state. An eight-bit AND logic element is connected to the output of register Rg_{32}, to the other input of which the inverted value of the rst_1 signal is supplied, as shown in Figure 5.

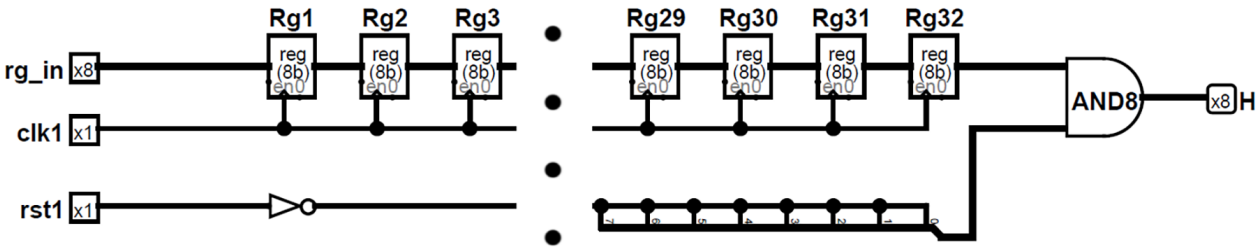

Figure 5. Fragment of the BRG block model

Source: developed by the authors

When $rst_1 = 1$, the AND element provides the BRG block output signal $H = 0$, which is necessary to set the initial state of the BRG block. Setting the initial state of BRG consists of sequentially writing the bytes of the initial hash value H_0 to the block registers. The H_0 bytes are sent to the HDG processor as an input signal m_i through the ADDB block, where they are added with zero bytes (output signal $H = 0$ of the BRG block) and written to the BRG block (signal *rg_in*). If $rst_1 = 0$, then the AND element ensures the transmission of the signal $H = Rg_{32}$. All these actions are reflected in the BRG block operation algorithm shown in Figure 6.

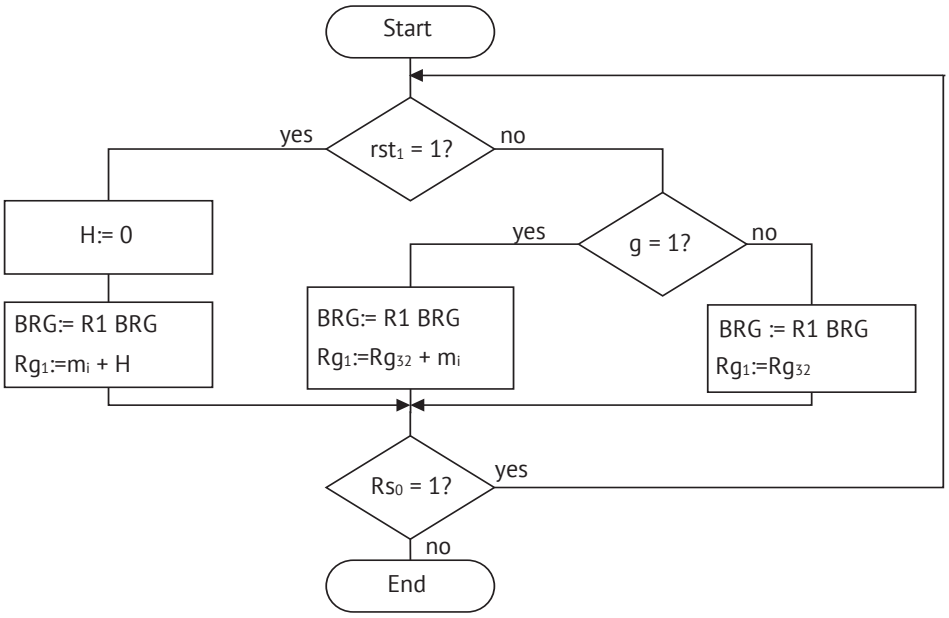

Figure 6. BRG block operation algorithm

Source: developed by the authors

The set of input signals *D* of the BRG block includes the *rg_in* byte. The output signal of the block is the *H* byte (the first 32 bytes H_i form an intermediate hash value, and the last 32 bytes *H* are the result of hashing). Among the internal states S of the BRG block are the values of the eight-bit registers ($Rg_1 \dots Rg_{32}$). The set *Y* includes the following operations: synchronisation (clock signal clk_1); initialisation of the BRG initial state (signal rst_1).

Thus, the model of the BRG block included in the HDG processor is described by the following set of sets.

1. Set of input signals $D = \{rg\_in\}$.
2. Set of output signals $R = \{H\}$.
3. Set of internal states $S = \{Rg_1 \dots Rg_{32}\}$.
4. Set of operations $Y = \{clk_1, rst_1\}$.

The ADDB block performs the addition of values *A* and *B*. ADDB is an 8-bit combinational adder with additional control of operand *B*, as shown in Figure 7. Operands A and B are 8-bit binary codes. Operand *B* is the current byte of the input message mi, and operand *A* is the output byte from the output of register Rg_{32} of the BRG block. The ADDB block has two control signals (signals rst_1 and *g*) that are fed to the inputs of the OR element. The value of the *g* signal is a pseudo-random bit coming from the LFSR block. The value $rst_1 = 1$ corresponds to the initialisation process of the BRG and LFSR blocks, during which $g = 0$ (provided by the control in the LFSR block), the zero code is supplied as operand *A* (provided by the control in the BRG block), therefore $Sum = B$. When $rst_1 = 0$, operand *B* is controlled by signal *g*. If $g = 1$, then operand *B* is sent to the ADDB input. Otherwise, a zero code will be sent instead of operand *B*.

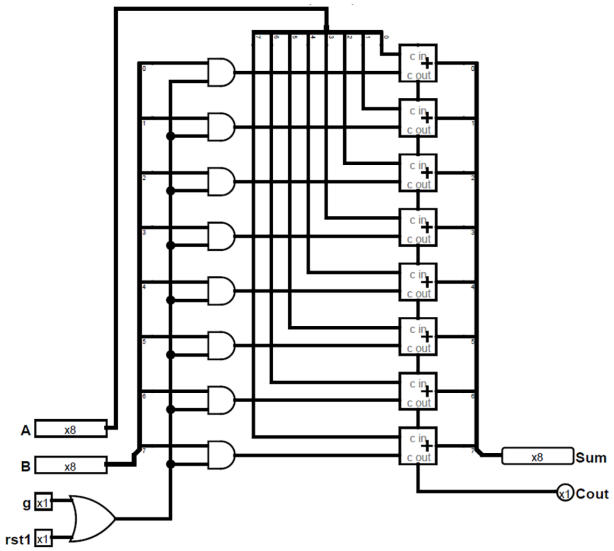

Figure 7. ADDB block model

Source: developed by the authors

The CB block controls the data hashing process and interacts with the central processor. Among the input signals *D* of the CB block, the clock signal clk_0, signals rst_0 and *stop* are described above for the HDG processor. The set of output signals of the block includes the clk_1 clock signal, the rst_2 signal, whose change indicates the completion of hashing of the *i*-th byte of data by the HDG processor, and rst_1, which controls other blocks. The internal state of the $S_{Counter1}$ and the states of the S_{Rs0}, S_{Rs1} flip-flops are included in the set *S* of block CB. The main operations are setting the triggers to the unit state (SET_{Rs0} and SET_{Rs1}), setting the triggers and counter to the zero state ($RESET_{Rs0}$, $RESET_{Rs1}$ та $RESET_{Counter1}$, respectively), and the *INC* counting operation.

The model of the CB block shown in Figure 8 is described by a set of sets.

1. Set of input signals D = {clk_0, rst_0, *stop*}.
2. Set of output signals R = {clk_1, rst_1, rst_2}.
3. Set of internal states S = { $S_{Counter1}$, S_{Rs0}, S_{Rs1}}.
4. Set of operations *Y* = {SET_{Rs0}, SET_{Rs1}, $RESET_{Rs0}$, $RESET_{Rs1}$, $RESET_{Counter1}$, *INC*}.

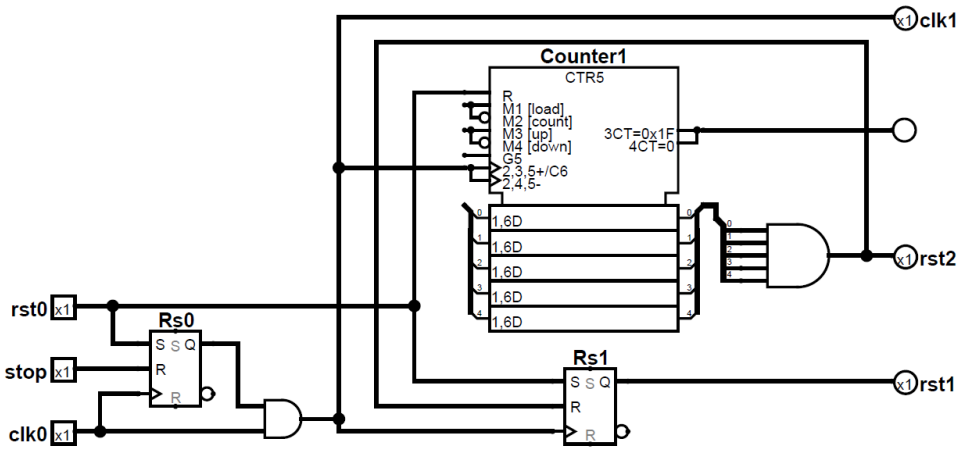

Figure 8. Model to CB block

Source: developed by the authors

The CB block operation algorithm is shown in Figure 9. The block initialisation begins with the rst_0 signal, which sets the Rs_0 trigger to the unit state, thanks to which the clk_1 clock signals begin to be received by the block. In addition, the Rs_1 flip-flop is set to the one state, and the *Counter1* counter is set to the initial zero state and starts counting. Thanks to the Rs_1 flip-flop, the rst_1 signal acquires a constant one value for 32 cycles, which allows the BRG and LFSR blocks to be initialised (Fig. 2). When the counter state reaches code 11111, a $rst_2 = 1$ signal is generated, which sets Rs_1 to zero, which in turn completes the initialisation mode. After initialisation is complete, the device enters the hash value

calculation mode. When the hashing of the entire message is complete, the central processor sends 32 bytes of zero values to the HDG processor to read the hashing result, and then sends a *stop* = 1 signal to complete the operation, which sets the Rs_0 flip-flop to zero, stopping the formation of clk_1 clock signals.

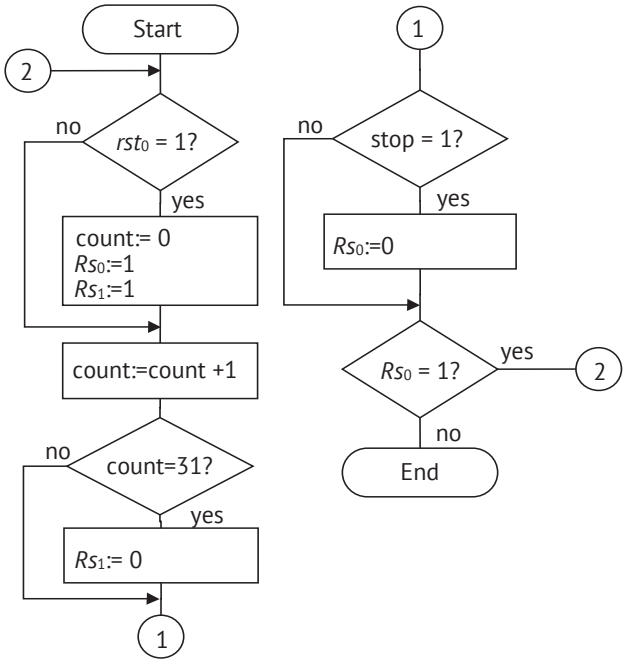

Figure 9. Algorithm of the CB block operation
Source: developed by the authors

The results of modelling using the Logisim-evolution environment confirmed the correctness of the logical structure of the specialised processor and the interaction of its components in accordance with the given algorithms. The developed models were used to calculate the hardware complexity of each component and the HDG processor as a whole. The total hardware complexity of the HDG processor was calculated using the formula:

$$S_{HDG} = S_{BRG} + S_{LFSR} + S_{ADDB} + S_{CB}, \quad (6)$$

where S_{BRG} – complexity of the BRG block; S_{LFSR} – complexity of the LFSR pseudorandom number generator; complexity of the ADDB addition block; S_{ADDB} – complexity of the CB control block. The BRG consists of 32 registers, each of which consists of eight *D Flip flops*, eight NAND elements and one NOT element, so the total complexity of the block is;

$$S_{BRG} = 32 \cdot 8 \cdot D + 8NAND + NOT. \quad (7)$$

The LFSR consists of 4 XOR elements, 32 *D*-triggers, a 2-in-1 MUX multiplexer, one NOT and one AND element, therefore:

$$S_{LFSR} = 4XOR + 32 \cdot D + MUX + NOT + AND. \quad (8)$$

The ADDB block consists of 8 AND elements, one OR element, and 8 single-bit adders, each of which, in turn, consists of two AND elements, one OR element, and two XOR elements:

$$S_{ADDB} = 8AND + OR + 8\,(2AND + OR + 2XOR). \quad (9)$$

The CB control module consists of two RS triggers, one AND element and a 5-bit counter. The RS trigger consists of 2 NAND elements. The controlled 5-bit counter consists of 4 AND elements and 5 *D* triggers. Complexity of the CB module:

$$S_{CB} = 4NAND + 5D + 5AND. \quad (10)$$

Based on formulas (7-10), the hardware complexity is as follows:

$$S_{HDG} = 293D + 2NOT + 29AND + 10NAND + 9OR + \\ + 20XOR + MUX. \quad (11)$$

The complexity of implementing an HDG processor that calculates 256-bit hash values on a microchip using 0.18 μm technology and the UMCL18G212T3 library in GE conventional units: $D = 5.33$; NOT = 0.67; NAND = 1; AND = 1.33; OR = 1.33; MUX = 2.33; XOR = 2.67. Based on this, the resulting complexity for the HDG-processor implementation is:

$$S_{HDG} = 1{,}683\text{GE}. \quad (12)$$

To compare the obtained estimate of the complexity of the hardware implementation of the HDG hash function, Table 1 shows estimates for known low-resource hash functions. Although the ISO/IEC 29192 standard allows the use of hash functions that generate hash values ranging from 80 to 256 bits for low-resource applications, there has been a recent trend towards higher bit-length hash values, namely up to 256 bits. Therefore, Table 1 lists the PHOTON, S-Quark, SPONGENT, and HVH hash functions, which generate output hash values with a length of 224 and 256 bits. At the same time, for the sake of completeness, functions with shorter hash values (160-224 bits), such as DeeR-Hash, Hash-One and GLUON, were also included.

Table 1. Hardware complexity of hashing devices

Hash function	Structure	Hash value length, bits	Hardware complexity, GE
DeeR-Hash	Sponge-based	160	984
Hash-One	Sponge-based	160	1,006
D-Quark	Sponge-based + NFSR	176	1,702
S-Quark		256	2,296

Continued Table 1.

Hash function	Structure	Hash value length, bits	Hardware complexity, GE
PHOTON	Sponge-based	224	1,736
		256	2,177
SPONGENT	Sponge-based	224	1,728
		256	1,950
GLUON	Sponge-based	160	2,800
		224	4,724
HVH	Sponge-based + block cipher VH	224	1,769
		256	2,009

Source: developed by the authors based on Y. Huang *et al.* (2021), S. Windarta *et al.* (2022), D.N. Gupta & R. Kumar (2023)

The PHOTON hash function proposed in the study by J. Guo *et al.* (2011) uses a sponge-type design with an AES-like internal permutation structure. PHOTON requires 2,177 GEs to compute 256-bit hash values. Each round of permutation uses operations based on S-blocks, bitwise addition of constants, cyclic row shifts, and linear column transformations. Compared to HDG, PHOTON has a more complex structure because it uses a classic cryptographic symmetric design with substitution tables. J.P. Aumasson *et al.* (2010) developed the D-Quark and S-Quark hash functions, which are also based on the "sponge" design, but unlike PHOTON, they use shift registers as their core. S-Quark uses three shift registers: two non-linear NFSRs and one linear LFSR, which provides bitwise processing without the use of substitution tables or matrices. Compared to HDG, which also uses register generation, S-Quark is limited to logical operations only. Unlike PHOTON, where the construction is based on S-blocks and linear transformations, transformations in S-Quark are implemented using a combination of gates. However, as the length of the hash value increases, the logic of registers and combined functions becomes more complex, leading to an increase in logical depth and, accordingly, to a more complex and slower hardware implementation.

As noted in the works of C. Manifavas *et al.* (2014) and I. El Gaabouri *et al.* (2022), for implementation on devices with very low resources (e.g., 4-bit microcontrollers), hardware complexity of up to 1,000 GE is considered acceptable, and for more general IoT platforms, implementations of up to 2,000-3,000 GE are acceptable. A separate section of the ISO/IEC 29192-5:2016 (2016) standard defines requirements for hash functions intended for use in environments with limited computing and energy resources. The key criterion here is to ensure cryptographic resistance to collisions and attacks on the first prototype with minimal hardware costs. The minimum acceptable hash length is considered to be 80 bits, while ensuring a basic level of collision resistance of 2^{40} operations, which is satisfactory for typical applications in the IoT environment. At the same time, given the growth in the computing capabilities of potential attackers and the increasing requirements for long-term security, modern scientific publications note a trend towards a gradual transition to hash functions with an output length of 256 bits, even in low-resource implementations (Bogdanov *et al.*, 2013; Buchanan *et al.*, 2017; Seleznov, 2023). This approach allows for collision resistance at the level of 2^{128} attempts, which is critical for secure systems with long lifecycles or high requirements for data confidentiality and integrity.

H. Martin *et al.* (2017) proposed a low-resource implementation of the Tav-128 hash function for use in RFID tags and sensor nodes. Tav-128 is based on the classic Merkle-Demgard design proposed in I. Damgård (1989) and R.C. Merkle (1990). The study presented three variants of the ASIC hardware implementation of this function, optimised to minimise area, maximise throughput, and achieve a balance between these parameters. The smallest area implementation of Tav-128 occupies 3,194 GE, which exceeds the authors' initial estimate (2,578 GE) due to the complexity of the control logic. A. Bogdanov *et al.* (2013) developed the SPONGENT hash function, which also implements a "sponge" structure with internal permutation based on the low-resource block cipher PRESENT. The main transformations are 4-bit S-boxes, bitwise permutations, and bitwise XOR operations, which makes SPONGENT similar to PHOTON. At the same time, unlike PHOTON, the authors use bitwise permutations as a linear diffusion transformation, which reduces the complexity of hardware implementation but worsens the diffusion properties. Research by D.N. Gupta & R. Kumar (2021) showed the possibility of implementing SPONGENT with a hardware complexity of 1,950 GE, which is one of the best indicators for 256-bit hashing.

The GLUON hash function calculates hash values based on a combination of Feedback with Carry Shift Registers (FCSR) and simple logic. As shown by research by D.N. Gupta & R. Kumar (2021), this structure allows for high entropy, but the authors only propose a hardware implementation, which provides faster hashing but comes with significant hardware costs: GLUON implementation requires 2,800 to 4,700 GE. Hash-One, developed by P.M. Mukundan *et al.* (2016), similar to the S-Quark hash function, implements a

"sponge" design with two non-linear feedback shift registers (NFSR). In addition, Boolean functions (P_j, Q_j, L_j) implemented exclusively using NAND gates are used as the main transformations. This allows for a hardware complexity of 1,006 GE to calculate 160-bit hash values. Unlike HDG, which uses a pseudorandom bit generator to control bytewise addition operations, Hash-One relies on fixed state update logic via NFSR. In turn, DeeR-Hash, presented by D.N. Gupta & R. Kumar (2023), also uses two NFSRs and additionally introduces one LFSR to increase diffusion. Both functions have a similar structure and implement state updates through Boolean functions, but DeeR-Hash uses a dynamic scheme for selecting variables from arrays, which allows for optimising the logic. A system of multiplexers and configuration arrays is used to select positions, forming updates with a minimal set of logic elements. Thanks to this, DeeR-Hash provides 160-bit hash value calculations with a hardware complexity of 984 GE. Proposed by Y. Huang *et al*. (2021), the HVH hash function implements a "sponge" design based on the lightweight VH block cipher with a block size of 64 bits. VH is constructed as a substitution-permutation network (SPN) that determines the nature of transformations during hashing. HVH supports five hash length options: 88, 128, 160, 224, and 256 bits, allowing it to be adapted to various resource-constrained scenarios. The authors estimate the complexity of the hardware implementation of HVH for 256-bit hash values at 2,009 GE, which is on par with the SPONGENT hash function.

Comparing the structural characteristics of all the hash functions considered, it can be noted that PHOTON, SPONGENT, and HVH use either AES-like transformations or block ciphers with S-blocks, which increases the complexity of the logic. On the other hand, GLUON, Quark, HASH-ONE, and DeeR-Hash, like HDG, are focused on using shift registers, bitwise operations, and simple logical functions. A common feature of all the hash functions studied is the use of a "sponge" construction or its modifications, except for HDG, which does not involve absorption and squeezing procedures, but instead implements the classic Merkle-Demgard iterative construction for forming hash values. Thus, the proposed HDG processor has a hardware complexity of 1,683 GE for a hash value length of 256 bits, which is the lowest complexity among all the low-resource hash functions considered. A comparative analysis of hardware complexities also showed that in some cases, the HDG processor is simpler than hardware devices that generate hash values of smaller bit lengths, in particular for the GLUON (160 bits), D-Quark (176 bits) and PHOTON (224 bits) hash functions.

CONCLUSIONS

The article presented the hardware implementation of the HDG hash function in the form of a specialised processor designed for use in low-resource devices thanks to its simple byte-oriented architecture, which ensures efficient data processing at the level of individual bytes. The development took into account the requirements for minimising hardware costs, simplicity of logic and ensuring an adequate level of cryptographic strength in accordance with the ISO/IEC 29192 standard. The proposed structure of the HDG processor for hashing was decomposed into four main modules: BRG, LFSR, ADDB, and a CB. Detailed design of each of the blocks allowed formalising the functional logic of the processor, adapted for digital implementation. The processor and its components were modelled in the Logisim-evolution environment, which allowed confirming the correct operation of the HDG processor and its compliance with the stated functional requirements.

During the modelling process, it was confirmed that the HDG processor correctly performs the algorithmically defined sequence of hashing operations and generates the expected hash value for the test input data, which indicates the functional consistency of the implemented hardware logic with the basic specification of the algorithm. The results of the hardware complexity analysis showed that the implementation of the HDG processor for calculating 256-bit hash values requires 1,683 GE, which meets the requirements for low-resource cryptography and indicates the possibility of using the HDG processor in environments with restrictions on hardware complexity.

A comparative assessment with hardware implementations of known hash functions, in particular PHOTON, SPONGENT, S-Quark, GLUON and HVH, showed that the proposed implementation provides a reduction in hardware costs of 15% or more, depending on the selected analogue function. In some cases, the HDG processor implementation even provides lower hardware complexity when generating a 256-bit hash than some known functions that work with shorter hash values, such as 224 or 160 bits, which demonstrates the effectiveness of the design underlying the HDG processor. This level of optimisation was achieved by using only elementary logical and arithmetic operations without involving complex nonlinear transformations. Thus, the developed HDG processor is suitable for use in devices with limited hardware resources and can be the basis for further research in the field of low-resource cryptography. It is promising to study the possibility of improving the architecture of the HDG processor for effective implementation using FPGA technology.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Aumasson, J.P., Henzen, L., Meier, W., & Naya-Plasencia, M. (2010). QUARK: A lightweight hash. In S. Mangard & F.-X. Standaert (Eds.), *Cryptographic hardware and embedded systems – CHES 2010* (pp. 1-15). Cham: Springer. doi: 10.1007/978-3-642-15031-9_1.

[2] Bogdanov, A., Knezevic, M., Leander, G., Toz, D., Varici, K., & Verbauwhede, I. (2013). SPONGENT: The design space of lightweight cryptographic hashing. *IEEE Transactions on Computers*, 62(10), 2041-2053. doi: 10.1109/TC.2012.196.

[3] Buchanan, W.J., Li, S., & Asif, R. (2017). Lightweight cryptography methods. *Journal of Cyber Security Technology*, 1(3-4), 187-201. doi: 10.1080/23742917.2017.1384917.

[4] Damgård, I. (1989). A design principle for hash functions. *Lecture Notes in Computer Science*, 435, 416-427. doi: 10.1007/0-387-34805-0_39.

[5] El Gaabouri, I., Senhadji, M., & Belkasmi, M. (2022). A survey on lightweight cryptography approach for IoT devices security. In *2022 5th international conference on networking, information systems and security: Envisage intelligent systems in 5g//6G-based interconnected digital worlds (NISS)* (pp. 1-8). Bandung: IEEE. doi: 10.1109/NISS55057.2022.10085144.

[6] Guo, J., Peyrin, T., & Poschmann, A. (2011). *The PHOTON family of lightweight hash functions*. Retrieved from https://eprint.iacr.org/2011/609.

[7] Gupta, D.N., & Kumar, R. (2021). Sponge based lightweight cryptographic hash functions for IoT applications. *In 2021 international conference on intelligent technologies (CONIT)* (pp. 1-5). Hubli: IEEE. doi: 10.1109/conit51480.2021.9498572.

[8] Gupta, D.N., & Kumar, R. (2023). DeeR-Hash: A lightweight hash construction for Industry 4.0 / IoT. *Journal of Scientific & Industrial Research*, 82(1), 142-150. doi: 10.56042/jsir.v82i1.69938.

[9] Harris, S.L., & Harris, D. (2021). Digital design and RISC-V computer architecture textbook. In *2021 ACM/IEEE workshop on computer architecture education (WCAE)* (pp. 1-5). Raleigh: IEEE. doi: 10.1109/wcae53984.2021.9707615.

[10] Huang, Y., Li, S., Sun, W., Dai, X., & Zhu, W. (2021). HVH: A lightweight hash function based on dual pseudo-random transformation. In *Security, privacy, and anonymity in computation, communication, and storage* (pp. 492-505). Cham: Springer. doi: 10.1007/978-3-030-68884-4_41.

[11] ISO/IEC 29192-1:2012. (2012). *Information technology – security techniques – lightweight cryptography – Part 1: General*. Retrieved from https://www.iso.org/standard/56425.html.

[12] ISO/IEC 29192-5:2016. (2016). *Information technology – security techniques – lightweight cryptography – Part 5: Hash-functions*. Retrieved from https://www.iso.org/standard/67173.html.

[13] Khan, S., Lee, W.-K., Karmakar, A., Mera, J.M.B., Majeed, A., & Hwang, S.O. (2023). Area-time efficient implementation of NIST lightweight hash functions targeting IoT applications. *IEEE Internet of Things Journal*, 10(9), 8083-8095. doi: 10.1109/jiot.2022.3229516.

[14] Lawhale, P.R., Kale, S.N., Kasturiwale, H., & Thakare, Y.N. (2025). FPGA implementation of compact architecture for lightweight hash algorithm for resource constrained devices. *Communications on Applied Nonlinear Analysis*, 32(2), 679-691. doi: 10.52783/cana.v32.1861.

[15] Logisim-evolution. (n.d.). *Digital logic design tool and simulator*. Retrieved from https://surl.li/pjzuiy.

[16] Manifavas, C., Hatzivasilis, G., Fysarakis, K., & Rantos, K. (2014). Lightweight cryptography for embedded systems – a comparative analysis. In *Data privacy management and autonomous spontaneous security* (pp. 333-349). Berlin: Springer. doi: 10.1007/978-3-642-54568-9_21.

[17] Martin, H., Lopez, P.P., San Millan, E., & Tapiador, J.E. (2017). A lightweight implementation of the Tav-128 hash function. *IEICE Electronics Express*, 14(11), article number 20161255. doi: 10.1587/elex.14.20161255.

[18] Merkle, R.C. (1990). One way hash functions and DES. *Lecture Notes in Computer Science*, 435, 428-446. doi: 10.1007/0-387-34805-0_40.

[19] Mukundan, P.M., Manayankath, S., Srinivasan, C., & Sethumadhavan, M. (2016). Hash-One: A lightweight cryptographic hash function. *IET Information Security*, 10(5), 242-252. doi: 10.1049/iet-ifs.2015.0385.

[20] Seleznov, V.I. (2023). Analysis of low-resource hashing methods. In *LII scientific and technical conference of VNTU departments: Proceedings of the scientific conference* (pp. 21-23). Vinnytsia: Vinnytsia National Technical University.

[21] Seleznov, V.I., & Luzhetskyi, V.A. (2023). Low-resource hashing method of the "data – generator" type. *Cybersecurity: Education, Science, Technique*, 28, 84-95.

[22] Virtual Silicon Inc. (2004). *0.18 µm VIP standard cell library tape out ready (part number: UMCL18G212T3). Process: UMC logic 0.18 µm generic II technology*. Retrieved from https://www.eetimes.com/virtual-silicon-ships0-18-micron-libraries/.

[23] Ward, R., & Molteno, T.C.A. (2012). *Table of linear feedback shift registers*. Otago: University of Otago.

[24] Widhiara, B., Kurniawan, Y., & Susanti, B.H. (2023). RM70: A lightweight hash function. *IAENG International Journal of Applied Mathematics*, 53(1), 94-102.

[25] Windarta, S., Suryadi, Ramli, K., Pranggono, B., & Gunawan, T.S. (2022). Lightweight cryptographic hash functions: Design trends, comparative study, and future directions. *IEEE Access*, 10, 82272-82294. doi: 10.1109/access.2022.3195572.

[26] Windarta, S., Suryadi, S., Ramli, K., Lestari, A.A., Wildan, W., Pranggono, B., & Wardhani, R.W. (2023). Two new lightweight cryptographic hash functions based on saturnin and beetle for the Internet of Things. *IEEE Access*, 11, 84074-84090. doi: 10.1109/access.2023.3301128.

Апаратна реалізація геш-функції HDG

Володимир Лужецький

Доктор технічних наук, професор
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0000-0001-7466-7738

Віталій Селезньов

Асистент
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0009-0004-0225-9697

Анотація. З огляду на зростання ролі Інтернету речей та пов'язаних із ним малоресурсних пристроїв, дослідження геш-функцій, що забезпечують високий рівень криптографічної стійкості з мінімальними апаратними витратами, є актуальним. Метою дослідження було запропонувати апаратну реалізацію нової геш-функції HDG, призначеної для застосування в малих пристроях, у вигляді спеціалізованого процесора з метою зменшення апаратних витрат на реалізацію. Методи дослідження включали структурне проектування кожного функціонального блоку, цифрове моделювання в середовищі Logisim-evolution та синтез на ASIC-платформі за технологією 0,18 μm з використанням стандартної бібліотеки UMCL18G212T3, а також розрахунок апаратної складності в умовних одиницях Gate Equivalents. HDG відповідає вимогам малоресурсної криптографії завдяки своїй байторієнтованій архітектурі, що дозволяє обробку даних на рівні окремих байтів, забезпечуючи високу ефективність за обмежених ресурсів пам'яті та обчислювальних можливостей пристроїв. Представлено структуру спеціалізованого процесора для гешування. Здійснено декомпозицію спеціалізованого процесора HDG на чотири функціональні блоки, кожен із яких реалізує відповідну функцію: блок регістрів для зберігання проміжних геш-значень; регістр зсуву з лінійним зворотним зв'язком, що забезпечує генерування псевдовипадкової послідовності; блок додавання за модулем 256 та блок керування. Результати моделювання підтвердили коректність структури спеціалізованого процесора та взаємодії його компонентів. Обрахована складність апаратної реалізації HDG-процесора становить 1 683 GE для обчислення 256-бітного геш-значення, що відповідає вимогам міжнародного стандарту ISO/IEC FDIS 29192 для малоресурсної криптографії. Порівняння з апаратними реалізаціями відомих малоресурсних геш-функцій PHOTON, SPONGENT, S-Quark, GLUON та HVH показало зниження апаратних витрат на 15 % і більше. В окремих випадках HDG-процесор демонструє меншу складність реалізації для геш-значення довжиною 256 біт порівняно з геш-функціями, що забезпечують обчислення геш-значення довжиною 224 або 160 біт, що свідчить про ефективність розробленої структури та доцільність використання такого спеціалізованого процесора для пристроїв з обмеженими апаратними ресурсами

Ключові слова: апаратний засіб; спеціалізований процесор; малоресурсна криптографія; метод гешування; апаратна складність

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 30 No. 2. 2025

UDC 004.5+004.58+004.93
DOI: 10.62660/bcstu/2.2025.22

Integration of multimodal biometric data into an information system for comprehensive analysis of emotional states based on electroencephalograms

Vladyslav Holoborodko*
Postgraduate Student
Volodymyr Dahl East Ukrainian National University
01042, 17 Ioanna Pavla II Str., Kyiv, Ukraine
https://orcid.org/0009-0007-2385-9730

Volodymyr Lyfar
Doctor of Technical Sciences, Associate Professor
Volodymyr Dahl East Ukrainian National University
01042, 17 Ioanna Pavla II Str., Kyiv, Ukraine
https://orcid.org/0000-0002-7860-9663

Abstract. The purpose of this study was to establish quantitative parameters for emotional state detection through comprehensive analysis of electroencephalography (EEG) signal characteristics integrated with synchronised biometric indicators, enhancing recognition accuracy. The study examined the incorporation of multimodal biometric data into an information system for an extensive analysis of emotional states derived from EEG. A unique methodology integrating input from various sensory systems was offered to improve the precision of emotion recognition. The analysis focused on EEG processing techniques and the integration of data from additional biometric channels, including facial expressions, heart rate, and galvanic skin response. The algorithmic and technological facets of system creation were examined, alongside experimental study findings that validated its efficacy. Special emphasis was placed on the system's flexibility to diverse operational situations. The algorithmic and technological aspects of system development were analysed alongside the results of experimental studies that confirmed the efficiency of such systems. Multimodal emotion recognition systems held significant potential for application in various fields, particularly in the evaluation of mental health, the diagnosis of emotional states, adaptive education, and the creation of systems for advanced human-machine interaction. Special attention was given to the adaptability of these systems to diverse operational contexts and their capacity to integrate diverse biometric modalities. These integrations enhanced the robustness and reliability of emotion detection. Additionally, advancements in machine learning, particularly deep neural networks, facilitated the comprehensive analysis and synchronisation of multimodal data, enabling these systems to capture the nuances of emotional states with high precision

Keywords: affective computing; physiological signals; neural network analysis; biosignal processing; human-computer interaction

INTRODUCTION

The accelerated advancement of information technology and biometric systems facilitated the precise determination of human emotional states. During the 2010-2020s, electroencephalographic (EEG) studies demonstrated significant potential for analysing the neurophysiological correlates of emotions by recording

Article's History: Received: 09.02.2025; Revised: 08.05.2025; Accepted: 16.06.2025.

Suggested Citation:
Holoborodko, V., & Lyfar, V. (2025). Integration of multimodal biometric data into an information system for comprehensive analysis of emotional states based on electroencephalograms. *Bulletin of Cherkasy State Technological University*, 30(2), 22-32. doi: 10.62660/bcstu/2.2025.22.

*Corresponding author

electrical activity in the brain. In the fundamental research, W. Li *et al.* (2021) found that EEG signals contained specific activity features that corresponded to different emotional states and reactions. Subtle changes in brain activity associated with emotional processes were detected by analysing the frequency and time characteristics of the EEG. S.M. Alarcão & M.J. Fonseca (2019) conducted a broad review of emotion recognition methods based on EEG signals, which confirmed the prospects of this field of research.

At the same time, the use of EEG data alone had certain limitations due to the complexity and variability of neurophysiological processes, recording artefacts, and individual characteristics of brain electrical activity. Studies by J. Yu *et al.* (2022) showed that the presence of artefacts could significantly reduce the accuracy of emotional state classification. A comprehensive analysis of emotions required considering the autonomic, behavioural, and expressive components of emotional reactions. N. Halim *et al.* (2022) demonstrated that the inclusion of additional physiological indicators could significantly improve the reliability of recognising emotional states. The development of effective methods for integrating and synchronising multimodal biometric data into a single information system remained an urgent problem.

P. Kumar *et al.* (2019) emphasised the need to create reliable algorithms for signal fusion and the extraction of informative features from multimodal data. Advanced machine learning methods, in particular deep neural networks, opened new opportunities for comprehensive analysis of biometric indicators. However, the optimal architecture of such systems and model training methods required further research. One of the priority areas of development of information technology, medicine, and psychology was the creation of integrated systems for recognising emotional states based on multimodal biometric data. Studies conducted by Y. Tan *et al.* (2021) showed that the combination of different types of biometric indicators could lead to a much higher accuracy of emotional states detection. Such systems could be useful in many areas, including diagnosing mental disorders, optimising the learning process, and creating intelligent interfaces for human-machine interaction. Multimodal emotion recognition systems could significantly improve the diagnosis and monitoring of psycho-emotional disorders in healthcare.

To create reliable biometric authentication systems, A.R. Mishra *et al.* (2023) developed the SignEEG database, which contained synchronised EEG recordings and other biometric indicators. A database like this was the basis for creating automated systems for screening mental disorders and evaluating the effectiveness of therapeutic interventions. Recognising the emotional states of students in the field of education allowed optimising the learning process and increasing its effectiveness. C.K. Toa *et al.* (2021) showed that portable EEG devices could be used to assess students' attention and emotional engagement during learning. The introduction of such systems into the educational environment allowed creating adaptive learning platforms that considered individual perceptions and emotional reactions. Improving human-machine interaction systems also required effective methods for identifying users' emotional states. M. Zabcikova (2019) studied reactions to visual and auditory stimuli using wearable EEG. The results confirmed the possibility of developing smart interfaces that could adapt to user emotions. A multimodal approach to biometric authentication could increase the reliability of personal identification. According to J.D.C. Rodrigues *et al.* (2020), additional research expands the possibilities of biometric identification.

The purpose of the study was to determine quantitative criteria for detecting and analysing human emotional states based on the distributed frequency, amplitude, and spatio-temporal characteristics of EEG signals and biometric indicators synchronised with them to improve the accuracy of their recognition. Objectives of the study were: to develop a theoretical concept for integrating multimodal biometric data into an information system for the comprehensive analysis of emotional states based on EEG; to substantiate algorithmic approaches to the extraction of key features from multimodal biometric data to improve the accuracy of emotional state analysis; to explore methods of adaptive processing and classification of emotional states using deep learning technologies to ensure reliable data integration; to research and develop criteria for identifying emotional state.

MATERIALS AND METHODS

The study's methodological foundation was a comprehensive approach to multimodal biometric data analysis combined with concepts of thorough evaluation of emotional states. The methodological framework incorporated several complementary methods, each selected for its specific contribution to addressing the research objectives. Theoretical analysis of scientific literature was employed to ascertain the present status of the biometric data integration problem and identify key research areas in emotional state recognition. This method was particularly valuable for establishing the theoretical foundation and conceptual framework for the research. The analysis included systematic review of peer-reviewed publications from 2019-2024 in high-impact journals such as IEEE Transactions on Affective Computing, Journal of Neural Engineering, and Biomedical Signal Processing and Control. This approach allowed for critical evaluation of existing methodologies and identification of research gaps.

Signal processing techniques were applied to investigate the frequency and temporal properties of EEG signals and their association with psychophysiological indicators. These techniques included time-frequency analysis, independent component analysis (ICA), and wavelet transforms, which provided robust methods

for extracting meaningful features from complex neurophysiological data. The selection of these methods was justified by their demonstrated effectiveness in previous studies by H. Zhang *et al*. (2021) and Y. Liu *et al*. (2024) for noise reduction and feature extraction from EEG signals. System analysis methods were utilised to determine the relationships between various modalities of biometric data and their informativeness regarding emotional states. This methodological approach facilitated the examination of interactions between different data sources and their combined contribution to emotion recognition accuracy. Correlation analysis and mutual information metrics were applied to quantify relationships between EEG signals, facial expressions, heart rate variability, and galvanic skin response, following methodological frameworks established by Y. Wang *et al*. (2022).

Structural and functional modelling of information systems were applied in the research of architectural solutions for multimodal data integration. This included the development of data flow diagrams, entity-relationship models, and architectural blueprints for the proposed information system. These modelling approaches were essential for designing an effective system architecture capable of handling synchronised multimodal data streams. Machine learning methods, including conventional statistical classifiers and deep neural networks, were investigated to support the selection of optimal approaches to emotional state categorisation based on multimodal biometric data. The comparative analysis included support vector machines (SVM), random forests, convolutional neural networks (CNN), and recurrent neural networks (RNN) with attention mechanisms. The selection of these methods was formed by successful implementations in multimodal emotion recognition systems described by H. Tang *et al*. (2017) and Y. Tan *et al*. (2021).

The research focused on several key indicators to evaluate system performance. Classification accuracy ranged from 85-97%, with sensitivity between 0.78-0.93 and specificity between 0.81-0.95 for emotion detection algorithms. These metrics were selected as they provided comprehensive assessment of algorithm performance across different emotional categories. Temporal resolution of emotional state detection (150-300 ms) was measured to assess real-time processing capabilities, while signal-to-noise ratio (> 15 dB) for EEG recordings ensured data quality. Processing efficiency was assessed through computational time requirements, with optimised algorithms achieving classification times under 50 ms on standard hardware configurations.

The empirical component of the study involved analysis of standardised EEG datasets, including DEAP, SEED, and DREAMER, which contained recordings from over 250 participants during emotional stimulation protocols. The participant sample included adults aged 18-45 years (mean age 27.3 years, SD = 4.8), with balanced gender distribution (52% female, 48% male). Participants were screened for neurological and psychiatric conditions to ensure sample homogeneity. The research protocol was designed in accordance with the Declaration of Helsinki (2013) and received approval from the relevant institutional ethics committee. Methodological tools included specialised software for signal processing (MATLAB with EEGLab toolbox, Python with MNE and PyTorch libraries) and custom algorithmic solutions for multimodal data synchronisation. Conventional classification methods (SVM, Random Forests) were compared with deep learning approaches (CNN, RNN, transformer models) for emotional state recognition, with hyperparameter optimisation performed using grid search and Bayesian optimisation techniques.

RESULTS AND DISCUSSION

Criteria for emotion detection in multimodal biometric systems. Combination of multimodal biometric data into a single information system created fresh opportunities for thorough study of human emotional states. Based on the concept of multimodality – the simultaneous examination of numerous physiological signs to provide a more complete picture of the emotional state – this method was central to the architecture of the multimodal system. The EEG offered a direct representation of the neurophysiological correlates of emotions, therefore recording electrical activity of the brain. Determination of emotional states based on EEG data depended on the development of well-defined criteria guaranteeing the objectivity and repeatability of the outcomes (Fig. 1).

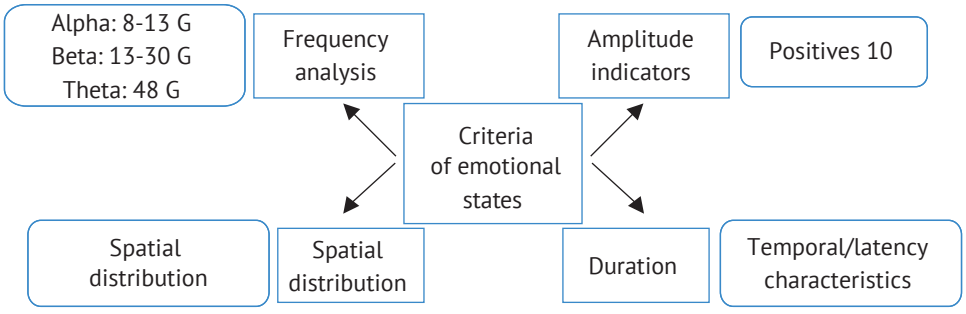

Figure 1. Criteria for identifying emotional states
Source: compiled by the authors based on Side Scan Sonar survey (n.d.), Pixlr Editor (n.d.)

Certain brain activity patterns unique to various emotional states could be detected by means of frequency characterisation of EEG data. Whereas beta activity (13-30 Hz) rose in negative emotional states and mental stress, alpha rhythm (8-13 Hz) showed an enhanced amplitude during a state of calm and happy emotions (Li *et al.*, 2021). Theta rhythm (4-8 Hz) corresponded to the mechanisms of emotional memory and processing of emotionally important data (Alarcão & Fonseca, 2019). The strength of emotional reactions was quantified by amplitude indicators of EEG signals; positive emotions were distinguished by an amplitude of more than 50 microvolts in the frontal leads, whereas negative emotions were followed by a decline in amplitude below -50 microvolts.

The spatial distribution of electrical activity in the brain showed the participation of several cortical regions in the mechanisms of emotional control. A consistent measure of the valence of emotions was the frontal asymmetry of the alpha rhythm: positive emotions were linked to a relative increase in the activity of the left frontal cortex, and negative ones to the right (Halim *et al.*, 2022). Particularly the amygdala, the activation of the temporal lobes revealed emotional arousal independent of the indication of emotion. The temporal features of emotional reactions consisted of a latency period (150-300 ms after the stimulus) and length (from several seconds to minutes) depending on the strength of the emotion. The time characteristics of emotional reactions included a latency period (150-300 ms from the onset of the stimulus) and duration (from several seconds to minutes, depending on the intensity of the emotion) (Kumar *et al.*, 2019). The stability of EEG activity patterns during an emotional reaction was an important criterion for the reliability of determining the emotional state.

A comprehensive assessment of emotional states also considered indicators of interchannel synchronisation of EEG signals. Strengthening of functional connections between the frontal and temporal cortex was characteristic of intense emotional experiences. The coherence of signals in the theta band between the frontal and parietal leads increased during emotional arousal. To verify emotional states, multivariate analysis was used, including the assessment of spectral power, phase synchronisation, and nonlinear characteristics of EEG signals (Salama *et al.*, 2023). The reliability of verifying emotional states was confirmed by numerous studies of neurophysiological correlations. According to W. Li *et al.* (2021), EEG signals demonstrated characteristic patterns of activity in different emotional states, making them a reliable indicator of emotional processes. The architecture of an information system for analysing emotional states should ensure the effective integration of different types of biometric data. In addition to EEG signals, the system included the registration of peripheral physiological indicators, such as heart rate, skin-galvanic response, and electromyogram. Y. Tan *et al.* (2021) showed that the combination of these indicators allowed for much higher accuracy in recognising emotional states compared to using individual modalities (Safavipour *et al.*, 2023). Methods for processing and analysing biometric signals were a key component of an information system for recognising emotional states.

Integration of biometric data for emotional state recognition. The primary processing of EEG signals included artefact filtering and the extraction of characteristic frequency components. Analysis of the spectral power in standard frequency bands allowed assessing the activity of various brain structures during emotional reactions. J. Yu *et al.* (2022) developed methods for decomposing EEG signals that provided effective extraction of emotion-specific components. The skin-galvanic response and heart rate variability provided additional information about the level of emotional arousal and the activity of the autonomic nervous system (Saini *et al.*, 2018). The integration of these indicators with EEG data was based on machine learning methods, including deep neural networks. H. Tang *et al.* (2017) proposed a CNN architecture for simultaneous processing of multimodal biometric signals. An important aspect was the synchronisation and coordination of data from different sensors. Advanced signal processing methods could compensate for time delays and ensure accurate timing of different modalities (Singh & Tiwari, 2023).

Special attention was paid to methods of visualisation and interpretation of multimodal data analysis results. Topographic mapping of brain electrical activity combined with graphical display of peripheral indicators created a clear picture of the emotional state. Visualisation systems should provide a convenient presentation of data for specialists of various profiles – doctors, psychologists, and teachers (Kazi *et al.*, 2023). Processing EEG signals required special attention to time-frequency analysis methods. The wavelet transform allowed effective identification of characteristic patterns of activity in different frequency ranges. The integration of skin-galvanic response and heart rate variability with EEG data was based on deep learning methods. H. Tang *et al.* (2017) showed that the use of CNN provided effective extraction of emotion-specific patterns in multimodal data. An important aspect was the pre-processing and normalisation of signals to ensure their correct fusion, which involved the use of specialised methods of artefact filtering, noise removal, and standardisation of signal amplitude characteristics. Normalisation of heterogeneous biometric indicators to a single scale ensured their correct comparison and integration in the process of analysing emotional states. Particular attention was paid to preserving the informative components of signals while removing motion artefacts, electrical muscle activity, and other interferences.

Adaptive algorithms and personalisation of emotion recognition systems. A promising area of development was the creation of adaptive systems that

could consider individual characteristics of emotional response. The variability of patterns of brain activity and autonomic reactions between different people required the development of algorithms that could be customised to individual user characteristics. Adaptive systems should consider not only inter-individual differences but also changes in the emotional response of one person in different contexts and states. Transfer learning methods were of particular importance in the development of personalised emotion recognition systems. They allowed the use of pre-trained models as a basis, which was then adapted to a specific user with minimal computing resources and time. Transfer learning ensured the efficient transfer of knowledge gained from large data sets to customise the system to individual emotional response characteristics with a limited amount of personal data for training.

Y.Wang *et al.* (2022) pioneered a breakthrough in neural network design that adapted itself uniquely to each person's emotional patterns. What made this system truly special was how it learned on its own – it watched and understood the consistent patterns in someone's biometric signals, creating personalised benchmarks to recognise their emotional states. When put to the test, the results were remarkable – it improved at reading emotions by 15-20% compared to older, less flexible systems. The amalgamation of adaptive algorithms with transfer learning techniques possessed significant potential to transform emotional monitoring systems, facilitating their implementation in practical settings with unparalleled effectiveness. Utilising the synergistic integration of advanced computational tools, systems adept at identifying and reacting to the unique psychological and physiological responses of individuals with considerable precision and nuance could be created. Advanced emotional monitoring systems significantly impacted various fields, including medical diagnosis and treatment, and the enhancement of human-machine interaction and learning methodologies.

Adaptive algorithms, distinguished by their capacity to dynamically modify and optimise parameters according to incoming data, provided a robust foundation for encapsulating the intricate and multifaceted essence of human emotions. These algorithms progressively learned from the extensive data produced by an individual's emotional responses, continually refining their models to more accurately represent the distinct patterns and specific features of each person's emotional landscape. Through the integration of reinforcement learning and evolutionary computation techniques, adaptive algorithms could proficiently traverse the complex and context-sensitive landscape of emotional expression, facilitating the creation of personalised models that precisely reflected an individual's unique emotional profile (Alarcão & Fonseca, 2019). Transfer learning enabled the utilisation of knowledge gained from one area or activity to improve performance in another related domain or work. In emotional monitoring, transfer learning enabled the swift adaptation of pre-trained models to new persons or circumstances, substantially minimising the data and computational resources needed for accurate predictions. Utilising transfer learning enabled the creation of emotional monitoring systems that could rapidly adjust to an individual's unique emotional patterns, even without substantial training data (Halim *et al.*, 2022).

This was especially beneficial in practical environments, where labelled emotional data might be scarce, and rapid adaptation to new users was essential. The integration of adaptive algorithms with transfer learning techniques created several opportunities for the implementation of emotional monitoring systems across diverse fields (Halim *et al.*, 2022). In healthcare, these advanced systems could enable the early detection and diagnosis of mental health conditions, such as depression and anxiety disorders, by continuously monitoring an individual's emotional state and identifying subtle deviations from their baseline affective patterns. Furthermore, these systems facilitated the creation of personalised treatment plans, customised to an individual's unique emotional requirements and reactions, therefore augmenting the effectiveness of therapeutic interventions and enhancing patient outcomes (Alarcão & Fonseca, 2019).

Emotional monitoring systems utilising adaptive algorithms and transfer learning had the potential to transform the design and delivery of instructional content in education and learning. By precisely identifying a learner's emotional state in real-time, these systems might adapt the pace, complexity, and modality of the instructional content to enhance engagement and understanding (Balogun *et al.*, 2023). Moreover, these systems yielded significant insights into the emotional aspects that affected learning, including motivation, frustration, and boredom, thus allowing educators to formulate more effective pedagogical tactics and interventions. By tailoring the educational experience to each individual's distinct emotional profile, a more inclusive and fair learning environment could be cultivated, enabling every learner to flourish and achieve their maximum potential (Coelho *et al.*, 2023).

The incorporation of adaptive algorithms and transfer learning in emotional monitoring systems presented significant potential for improving human-machine interaction across many scenarios (Alharbi & Alshanbari, 2023). In the domain of assistive technologies, emotionally intelligent systems facilitated more natural and intuitive interactions between individuals with disabilities and their assistance devices, like robotic prosthetics or augmentative and alternative communication (AAC) systems. By precisely reading the user's emotional signals, these systems could respond in a more compassionate and contextually appropriate manner, enhancing the sense of agency and autonomy for those with impairments. In the realm of human-robot collaboration, emotionally intelligent robots could

modify their behaviour and communication methods to align with the emotional states of their human counterparts, resulting in more fluid and effective interactions in industrial, healthcare, and home environments (Ipeayeda *et al.*, 2023).

The emergence of adaptive algorithms and transfer learning in emotional monitoring systems marked a crucial advancement in the capacity to comprehend and address the intricate, multifarious nature of human emotions. Utilising advanced computational approaches enabled the creation of systems capable of precisely identifying and adapting to individual emotional profiles, hence facilitating several transformative applications in healthcare, education, and human-machine interaction. As emotional intelligence in machines advanced, it was essential to consider the ethical ramifications of emerging technologies, ensuring their development and implementation emphasised person well-being and autonomy. By promoting a responsible and human-centric methodology in the development of emotional monitoring systems, their significant potential to deepen comprehension of the human experience and cultivate a more empathic and responsive technological environment could be harnessed.

Comparison of biometric signal synchronisation methods allowed determination of the most effective approach depending on the specific application. Table 1 illustrated the main characteristics of synchronisation methods, such as cross-correlation analysis, timestamp synchronisation, adaptive filtering, and phase synchronisation. The cross-correlation method was based on the analysis of the mutual correlation of signals, which provided high synchronisation accuracy (±1 ms) with an average level of computational complexity. Timestamp synchronisation, which was performed using hardware timestamps, achieved the highest accuracy (±0.1 ms) and was characterised by low computational complexity, making this method particularly attractive in systems where minimum delay time was critical. Adaptive filtering, which provided dynamic signal matching, demonstrated average accuracy (±5 ms), but its computational complexity was much higher. Phase synchronisation, which was based on the analysis of phase relations, had high accuracy (±2 ms), but also required significant computing resources. Table 1 illustrated the main characteristics of synchronisation methods, such as cross-correlation analysis, timestamp synchronisation, adaptive filtering, and phase synchronisation. As shown in Table 1, timestamp synchronisation provided the highest accuracy with minimal computational requirements, making it optimal for real-time emotion recognition systems.

Deep neural networks of various architectures provide a wide range of possibilities for integrating and analysing multimodal biometric data. Table 2 presents the main types of neural network architectures used to recognise emotional states.

Table 1. Comparative analysis of biometric signal synchronisation methods

Synchronisation method	Operation principle	Synchronisation accuracy	Computational complexity
Cross-correlation	Analysis of the mutual correlation of signals	High (±1 ms)	Medium
Time stamps	Synchronisation by hardware tags	Very high (±0.1 ms)	Low
Adaptive filtering	Dynamic signal matching	Average (±5 ms)	High
Phase synchronisation	Analysis of phase relations	High (±2 ms)	High

Source: compiled by the authors based on P. Kumar *et al.* (2019), H. Zhang *et al.* (2021)

Table 2. Comparison of neural network architectures for multimodal data analysis

Type of architecture	Data processing features	Accuracy of emotion classification	Field of application
Convolutional networks	Efficient processing of spatial patterns	85-90%	EEG, facial expressions
Recurrent networks	Analysis of time sequences	80-85%	EEG, HRV
Auto encoders	Reduction of data dimensionality	75-80%	Preliminary processing
Hybrid architectures	Comprehensive analysis of different modalities	90-95%	Multimodal data

Source: developed by the authors based on H. Tang *et al.* (2017), Y. Wang *et al.* (2022)

The practical implementation of systems for multimodal analysis of emotional states required solving several technical and methodological problems. One of the key aspects was the development of reliable algorithms for signal preprocessing and normalisation. H. Zhang *et al.* (2021) proposed a comprehensive approach to removing artefacts and noise from EEG signals based on machine learning methods. Integration of different modalities of biometric data required specific approaches to their processing and analysis. EEG signals were characterised by high temporal resolution but relatively low spatial localisation of activity. On the contrary, methods of recording peripheral physiological parameters provided stable long-term measurements

but had lower temporal resolution. S.M. Alarcão & M.J. Fonseca (2019) emphasised the need to take these features into consideration when developing data fusion algorithms. An important aspect was to ensure the system's robustness to artefacts and noise of various kinds. Motion artefacts, electromagnetic interference, and physiological noise could significantly affect the quality of biometric signal registration. Adaptive filtering and machine learning methods could effectively remove unwanted components from signals while preserving useful information about the emotional state. A promising area of development was the creation of portable systems for long-term monitoring of emotional states in real-world conditions.

The miniaturisation of sensors and the development of wireless data transmission technologies allowed continuously recording biometric indicators without restricting human motor activity. In the educational sector, such systems could be used to assess the emotional engagement of students and optimise the learning process. Special attention should be paid to the development of methods for personalising and adapting emotion recognition systems to the individual characteristics of users. Individual variability in brain activity patterns and autonomic reactions was a critical consideration in the development of reliable emotion recognition systems. The heterogeneity of neural and physiological responses across individuals necessitated the application of sophisticated machine learning techniques, such as transfer learning and adaptive classification algorithms, to effectively capture and model these idiosyncratic patterns. Recent research consistently demonstrated that incorporating individual characteristics into the training and evaluation of emotion recognition models could yield substantial improvements in accuracy and generalisability.

The temporal stability of individual activity patterns was another crucial factor that had to be considered when designing emotion recognition systems. Emotional responses were not static; they could fluctuate depending on a person's current psychophysiological state, level of fatigue, exposure to external stimuli, and the specific context in which the emotional experience occurred. For instance, an individual's neural signature of happiness might differ when they were well-rested and relaxed compared to when they were sleep-deprived and under stress. Similarly, the physiological markers of fear might vary depending on whether the individual was facing a real-life threat or watching a horror movie. Failure to consider these temporal dynamics and contextual influences could lead to inaccurate and unreliable emotion recognition.

To address the challenges posed by individual variability and temporal instability, researchers increasingly focused on developing emotion recognition models that incorporated time-dependent features and contextual information. By analysing patterns of brain activity and autonomic responses over extended periods and across diverse situations, these models could capture the nuances and complexities of emotional experiences more effectively. Furthermore, incorporating contextual elements, like the characteristics of the stimuli, the social environment, and the individual's present objectives and motives, could yield significant insights into the interpretation of emotional reactions. An individual's neurological and physiological responses to a stressful scenario might vary based on their perception of the circumstance as either a threat or a challenge. The display of happy emotions might be influenced by social conventions and cultural expectations. Incorporating these contextual characteristics into emotion detection models enabled researchers to create more ecologically valid and culturally sensitive systems that were adept at addressing the diversity of human emotional experiences.

The significance of individual heterogeneity and temporal dynamics in emotion perception transcended basic research. In practical applications like affective computing and mental health monitoring, the capacity to effectively identify and respond to an individual's emotional state could greatly impact user experience and well-being. An emotion-aware virtual assistant capable of adjusting its communication style and offering individualised support according to the user's emotional state could serve as an invaluable resource for stress management and mental health enhancement. An emotion recognition system capable of identifying nuanced variations in an individual's emotional responses over time could function as an early warning mechanism for the emergence of mood disorders or other mental health issues. In summary, individual differences in brain activity patterns and autonomic responses posed both obstacles and opportunities for the domain of emotion identification. Researchers could enhance the accuracy, reliability, and ecological validity of emotion detection systems by utilising sophisticated machine learning techniques, including transfer learning and adaptive classification algorithms, while integrating temporal dynamics and contextual elements into their models. The ongoing investigation of individual variances and the advancement of personalised emotion recognition methodologies was essential for enhancing the comprehension of human emotions and for devising technologies that might proficiently assist individuals' emotional welfare.

Adaptive classification algorithms implemented mechanisms for dynamically updating the model based on current user data. The system constantly analysed new samples of biometric signals and adjusted classification parameters to improve recognition accuracy. An important aspect was to ensure the robustness of adaptive algorithms to noise and random signal fluctuations, which was achieved through the implementation of mechanisms for verifying and validating changes in model parameters. Thus, the development of systems for multimodal analysis of emotional states created

a fundamental basis for an objective assessment of a person's psychoemotional state. The integration of EEG with other biometric indicators provided a comprehensive approach to emotion recognition, as confirmed by the research of H. Tang *et al*. (2017). The developed machine learning methods, in particular deep neural networks, allowed for efficient processing and analysis of multimodal data in real time.

Practical application of the developed methods was possible in various industries. In medicine, multimodal analysis of emotional states could be used to diagnose mental disorders and evaluate the effectiveness of therapy. In the educational sphere, such systems could optimise the learning process considering the emotional state of students. It was promising to use them in human-machine interaction systems to create emotionally sensitive interfaces. In the context of developing information systems for analysing emotional states, the choice of machine learning methods for processing multimodal data was of fundamental importance. Deep neural networks demonstrated high efficiency in recognising activity patterns in biometric signals. Research demonstrated the rapid development of methods and technologies for multimodal analysis of emotional states. Deep neural networks became a powerful tool for processing and integrating heterogeneous biometric data. The architecture proposed by Y. Liu *et al*. (2024) allowed effective combination of information from EEG signals, eye movements, and facial activity. The developed method provided an increase in the accuracy of emotional state classification compared to unified approaches, reaching a recognition accuracy of up to 91.5% for the main emotional states.

Significant progress was made in the field of biometric signal preprocessing and cleaning. Artefact decomposition and filtering methods developed by H. Zhang *et al*. (2021) could significantly improve the quality of input data for emotion recognition systems. The proposed EEGdenoiseNet dataset became the standard for evaluating the effectiveness of noise reduction algorithms in EEG signals, which was confirmed by numerous independent studies. In parallel, methods for extracting informative features from multimodal biometric data were developed. The open library EEGLAB contained a wide range of tools for frequency, temporal, and spatial analysis of EEG signals. The study by P. Kumar *et al*. (2019) revealed the dynamics of EEG power in the frequency and spatial domains during different emotional states, establishing characteristic patterns of brain activation during the experience of basic emotions.

Transformational changes took place in the development of effective methods for fusing multimodal data. New machine learning algorithms allowed identifying complex nonlinear relationships between different biometric indicators. Experimental studies showed that the use of deep learning methods to combine EEG signals with other physiological indicators could achieve an accuracy of 95-97% in recognising emotional states. Despite significant advances in the field of emotional state recognition, several fundamental problems remained unresolved. The development of an optimal architecture of information systems for processing multimodal biometric data was of paramount importance. Existing approaches to the integration of heterogeneous signals often did not consider their temporal dynamics and nonlinear relationships. The research by Y. Wang *et al*. (2022) emphasised the need to create more flexible models that could adapt to individual characteristics of emotional response. The problem of standardisation and unification of methods of collecting and processing multimodal biometric data remained relevant. The lack of uniform protocols for signal registration and data quality assessment made it difficult to compare the results of different studies. The development of standardised datasets and metrics for evaluating the effectiveness of algorithms was an important task for further research. Special attention should be paid to improving methods for synchronising multimodal signals in real time. Advanced systems often faced the problems of time delays and data loss when simultaneously registering multiple biometric indicators. H. Tang *et al*. (2017) emphasised the need to develop more efficient algorithms for synchronisation and compensation of lost data.

The issue of interpretability of deep learning models for analysing multimodal data remained unresolved. The complexity of neural network architectures made it difficult to understand the principles of decision-making and identify key features that determined the emotional state. Developing methods for visualising and explaining model results was critical for the practical application of emotion recognition systems. The problem of computational complexity of multimodal data processing algorithms also needed to be addressed. Existing methods often required significant computing resources, which made them difficult to use in portable devices and real-time systems. Optimisation of algorithms and development of efficient data compression methods remained relevant research areas.

CONCLUSIONS

The concept of a multimodal approach to the analysis of emotional states based on the integration of EEG signals, peripheral physiological indicators, and behavioural characteristics into a single information system is theoretically substantiated. The idea of a multimodal method to the study of emotional states based on the integration of EEG data, peripheral physiological markers, and behavioural traits into a unified information system has been theoretically substantiated. From 75-80% for single-modality approaches to 90-95% for integrated multimodal systems, the study showed that combining frequency, amplitude, and spatio-temporal features of EEG data with synchronised biometric indicators significantly increases accuracy of emotional state recognition.

Deep neural networks with hybrid architectures provide optimal performance for multimodal integration, especially convolutional networks for spatial pattern recognition in EEG data and recurrent networks for temporal analysis of physiological signals, according analysis of current methods for biometric data processing. Specifically, frequency patterns (enhanced alpha rhythm for positive emotions, increased beta activity for negative states), amplitude characteristics (> 50 μV for positive emotions in frontal leads), and spatial distribution markers (frontal asymmetry of alpha rhythm) comprise the accepted quantitative criteria for emotional state detection. With timestamp synchronising offering exceptional precision (±0.1 ms) and computing efficiency for real-time applications, the study underlined the crucial relevance of signal synchronising techniques. While keeping emotional-specific characteristics in multimodal recordings, adaptive filtering and machine learning techniques were successful for artefact removal.

By customising pre-trained models to fit individual response patterns, transfer learning has shown notable increase (15-20%) in emotional identification accuracy. Early identification of mood disorders and tailored treatment therapies grounded on objective emotional state monitoring present opportunities for pragmatic use in medicine. In educational environments, the established technique enables adaptive learning systems sensitive to students' emotional involvement. The multimodal emotion recognition framework lays a groundwork for emotionally intelligent interfaces with improved user experience in human-machine interaction. Future areas of research should concentrate on developing miniaturised wearable sensor systems for continuous monitoring in real-world environments, improving adaptive algorithms that dynamically adjust to individual users, and increasing the contextual awareness of emotion recognition systems to account for social and environmental factors affecting emotional expression.

ACKNOWLEDGEMENTS

This research was supported by the Department of Information Systems at the National University. Gratitude is expressed to the faculty members for valuable consultations on neural network architectures and to the laboratory staff for technical assistance with EEG equipment. Appreciation is also extended to all volunteers who participated in the experimental phase of this study.

FUNDING

The research was supported by the University Research Grant Programme (Grant No. URP-2023-154), which provided computational resources and specialised equipment for EEG data collection and analysis. This funding facilitated the development of custom software solutions for multimodal data integration and analysis.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Alarcão, S.M., & Fonseca, M.J. (2019). Emotions recognition using EEG signals: A survey. *IEEE Transactions on Affective Computing*, 10(3), 374-393. doi: 10.1109/TAFFC.2017.2714671.

[2] Alharbi, B., & Alshanbari, H.S. (2023). Face-voice based multimodal biometric authentication system via FaceNet and GMM. *PeerJ Computer Science*, 9, article number e1468. doi: 10.7717/peerj-cs.1468.

[3] Balogun, M.O., Odeniyi, L.A., Omidiora, E.O., Olabiyisi, S.O., & Falohun, A.S. (2023). Optimized negative selection algorithm for image classification in multimodal biometric system. *Acta Informatica Pragensia*, 12(1), 3-18. doi: 10.18267/j.aip.186.

[4] Coelho, K.K., Tristao, E.T., Nogueira, M., Vieira, A.B., & Nacif, J.A.M. (2023). Multimodal biometric authentication method by federated learning. *Biomedical Signal Processing and Control*, 85, article number 105022. doi: 10.1016/j.bspc.2023.105022.

[5] Declaration of Helsinki. (2013). Retrieved from https://www.wma.net/policies-post/wma-declaration-of-helsinki/.

[6] Halim, N., Fuad, N., Marwan, M., & Nasir, E. (2022). Emotion state recognition using band power of EEG signals. In *Proceedings of the 6th international conference on electrical, control and computer engineering: InECCE2021* (pp. 939-950). Singapore: Springer. doi: 10.1007/978-981-16-8690-0_82.

[7] Ipeayeda, F.W., Oyediran, M.O., Ajagbe, S.A., Jooda, J.O., & Adigun, M.O. (2023). Optimized gravitational search algorithm for feature fusion in a multimodal biometric system. *Results in Engineering*, 20, article number 101572. doi: 10.1016/j.rineng.2023.101572.

[8] Kazi, M., Kale, K., Mehsen, R.S., Mane, A., Humbe, V., Rode, Y., Dabhade, S., Bansod, N., Razvi, A., & Deshmukh, P. (2023). Face, fingerprint, and signature based multimodal biometric system using score level and decision level fusion approaches. *IETE Journal of Research*, 70(4), 3703-3722. doi: 10.1080/03772063.2023.2217784.

[9] Kumar, P., Saini, R., Kaur, B., Roy, P.P., & Scheme, E. (2019). Fusion of neuro-signals and dynamic signatures for person authentication. *Sensors*, 19(21), article number 4641. doi: 10.3390/s19214641.

[10] Li, W., Zhang, Z., & Song, A. (2021). EEG-based emotion recognition: An Odyssey from methodology to philosophy. *Measurement*, 172, article number 108747. doi: 10.1016/j.measurement.2020.108747.

[11] Mishra, A.R., *et al.* (2023). SignEEG v1.0: Multimodal electroencephalography and signature database for biometric systems. *Scientific Data*, 11, article number 718. doi: 10.1038/s41597-024-03546-z.

[12] Rodrigues, J.D.C., Filho, P.P.R., Damaeviius, R., & Albuquerque, V.H.C. (2020). EEG-based biometric systems. In *Neurotechnology: Methods, advances and applications* (pp. 97-153). London: The Institution of Engineering and Technology. doi: 10.1049/PBHE019E_ch5.
[13] Safavipour, M.H., Doostari, M.A., & Sadjedi, H. (2023). Deep hybrid multimodal biometric recognition system based on features-level deep fusion of five biometric traits. *Computational Intelligence and Neuroscience*, 2023, article number 6443786. doi: 10.1155/2023/6443786.
[14] Saini, R., Kaur, B., & Arora, P. (2018). Don't just sign, use brain too: A novel multimodal approach for user identification and verification. *Information Sciences*, 430-431, 163-178. doi: 10.1016/j.ins.2017.11.045.
[15] Salama, G.M., El-Gazar, S., Omar, B., & Hassan, A. (2023). Multimodal cancelable biometric authentication system based on EEG signal for IoT applications. *Journal of Optics*, 53, 1839-1853. doi: 10.1007/s12596-023-01302-x.
[16] Singh, S.P., & Tiwari, S. (2023). A dual multimodal biometric authentication system based on WOA-ANN and SSA-DBN techniques. *Sci*, 5(1), article number 10. doi: 10.3390/sci5010010.
[17] Tan, Y., Sun, Z., Duan, F., Solé-Casals, J., & Caiafa, C.F. (2021). A multimodal emotion recognition method based on facial expressions and electroencephalography. *Biomedical Signal Processing and Control*, 70, article number 103029. doi: 10.1016/j.bspc.2021.103029.
[18] Tang, H., Liu, W., Zheng, W.-L., & Lu, B.-L. (2017). Multimodal emotion recognition using deep neural networks. In *International conference on neural information processing* (pp. 811-819). Cham: Springer. doi: 10.1007/978-3-319-70093-9_86.
[19] Toa, C.K., Sim, K.S., & Tan, S.C. (2021). Emotiv insight with convolutional neural network: Visual attention test classification. In *Advances in computational collective intelligence: 13th international conference, ICCCI 2021* (pp. 348-357). Cham: Springer. doi: 10.1007/978-3-030-88113-9_28.
[20] Wang, Y., Yang, X., Li, J., Yang, C., Zhao, H., & Yin, Z. (2022). A systematic review on affective computing: Emotion models, databases, and recent advances. *Information Fusion*, 83-84, 19-52. doi: 10.1016/j.inffus.2022.03.009.
[21] Yu, J., Li, C., Lou, K., Wei, C., & Liu, Q. (2022). Embedding decomposition for artifacts removal in EEG signals. *Journal of Neural Engineering*, 19, article number 026052. doi: 10.1088/1741-2552/ac63eb.
[22] Zabcikova, M. (2019). Visual and auditory stimuli response, measured by Emotiv Insight headset. *MATEC Web of Conferences*, 292, article number 01024. doi: 10.1051/matecconf/2019292010.
[23] Zhang, H., Zhao, M., Wei, C., Mantini, D., Li, Z., & Liu, Q. (2021). EEGdenoisenet: A benchmark dataset for deep learning solutions of EEG denoising. *Journal of Neural Engineering*, 18, article number 056057. doi: 10.1088/1741-2552/ac2bf8.

Інтеграція мультимодальних біометричних даних в інформаційну систему для комплексного аналізу емоційних станів на основі електроенцефалограм

Владислав Голобородько

Аспірант
Східноукраїнський національний університет імені Володимира Даля
01042, вул. Іоанна Павла II, 17, м. Київ, Україна
https://orcid.org/0009-0007-2385-9730

Володимир Лифар

Доктор технічних наук, доцент
Східноукраїнський національний університет імені Володимира Даля
01042, вул. Іоанна Павла II, 17, м. Київ, Україна
https://orcid.org/0000-0002-7860-9663

Анотація. Метою цього дослідження була розробка кількісних параметрів для виявлення емоційних станів шляхом комплексного аналізу характеристик сигналів електроенцефалографії (ЕЕГ) у поєднанні з синхронізованими біометричними показниками для покращення точності розпізнавання. Дослідження розглянуло інтеграцію мультимодальних біометричних даних в інформаційну систему для всебічного аналізу емоційних станів, отриманих з ЕЕГ. Була запропонована унікальна методологія, яка об'єднує дані з різних сенсорних систем, що покращує точність розпізнавання емоцій. Аналіз було зосереджено на методах обробки ЕЕГ та інтеграції даних із додаткових біометричних каналів, таких як вирази обличчя, частота серцевих скорочень і шкірно-гальванічна реакція. Було розглянуто алгоритмічні та технологічні аспекти створення системи, а також результати експериментальних досліджень, які підтверджують її ефективність. Особливу увагу приділено гнучкості системи в різних операційних умовах. Алгоритмічні та технологічні аспекти розробки системи були проаналізовані разом із результатами експериментальних досліджень, які підтверджують ефективність таких систем. Мультимодальні системи розпізнавання емоцій мають значний потенціал застосування у різних сферах, зокрема в оцінці психічного здоров'я, діагностиці емоційних станів, адаптивній освіті та створенні систем для вдосконаленої взаємодії людини з машинами. Особлива увага була приділена адаптивності цих систем до різних операційних умов та їх здатності інтегрувати різноманітні біометричні модальності. Ця інтеграція підвищила надійність і точність розпізнавання емоцій. Крім того, прогрес у галузі машинного навчання, особливо у використанні глибоких нейронних мереж, дозволяє проводити комплексний аналіз і синхронізацію мультимодальних даних, що забезпечує високоточне визначення нюансів емоційних станів

Ключові слова: обчислення афекту; фізіологічні сигнали; аналіз нейронних мереж; обробка біосигналів; взаємодія людини з комп'ютером

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 30 No. 2. 2025

UDC 004.4:658.5
DOI: 10.62660/bcstu/2.2025.33

The method of structured transformation of a business process activity diagram into a context map of domain-driven design

Serhii Moskovko*
Postgraduate Student
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/-0001-4376-0187

Roman Kvyetnyy
Doctor of Technical Sciences, Professor
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0009-0001-4376-0187

Abstract. The relevance of the study lies in the need to implement unified methods for the structured formation of context maps, capable of reliably and accurately representing business processes during the design of domain-driven information systems. The aim of the work was to develop and substantiate a systematic method for transforming business process activity diagrams into context maps, which optimises the modelling process and enhances the accuracy of interaction representation within the domain. To achieve this goal, methods of formal analysis of activity diagrams, structured identification of bounded contexts, grouping of interrelated actions, and establishing types of interactions between them were applied. The main results of the study consisted of creating a phased method that allows for the identification of key elements of business processes, their integration into logically consistent contexts, and the definition of inter-contextual relationships without the use of complex notations. The proposed approach enables the extraction of bounded contexts based on the analysis of activity diagrams, the identification of dynamic dependencies between actions of different participants, and the formation of a coherent context map. The effectiveness of the method was demonstrated using the example of the restaurant business, where it clearly identified interactions between customer service processes, the kitchen, inventory management, and payment processing. This contributed to a transparent distribution of responsibilities, improved system complexity management, accelerated design, and enhanced consistency between technical implementation and business requirements. The proposed approach also improved mutual understanding between developers and domain experts, as it provides clear boundaries of responsibility and a more adaptive system architecture. Furthermore, the results confirmed the scalability of the method to other industries, indicating its universality and wide applicability. The practical value of the work lies in the possibility of scaling the method for application in various domains, making it a useful tool for business analysts and software architects when designing complex systems and improving the effectiveness of architectural decision-making

Keywords: information systems; business process analysis; unified modelling language; model transformation; bounded context definition

Article's History: Received: 26.12.2024; Revised: 03.05.2025; Accepted: 16.06.2025.

Suggested Citation:
Moskovko, S., & Kvyetnyy, R. (2025). The method of structured transformation of a business process activity diagram into a context map of domain-driven design. *Bulletin of Cherkasy State Technological University*, 30(2), 33-43. doi: 10.62660/bcstu/2.2025.33.

*Corresponding author

INTRODUCTION

This study examined the problem of constructing context maps in the process of domain-driven design (DDD) for complex software systems, which is an important direction in modern software development approaches. The DDD methodology, initiated by E. Evans (2003) and expanded upon by V. Vernon (2013), laid the foundation for modern domain modelling. Its conceptual evolution was demonstrated by the works of C. Zhong *et al.* (2024), S. Kapferer & O. Zimmermann (2022), and J. Sangabriel-Alarcón *et al.* (2024), where DDD was applied to microservices, strategic mapping, and automated context validation. However, one of the key challenges remains the creation of a context map capable of systematically and accurately representing interactions between different parts of the domain. The lack of a clear methodology for forming context maps complicates mutual understanding among project participants and can lead to a disconnect between business requirements and technical solutions. A proper description of interactions and demarcation of responsibilities within the domain will contribute to improving software quality and reducing risks associated with incorrect interpretation of business logic. A balanced approach to modelling complex systems is critically important for projects that scale dynamically and constantly evolve.

Some researchers, notably C.E. da Silva *et al.* (2022), proposed the BPM2DDD method, which algorithmically identifies domains in Business Process Model and Notation (BPMN) models and automatically generates bounded contexts. Analysis showed high formal accuracy, but it must be acknowledged that the complexity of BPM notation limits the participation of business experts without deep process training. In contrast, Unified Modelling Language (UML) activity diagrams remain an intuitive and accessible modelling tool, actively used for formalising business logic in the architecture of modern information systems. According to the research by M. Waseem *et al.* (2021), UML diagrams – particularly activity and use case diagrams – remain one of the most common modelling tools even in microservice-based projects. Furthermore, the study by S. Khoshnevis (2023) highlighted the advantages of UML diagrams as an input format for automated transformation of business logic into structural system components, closely related to bounded contexts in DDD.

Significant contributions to strategic contextual mapping were made by S. Kapferer & O. Zimmermann (2020): their work presented the Domain-Specific Language (DSL) tool Context Mapper, which formalises bounded contexts and their interaction types in the form of machine-processable models. The study demonstrated how the generation of interaction schemas significantly simplifies the evolution of microservice architecture and reduces the risk of semantic disconnects between teams. However, existing approaches predominantly ignore the possibility of using simplified models at the initial stage, accessible to business analysts and domain experts, which reduces the effectiveness of their involvement in design.

Between 2020 and 2024, a number of works were published, some of which directly addressed the issue of forming context maps. H. Vural & M. Koyuncu (2021) demonstrated that the correct definition of bounded contexts in a microservice architecture reduces the cognitive complexity of code and accelerates evolutionary changes. However, the authors emphasised the absence of clear rules for transitioning from business processes to contexts. In a broader overview, J. Sangabriel-Alarcón *et al.* (2024) highlighted that practitioners either fragment the domain excessively, leading to unnecessary dependencies, or leave contexts too large, hindering development. C. Zhong *et al.* (2024), based on a statistical cross-section of a large number of companies, showed that early formation of a context map reduces the number of changes in service contracts by 23% in the first year, but researchers note that typical teams lack methodologies that combine the clarity of UML with the strict rules of DDD. In this same context, the research by H. Vural & M. Koyuncu (2021) complemented the results of S. Kapferer & O. Zimmermann (2020), demonstrating that formalised DSL solutions are effective only if the initial models are sufficiently accurate, which is often lacking in the pre-design analysis phase.

The aim of the article was to develop a new method for transforming UML activity diagrams into a context map within DDD, which will simplify the initial modelling of complex software systems and enhance the effectiveness of collaboration between business analysts and developers. It was necessary to analyse existing approaches to building context maps within Domain-Driven Design, paying particular attention to methods based on BPMN, and to evaluate their limitations. The next step involved developing and substantiating a method for systematically transforming UML activity diagrams into a DDD context map, suitable for application in the initial stages of modelling. To verify the effectiveness of the proposed method, it was planned to use a typical domain from the restaurant business to identify key bounded contexts and the relationships between them. Following this, a comparative analysis of the developed method with other approaches was planned to identify its advantages and disadvantages regarding effectiveness and flexibility. Finally, prospects for the further development of the method were identified, including possibilities for automation, the use of CASE tools, and integration with modern technologies.

MATERIALS AND METHODS

The research was based on an approach that involves the systematic transformation of business process activity diagrams into a context map within the framework of DDD. The material for analysis consisted of UML activity diagrams, which were manually created based on the primary business requirements of the

restaurant domain; during modelling, completeness of scenarios, the presence of all key actors, and the display of main control and data flows were taken into account. This domain is characterised by the presence of various participants (clients, cooks, waiters, suppliers, payment systems), as well as extensive interaction between them, which allows for a comprehensive demonstration of the proposed method's effectiveness.

Preliminary identification of key domain events was conducted using the Event Storming method, proposed by S. Gadiyar (2024), which ensured rapid detection of cause-and-effect relationships in the process. The study used a five-stage methodology for transforming a UML activity diagram into a context map within the framework of DDD. A schematic representation of this methodology is provided in Figure 1.

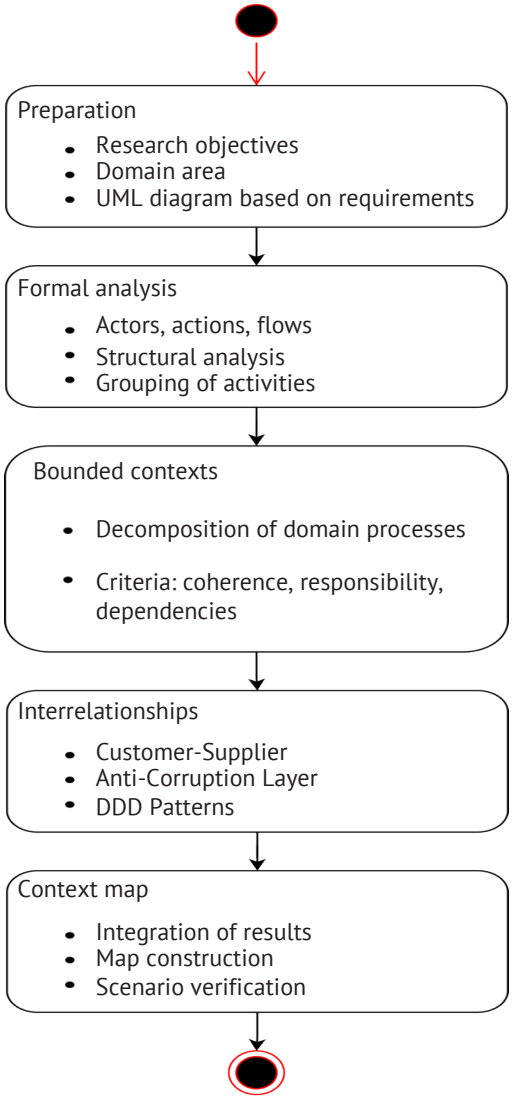

Figure 1. Diagram of the phased transformation of an activity diagram into a context map
Source: developed by the authors

A detailed description of the actions performed at each stage is provided below. During the preparatory stage, the research objectives were formulated, which involved developing a systematic approach to forming a context map for complex business systems, particularly within the restaurant service domain. The restaurant business was chosen as the domain area due to its clear structure, the presence of multiple actors with distinct areas of responsibility, and the typicality of its processes. The UML activity diagram used in the study was created based on an analysis of primary business requirements, formulated through an Event-Storming session (Gadiyar, 2024). When creating the diagram, consideration was given to the completeness of process coverage (customer service, food preparation, inventory management, payment processing), the participation of all key actors (client, waiter, cook, supplier, accountant, payment system), and the presence of control points for data transfer and management. Table 1 provides a generalised description of the main elements used in the study during the transformation of business logic into a context map.

Table 1. Key elements of UML activity diagrams

UML element	Description
Action	The basic unit of the diagram, representing a single step performed by a user or system.
Control Flow	Arrows indicating the sequence of action execution.
Initial/Final Node	Define the start and end of the process respectively.
Fork/Join Node	Used for parallelising or merging execution flows.
Decision Node	A branching point where the process divides depending on a condition.
Object Node	Displays objects (data) that are passed between actions.
Swimlane	Sections that divide actions among process participants, indicating who is responsible for each action.

Source: developed by the authors

At the formal analysis stage of the activity diagram, key actors and the corresponding activities they perform were identified. Grouping was based on the analysis of functional commonality, target outcomes, and sequences of actions. The key criterion here is the presence of a clear business objective that each group of actions realises. The methodological basis for this stage was the principle of transitive unification of actions into service layers, as outlined by O. Özkan *et al.* (2023). This forms the prerequisite for distinguishing bounded contexts – an autonomous part of the system with its own internal logic. At the stage of identifying bounded contexts, the processes of the restaurant business were decomposed based on the identified groups of activities. The criteria for defining contexts were the internal coherence of actions, clear boundaries of responsibility, and the minimisation of external dependencies. For the transitive unification of commands into service layers, the approach by O. Özkan *et al.* (2023) was used, which allowed related functions to be consolidated into a single logical whole.

At the stage of establishing inter-contextual relationships, the directions of interaction between the identified contexts were determined. A comparison of connection types with known interaction patterns (Alfadel *et al.*, 2020; Farshidi *et al.*, 2020) was applied. The determination of critical synchronisation points was carried out considering the queue-time/flow-time ratio (Albuquerque *et al.*, 2020), and patterns (Valdivia *et al.*, 2020) were used for formal typification. In the final stage, a context map was constructed, where each of the bounded contexts is represented as a separate component, and the interrelationships between them as labelled lines indicating the type of dependency. To verify the integrity of the model, the Specification by Example methodology was applied, which allowed for ensuring the correctness and consistency of the interaction logic between system elements (Irshad *et al.*, 2022).

RESULTS AND DISCUSSION

In the process of developing the method, the main components of activity diagrams were analysed, which are a key means of visualising business processes within UML (Ambler, 2005; Ozkaya & Erata, 2020). This type of diagram allows for modelling sequences of actions, parallel flows, branching conditions, and also for defining the responsibilities of process participants using swimlanes. An activity diagram is intuitive and has a lower barrier to entry compared to the more complex BPMN notation, making it convenient for use in designing business processes. DDD is an approach to developing software systems that focuses on a deep understanding of the domain and modelling business processes according to its real needs (Evans, 2003; Vernon, 2013). DDD involves close interaction between developers and domain experts to create a model that accurately reflects the entities and processes existing within the domain. The main tool is a Ubiquitous Language, which both parties use to describe and interact with the domain.

A Context Map is an important DDD tool that allows for visualising the boundaries and interactions between different parts of the system, or so-called bounded contexts. Each bounded context represents a separate, clearly defined business area within which specific terms, models, and rules have a fixed meaning and are interpreted independently of other parts of the system. This helps to avoid ambiguities and conflicts in the interpretation of concepts, especially in large projects with multiple teams. Within a single bounded context, its own model is formed, which should not be directly used outside of this context without special alignment mechanisms (e.g., via anti-corruption layers or published languages). As an example in the restaurant business, the "Customer Service" context might contain the term "order", which means a list of dishes chosen by a visitor and is accompanied by a waiter's actions. At the same time, in the "Kitchen" context, the same term "order" would mean an internal technical task for the cook, which includes preparation instructions. Despite having the same name, these models are independent and correspond to different business needs, thus they should exist within separate contexts. Bounded contexts in DDD are classified into three main types depending on their role in the system (Kapferer & Zimmermann, 2020).

The Core Domain is the most important part of the system, reflecting key business capabilities and strategic

aspects. This is a critical area for the product's success, and it is here that the largest investment of resources in modelling has the greatest value. Unique competitive advantages are formed in this domain, so it requires the deepest elaboration and continuous involvement of domain experts. The Supporting Domain encompasses auxiliary contexts that support the core functions but are not strategically important themselves. They perform less critical roles but are necessary for the effective functioning of the system. Typically, these are technical or administrative components that facilitate the operation of the core domain (e.g., logistics or accounting). The Generic Domain includes standard contexts that use typical, generally accepted models, not specific to a particular business. They can be reused in different projects without significant adaptation or deep modelling. Examples might include authorisation modules, user registration, or message processing.

Contexts do not exist in isolation; various types of interrelationships can arise between them. It is important to clearly define these interrelationships to ensure a coherent and flexible architecture. Incorrectly designed interactions between contexts can lead to excessive dependency, duplication of functionality, or difficulties with scaling. Clear definition of boundaries and integration mechanisms help to avoid terminology conflicts, preserve module autonomy, and ensure system stability during changes. Research by A. Hlybovets & I. Paprotskyi (2024) demonstrated that correctly implemented fault tolerance at the microservice level significantly reduces the risk of such failures. The main types of interactions are shown in Table 2.

Table 2. Types of interactions between bounded contexts

Type of interaction	Description	Features of interaction
Customer-Supplier	One context supplies data or services, the other consumes them.	The supplier is obliged to maintain the service; the consumer depends on the reliability and stability of the Application Programming Interface (API).
Conformist	The consumer adopts the supplier's model without changes.	Reduces integration complexity, but limits flexibility for future changes.
Partnership	Contexts closely collaborate to achieve a common goal.	Both sides are mutually dependent; high coordination of actions is required.
Open Host Service	A context publishes a service, accessible to several other contexts.	Provides shared access to functionality without tight integration.
Published Language	Interaction via a standardised language or protocol, agreed upon between contexts.	Reduces dependency, facilitates scaling and interaction with multiple systems.
Anti-Corruption Layer (ACL)	The downstream context is isolated from the upstream context using an adaptation layer.	Ensures translation between models; protects the internal model from external influences.

Source: developed by the authors

The main components of a context map in DDD (Vernon, 2013) play a crucial role in forming a clear system architecture and ensuring consistency between its parts. A context map serves as a visual tool demonstrating the structure of interactions between autonomous modules – bounded contexts – and defines their relationships, areas of responsibility, and shared components. The first and central element is the bounded context, which has already been described above. The importance of this component lies in its ability to avoid semantic conflicts and ensure the autonomy of teams working on different modules. Dividing into contexts helps prevent confusion and mixing of meanings. The next critical element of the map is the Relationships between bounded contexts. These relationships describe exactly how individual parts of the system interact, who is the supplier, and who is the consumer of data or services.

Types of interactions, such as Customer-Supplier, Conformist, Partnership, Open Host Service, or Published Language, define the level of dependency, coordination, and shared responsibility between contexts. Clear modelling of such interrelationships helps avoid excessive integration dependencies and ensures system flexibility when changes occur in specific modules. The third important component is the Shared Kernel – a set of models, objects, or functions that are used by several bounded contexts simultaneously. Its main purpose is to ensure consistency and reduce duplication of business logic in cases where complete isolation of contexts is impractical or too costly. A shared kernel requires careful management, as changes within it can affect all related contexts. However, in cases where contexts have closely related logic, a shared kernel becomes an effective mechanism for code reuse and maintaining term consistency.

A context map helps divide complex systems into independent areas, which reduces the complexity of interactions between different components and provides clear boundaries of responsibility. This is important for constantly evolving systems, as it allows changes within one context to be easily managed without disrupting the operation of others. Creating a context map within

DDD involves the sequential execution of a series of analytical stages, each of which plays an important role in forming a correct architectural model.

At the initial stage, bounded contexts were identified and grouped into corresponding subdomains. This involves detailed identification of the actors represented in the UML activity diagram, as well as the activities these actors perform. Identification was based on an analysis of swimlanes, which allow for the visual distribution of responsibilities between users and system components. Particular attention was paid to establishing the sequence of actions that form logical process blocks. In the next step, an analysis of the goals and expected outcomes of each process participant was carried out. This approach allows for understanding the motivation of each actor, as well as determining which results are critical from a business value perspective. This analysis serves as the basis for identifying bottlenecks and priority elements of the system. It is in this context that the identification of critical process points and potentials for its optimisation becomes possible, contributing to the qualitative separation of logic by domain boundaries.

Next, an investigation of interactions between actors was conducted. The primary focus was on identifying the roles of supplier and consumer in each specific segment of the process, which allows for the classification of dependencies between participants. Data flows and their directions are analysed to determine the directions of integration and coordination between contexts. This is critical for understanding precisely how information is transferred between system parts, as well as for identifying areas of potential duplication or conflicts. Following this, semantic grouping of activities within logical business functions was carried out. Based on the resulting groups of functionally related activities, bounded contexts are identified. It is important not only to ensure the internal coherence of each context but also to minimise its dependencies on other parts of the system. This provides greater autonomy, simplifies testing, accelerates independent deployment, and reduces the risks of cascading changes. At this stage, bounded contexts were also integrated into broader subdomains. Such consolidation is based on the relatedness of functions and conformity to the overall business structure, and also allows for classifying contexts as core, supporting, or generic – in accordance with DDD principles.

Further analysis involved identifying internal relationships between bounded contexts. For this purpose, data exchange paths between previously defined groups of actions were analysed, their functional interdependence and frequency of interactions were evaluated. Based on this information, specific communication channels were identified, their roles (supplier/consumer) were formalised, and the direction of information flows is established. The concepts of upstream and downstream contexts are applied to model who initiates the interaction and who receives it, which is crucial for maintaining system stability in changing conditions. The final analytical step is to assess the nature of the interrelationships between contexts. Here, it is appropriate to apply the DDD typology, which includes models such as Customer-Supplier, Conformist, Partnership, Anti-Corruption Layer, Open Host Service, and Published Language. Each type of interaction implies a different level of integration, isolation, and coordination, and the selection of the appropriate model depends on the requirements for flexibility, independence, and scalability. The analysis of these interrelationships allows for the justified formation of a context map, which not only visualises the system's architecture but also serves as a means of strategic management for its development.

Based on the conducted analysis, a detailed list of bounded contexts, subdomains, groups of related actions was obtained, and the interrelationships between them were established. As a result of the preceding stages, context map within DDD was created. This map serves as the basis for visualising and understanding the interaction between different parts of the system. Furthermore, an example of applying the method for transforming an activity diagram into a domain-driven context map using the restaurant business as a case study was provided. The activity diagram for the restaurant business is presented in Figure 2.

The first step in building a contextual map of the restaurant business was to identify key actors and related activities, which were recorded in the diagram (Fig. 2). Analysis of the participants in the process made it possible to determine which roles (both human and technical) are involved in the business process and what functional actions they perform. In particular, roles such as customer, waiter, cook, supplier, payment processing system, and accountant were covered. For each of them, a typical set of activities was identified that represent their functions within the service process: from the customer's arrival to the confirmation of payment by the accountant.

Further analysis was aimed at uncovering the motivations and expectations of each actor. At this stage, the goals of each role were defined – for example, client satisfaction, ensuring the quality of prepared dishes, accurate financial reporting, etc. – as well as the expected outcomes of participation in the business process. This allowed for identifying the strategic significance of individual functions and understanding which aspects of the process hold the highest business value and influence the overall user experience. Particular attention was paid to analysing the interactions between actors. It was established that supplier-consumer relationships exist between the roles, defining functional dependencies within the system. For instance, the client consumes services provided by the waiter, and the latter, in turn, relies on the results of the cook's work. Such chain dependencies allow for clear outlining of information flows and responsibilities between parts of the system. Data flow analysis enabled the identification of control directions and the

boundaries within which services or requests are transferred. The next logical step was the conceptual grouping of activities within higher logical units – business functions. Activities that were related by purpose or outcome were combined into four main business functions: customer service, food preparation, inventory management, and payment processing. This grouping was critical from the perspective of forming bounded contexts, as it allowed for the identification of autonomous functional areas with a high degree of internal coherence.

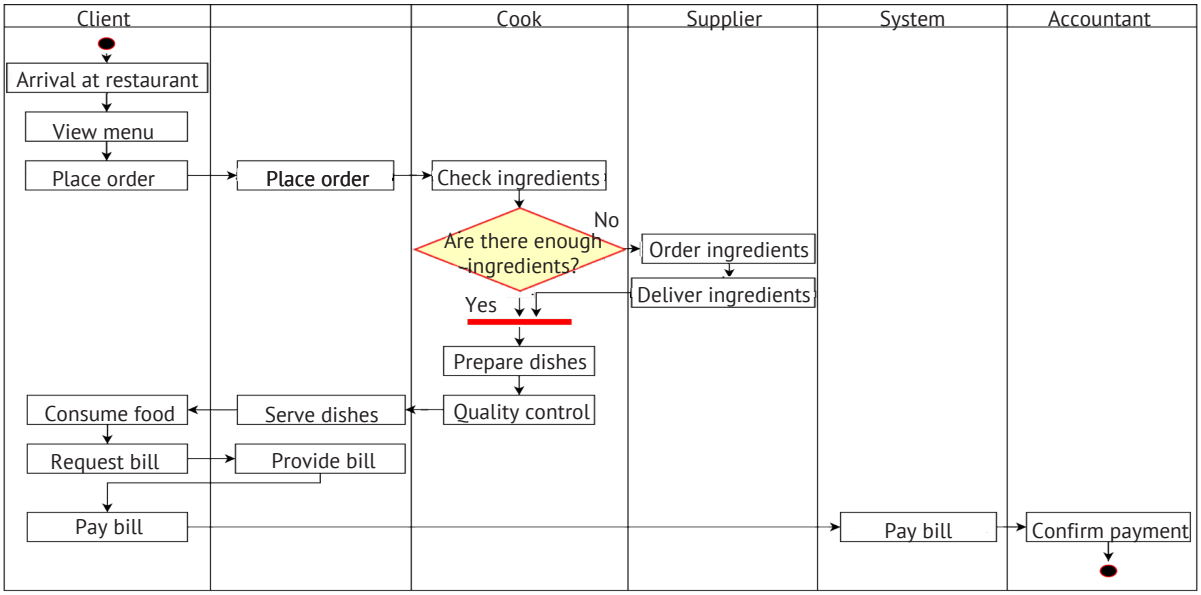

Figure 2. Activity diagram of the restaurant business

Source: developed by the authors

Based on the formed functional groups, the bounded contexts were defined: "Customer Service", "Kitchen", "Inventory Management", and "Payment". Each of these demonstrates a clearly expressed autonomy, which allows for minimising coupling between contexts and reducing potential risks during changes. Further analysis allowed these contexts to be classified by subdomain types according to DDD approaches. The "Customer Service" and "Kitchen" contexts were assigned to the core subdomains, as they represent key business value. In contrast, "Inventory Management" and "Payment" were classified as supporting subdomains, because, although important for system functioning, they do not carry strategic uniqueness.

In the next stage, inter-contextual relationships were modelled based on previously identified relationships between actors. The "Customer Service" context interacts with the "Kitchen" context using a Customer-Supplier model, where "Kitchen" supplies prepared dishes, and "Customer Service" consumes them. A similar dependency is observed between "Kitchen" and "Inventory Management", although an additional mechanism – an anti-corruption layer – was applied here, which allowed for isolating the model of one context from changes in another. Interaction between "Customer Service" and "Payment" was also modelled as Customer-Supplier with elements of an ACL, as the payment processing process is critical but should not affect the internal logic of the main service. The final stage was the construction of the context map, which visualises the derived bounded contexts, their interrelationships, integration types, and directions of dependencies (Fig. 3).

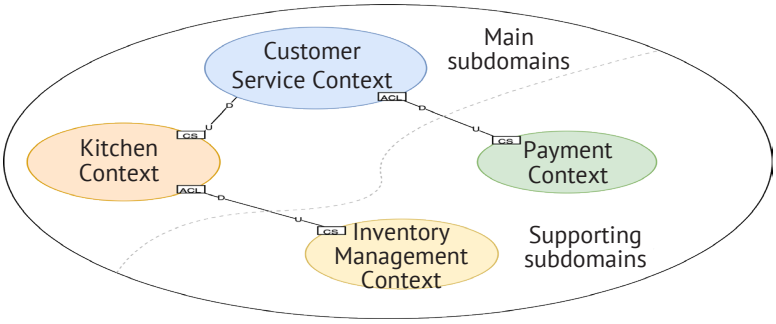

Figure 3. Context map of the restaurant business

Source: developed by the authors

It is important to involve domain experts to verify and refine the created context map. Their experience and knowledge of business specifics will help identify possible inaccuracies or gaps in the model, and ensure that all nuances of business processes are taken into account. This will contribute to the context map accurately reflecting the real needs and structure of the business. It should be noted that the context map may change during further analysis and development. In-depth study of the domain, changes in business requirements, or newly identified details may require a review of context boundaries and their interrelationships. The proposed method for transforming activity diagrams into a DDD context map is convenient for modelling business systems. The use of UML activity diagrams simplifies the process, as UML is a widely known and understood tool, facilitating collaboration between business analysts and developers.

Unlike the BPM2DDD method, which is based on detailed BPMN models and can be more complex for organisations without established BPMN usage practices, the approach proposed in this study is more flexible and accessible (Da Silva *et al.*, 2022). It allows for precisely defining bounded contexts and accounting for the nuances of business processes without excessive formalisation. Furthermore, this method emphasises the importance of involving domain experts and readiness to adapt the context map during the analysis process, which increases the accuracy and relevance of the model. Compared to BPM2DDD, the proposed approach ensures faster and simpler transformation of business processes into technical models, making it an attractive choice for many organisations.

The obtained results regarding the systematic transformation of UML activity diagrams into a context map for DDD should be considered in comparison with the research of other authors who also studied the problem of forming bounded context boundaries and integrating business processes into DDD. A comparative analysis made it possible to determine the extent to which the chosen approach is consistent with previous works and what the main differences are. First and foremost, the findings should be compared with the results of the study by C.E. Da Silva *et al.* (2022), which proposed a formalised transformation of BPMN models into domains within the BPM2DDD method. The author noted high accuracy in domain identification but emphasised that the complexity of BPMN complicates the involvement of specialists without in-depth expertise in process notation. In the presented approach, a similar goal – the formation of bounded contexts – was achieved, but by using more accessible UML activity diagrams. Compared to the BPM2DDD method, the presented solution benefits from a lower barrier to entry for teams, but it may lose out on the level of detail in formal process descriptions.

The views in the study by F. Gonzalez-Lopez *et al.* (2022), which proposed domain-oriented design of business process architecture, also turned out to be similar in their emphasis on business logic. The importance of a clear description of the domain before forming the system architecture was highlighted. The proposed approach also underscored the decisive role of initial business process analysis through UML models, where grouping activities by logical attributes contributed to the isolation of contexts. The difference lies in the emphasis on simplifying the initial modelling steps, whereas the study by F. Gonzalez-Lopez *et al.* (2022) aimed at deep integration of domain concepts in large-scale corporate systems.

Compared to the work of S. Kapferer & O. Zimmermann (2020), where the main focus was on creating specialised languages and tools for strategic contextual mapping, the obtained results are less "technologically specialised." The authors demonstrated the benefits of DSL for describing interactions between contexts, which reduced the risk of incorrect terminology interpretation. In this study, priority was given to the simplicity of UML diagrams, which did not require additional description languages. Such an approach proved effective for application in teams with varying levels of technical expertise, although it potentially yields in terms of formalisation level, as discussed by S. Kapferer & O. Zimmermann (2020).

Important parallels can be seen with the developments of C. Zhong (2024), which conducted a systematic analysis of DDD application in microservices. This study focused on how clear context boundaries help avoid duplication of functions and incorrect distribution of data between services. The obtained results confirmed the validity of these conclusions, as the formation of a context map at an early stage safeguarded against unnecessary fragmentation. However, that study was based on the analysis of existing real-world systems, whereas this work focused on a method for transforming activity diagrams into a contextual model.

The studies by H. Vural & M. Koyuncu (2021) were aimed at investigating the impact of DDD on the modularity of microservices, demonstrating that correct identification of bounded contexts defines the boundaries of service responsibilities. The obtained conclusions align with the authors' position on the crucial role of proper grouping of business processes in preventing duplication and terminology conflicts. At the same time, the works of these authors provided quantitative indicators evaluating system modularity, whereas the proposed methodology placed greater emphasis on qualitative analysis (logical grouping of actions and decomposition of processes).

It is also interesting to compare the approach to domain decomposition highlighted in the work of J. Sangabriel-Alarcón *et al.* (2024), where it was noted that strategic contextual mapping and precise definition of bounded contexts contribute to better system scalability and reduced technical debt. The obtained results confirmed this thesis, as transparent context boundaries facilitated working with domains that change over time. At the same time, the risks of oversaturating the

system with overly granular contexts, which can complicate coordination between teams, were considered. The presented methodology included a criterion for minimising inter-contextual dependencies, but the method prioritised preliminary business analysis, whereas the scholar's work favoured a multi-step iterative verification of context boundaries based on real data.

A comparison with the results of previous studies demonstrated that the field of integrating DDD with business process models is evolving in several directions. Some researchers focused on more formalised transformation mechanisms, using, for example, BPMN, specialised DSLs, and algorithmic modularity criteria. Others, however, emphasised a simplified initial analysis that fosters better mutual understanding among teams with varying levels of technical expertise. Ultimately, it is the balance between formal depth and the accessibility of the proposed methods that determines success in domain decomposition and the formation of a flexible, scalable architecture within DDD.

CONCLUSIONS

Throughout the study, a phased methodology was developed that effectively combines business analysis with technical DDD practices, allowing for the clear and prompt identification and description of key operations within the restaurant business domain, all within their respective bounded contexts. In the initial stage, business requirements and primary process models were thoroughly analysed, and key participants and their main activities were identified. The subsequent stage involved a formal analysis of UML activity diagrams, which included systematising actions based on functional commonality, business objectives, and interdependencies. This enabled the separation of several autonomous bounded contexts and the formation of a clear understanding of the nature of interactions between them, serving as the foundation for creating a coherent context map.

The research results demonstrated the practical effectiveness of the proposed method and confirmed that combining business analysis with UML diagrams significantly simplifies and accelerates the modelling of complex systems at early design stages. The use of intuitive UML diagram tools ensured the accessibility and transparency of the process for both business analysts and software developers, fostering improved communication and mutual understanding between multidisciplinary teams. The proposed method not only allowed for clear definition of responsibilities between system parts but also significantly reduced the risks of semantic conflicts and errors arising from ambiguous interpretations of business requirements. Furthermore, this approach can contribute to building flexible and scalable architectural solutions that can adapt to changes in business processes without substantial costs for system rework and adaptation. However, it is important to note that the practical implementation of the method in this study was conducted using the example of a single domain – the restaurant business – which requires additional verification of the method's universality and scalability in other industries and business contexts. For further refinement of the method, it is recommended to conduct similar studies in other domains, such as financial services, healthcare, logistics, and so forth.

In future research, a promising direction is the automation of the process for transforming UML activity diagrams into context maps using modern CASE tools and software platforms. This can significantly enhance the efficiency and accuracy of initial modelling, reduce time expenditure, and facilitate the integration of this method into existing software development processes. The integration of the method with cutting-edge technologies, such as machine learning and artificial intelligence, holds promise for further enhancing business process analysis and enabling the automatic identification of key contexts and interrelationships.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Albuquerque, F., Torres, A.S., & Berssaneti, F.T. (2020). Lean product development and agile project management in the construction industry. *Revista de Gestão*, 27(2), 135-151. doi: 10.1108/REGE-01-2019-0021.

[2] Alfadel, M., Aljasser, K., & Alshayeb, M. (2020). Empirical study of the relationship between design patterns and code smells. *PLoS ONE*, 15(4), article number e0231731. doi: 10.1371/journal.pone.0231731.

[3] Ambler, S.W. (2005). UML activity diagrams. In *The elements of UML™ 2.0 style* (pp. 113-131). Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511817533.011.

[4] Da Silva, C.E., Gomes, E.L., & Basu, S.S. (2022). BPM2DDD: A systematic process for identifying domains from business processes models. *Software*, 1(4), 417-449. doi: 10.3390/software1040018.

[5] Evans, E. (2003). *Domain-driven design: Tackling complexity in the heart of software*. Boston: Addison-Wesley.

[6] Farshidi, S., Jansen, S., & van der Werf, J.M. (2020). Capturing software architecture knowledge for pattern-driven design. *Journal of Systems and Software*, 169, article number 110714. doi: 10.1016/j.jss.2020.110714.

[7] Gadiyar, S. (2024). Leveraging event storming for enhanced product development in FinTech: A collaborative approach to simplifying complex business processes. *Journal of Mathematical & Computer Applications*, 3(5), article number E146. doi: 10.47363/JMCA/2024(3)E146.

[8] Gonzalez-Lopez, F., Bustos, G., Muñoz-Gama, J., & Sepúlveda, M. (2022). Domain model based design of business process architectures. *Applied Sciences*, 12(5), article number 2563. doi: 10.3390/app12052563.
[9] Hlybovets, A., & Paprotskyi, I. (2024). Increasing the fault tolerance in microservice architecture. *Cybernetics and Systems Analysis*, 60(3), 480-488. doi: 10.1007/s10559-024-00689-0.
[10] Irshad, M., Börstler, J., & Petersen, K. (2022). Supporting refactoring of BDD specifications – an empirical study. *Information and Software Technology*, 141, article number 106717. doi: 10.1016/j.infsof.2021.106717.
[11] Kapferer, S., & Zimmermann, O. (2020). Domain-specific language and tools for strategic domain-driven design, context mapping and bounded context modeling. In *Proceedings of the 8th international conference on modeling and software development* (pp. 299-306). Valetta: SciTePress. doi: 10.5220/0008910502990306.
[12] Khoshnevis, S. (2023). A search-based identification of variable microservices for enterprise SaaS. *Frontiers of Computer Science*, 17(3), article number 173208. doi: 10.1007/s11704-022-1390-4.
[13] Özkan, O., Babur, Ö., & van den Brand, M. (2023). Refactoring with domain-driven design in an industrial context. *Empirical Software Engineering*, 28, article number 94. doi: 10.1007/s10664-023-10310-1.
[14] Ozkaya, M., & Erata, F. (2020). A survey on the practical use of UML for different software architecture viewpoints. *Information and Software Technology*, 121, article number 106275. doi: 10.1016/j.infsof.2020.106275.
[15] Sangabriel-Alarcón, J., Ocharán-Hernández, J.O., Limón, X., & Cortés-Verdín, M.K. (2024). Domain-driven design in microservices-based systems development: A systematic literature review and thematic analysis. *Programming and Computer Software*, 50(8), 742-770. doi: 10.1134/S0361768824700749.
[16] Valdivia, J.A., Lora-González, A., Limón, X., Cortés-Verdin, M.K., & Ocharán-Hernández, J.O. (2020). Patterns related to microservice architecture: A multivocal literature review. *Programming and Computer Software*, 46(8), 594-608. doi:10.1134/S0361768820080253.
[17] Vernon, V. (2013). *Implementing domain-driven design*. Boston: Addison-Wesley Professional.
[18] Vural, H., & Koyuncu, M. (2021). Does domain-driven design lead to finding the optimal modularity of a microservice? *IEEE Access*, 9, 32721-32733. doi: 10.1109/ACCESS.2021.3060895.
[19] Waseem, M., Liang, P., Shahin, M., Di Salle, A., & Márquez, G. (2021). Design, monitoring, and testing of microservices systems: The practitioners' perspective. *Journal of Systems and Software*, 182, article number 111061. doi: 10.1016/j.jss.2021.111061.
[20] Zhong, C., Li, S., Huang, H., Liu, X., Chen, Z., Zhang, Y., & Zhang, H. (2024). Domain-driven design for microservices: An evidence-based investigation. *IEEE Transactions on Software Engineering*, 50(6), 1425-1449. doi: 10.1109/TSE.2024.3385835.

Метод структурованого перетворення діаграми активності бізнес-процесів у карту контексту предметно-орієнтованого проєктування

Сергій Московко

Аспірант
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0009-0001-4376-0187

Роман Квєтний

Доктор технічних наук, професор
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0000-0002-9192-9258

Анотація. Актуальність дослідження полягає в потребі впровадження уніфікованих методів для структурованого формування контекстних карт, здатних надійно та точно відображати бізнес-процеси під час проєктування предметно-орієнтованих інформаційних систем. Метою роботи було розроблення та обґрунтування систематичного методу перетворення діаграм активності бізнес-процесів у контекстні карти, що оптимізує процес моделювання та підвищує точність відображення взаємодій у домені. Для досягнення цієї мети було застосовано методи формального аналізу діаграм активності, структурованого виявлення обмежених контекстів, групування взаємопов'язаних дій і встановлення типів взаємодій між ними. Основні результати дослідження полягають у створенні поетапного методу, який дозволяє ідентифікувати ключові елементи бізнес-процесів, об'єднувати їх у логічно узгоджені контексти та визначати міжконтекстні взаємозв'язки без використання складних нотацій. Запропонований підхід дає змогу виокремлювати обмежені контексти на основі аналізу діаграм активності, виявляти динамічні залежності між діями різних учасників та формувати узгоджену контекстну карту. Ефективність методу продемонстрована на прикладі ресторанного бізнесу, де за його допомогою було чітко ідентифіковано взаємодії між процесами обслуговування клієнтів, кухнею, управлінням запасами та обробкою платежів. Це сприяло прозорому розподілу відповідальності, покращило керування складністю системи, прискорило проєктування та підвищило узгодженість між технічною реалізацією та бізнес-вимогами. Запропонований підхід також покращив взаєморозуміння між розробниками та експертами предметної області, оскільки забезпечує чіткі межі відповідальності та більш адаптивну архітектуру системи. Крім того, результати підтвердили можливість масштабування методу на інші галузі, що свідчить про його універсальність та широку застосовність. Практична цінність роботи полягає у можливості масштабування методу для застосування в різних доменах, що робить його корисним інструментом для бізнес-аналітиків і архітекторів програмного забезпечення під час проєктування складних систем та підвищення ефективності прийняття архітектурних рішень

Ключові слова: інформаційні системи; аналіз бізнес-процесів; уніфікована мова моделювання; трансформація моделей; визначення обмежених контекстів

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 30 No. 2. 2025

UDC 004.056.5:351.862
DOI: 10.62660/bcstu/2.2025.44

Integrated method for assessing expert competence (IMAEC) in state information security

Dmytro Prokopovych-Tkachenko*
PhD in Technical Sciences, Associate Professor
University of Customs and Finance
49000, 2/4 Volodymyra Vernadskoho Str., Dnipro, Ukraine
https://orcid.org/0000-0002-6590-3898

Liudmyla Rybalchenko
PhD in Economic Sciences, Associate Professor
University of Customs and Finance
49000, 2/4 Volodymyra Vernadskoho Str., Dnipro, Ukraine
https://orcid.org/0000-0003-0413-8296

Volodymyr Zvieriev
PhD in Technical Sciences, Associate Professor
State University of Trade and Economics
02156, 19 Kyoto Str., Kyiv, Ukraine
https://orcid.org/0000-0002-0907-0705

Valerii Bushkov
Postgraduate Student
State University of Trade and Economics
02156, 19 Kyoto Str., Kyiv, Ukraine
https://orcid.org/0009-0005-5097-2689

Borys Khrushkov
Postgraduate Student
University of Customs and Finance
49000, 2/4 Volodymyra Vernadskoho Str., Dnipro, Ukraine
https://orcid.org/0009-0002-3978-5012

Abstract. The relevance of this study is grounded in the need to improve methods for evaluating expert competence in the field of state information security. The increasing number of threats in cyberspace and the growing demands on specialist qualifications necessitate innovative decision-making approaches that can account for subjective factors and data uncertainty. This research aimed to develop and test an integrated approach combining the Analytic Hierarchy Process (AHP) with fuzzy logic for the assessment of candidates for expert positions in information security. The AHP enabled the problem to be structured as a hierarchical model encompassing the goal, criteria (experience, certifications, communication skills), and alternatives (candidates). The AHP methodology involved pairwise comparison of criteria, calculation of weights, and consistency checks of the matrices. Fuzzy logic complemented AHP by enabling the processing of imprecise data through fuzzification, the application of “If-Then" rules, and defuzzification. Practical validation of the method was carried out using the example of

Article's History: Received: 08.01.2025; Revised: 24.04.2025; Accepted: 16.06.2025.

Suggested Citation:
Prokopovych-Tkachenko, D., Rybalchenko, L., Zvieriev, V., Bushkov, V., & Khrushkov, B. (2025). Integrated method for assessing expert competence (IMAEC) in state information security. *Bulletin of Cherkasy State Technological University*, 30(2), 44-52. doi: 10.62660/bcstu/2.2025.44.

*Corresponding author

candidate evaluation based on the specified criteria. The research findings demonstrated that the proposed approach enabled the integration of both precise and imprecise elements of assessment, thereby improving the accuracy and justification of decisions. The Monte Carlo method was employed to verify the model's reliability, ensuring result stability through repeated simulations of data variations. The results confirmed the method's high adaptability to handling uncertain data. The practical value of the study lies in the application of the integrated approach to enhance the efficiency of personnel selection – an essential component in safeguarding information security within state institutions. The proposed method may also be applied to other multi-criteria decision-making problems under conditions of uncertainty

Keywords: analytic hierarchy process; fuzzy logic; decision-making; criteria; fuzzification; defuzzification; Monte Carlo

INTRODUCTION

The assessment of expert competence in the field of information security is one of the critical tasks for ensuring the effective functioning of state institutions. The increasing number of cyberattacks and the development of complex technologies demand the involvement of qualified specialists who can rapidly adapt to emerging challenges. According to researcher D. Landoll (2021), the spread of hybrid threats – now a frequent occurrence in contemporary life – is influenced by the global digital environment, which necessitates reliable methods of protection. Various approaches exist for assessing risks in the field of information security. For example, the study by V.V. Yemelyanov *et al*. (2021) is based on traditional criteria such as formal qualifications, professional experience, and certification outcomes.

However, K. Latysh *et al*. (2022) argued that substantial errors may occur in the examination of information technology-related issues due to the use of outdated methods, obsolete logical protection tools for technical media, and flawed approaches to analysing past mistakes. In the academic study, I. Hasiuk & O. Ivanii (2024) observed that traditional criteria fail to consider adaptability to frequently changing threats and that their effectiveness in crisis situations is limited, as not all cyber incidents can be taken into account. The methodology proposed by K. Smith *et al*. (2023) for multi-criteria decisionmaking – namely, the AHP – is considered the most effective. It enables the structuring of tasks by dividing them into levels of objectives, criteria, and alternatives, and allows for pairwise comparison of criteria, determination of their relative weights, and calculation of consistency ratios.

Most existing methods for assessing experts are based on traditional approaches that consider only quantitative indicators, overlooking subjective factors such as professional intuition, adaptability, and the ability to operate under uncertainty. This results in limitations when forming an objective picture of a candidate's competence. The integration of new decision-making methods that account for both precise and imprecise aspects of evaluation is therefore of particular importance. Traditional approaches lack the flexibility required to handle incomplete or contradictory data, which is typical in the field of information security. Consequently, there is a need to implement combined methods that ensure objectivity, reliability, and a high degree of accuracy in expert assessment. The AHP is among the most effective methods for multi-criteria decision-making. It enables pairwise comparison of criteria, determination of their relative weights, and calculation of consistency ratios (Smith *et al.*, 2023).

According to V.O. Burenko (2024), despite its numerous advantages, the classical AHP is not always effective when dealing with imprecise and subjective data, which is typical in the assessment of expert competence in the field of information security. To address this issue, researchers M. Averkyna & M. Prystupa (2019) considered the integration of AHP with fuzzy logic methods to be a promising solution. These methods enable the processing of uncertain and vague evaluations through the fuzzification of input data, the use of "If-Then" fuzzy rules, and defuzzification. Scholars N.I. Domantsevych & H.S. Shestopal (2023) noted that such an approach allows for the integration of both quantitative and qualitative aspects of expert evaluation, thereby enhancing the justification and accuracy of decisionmaking. Researcher I.M. Debela (2021) applied a Bayesian approach for assessing alternative decisions involving continuous random variables, which takes into account the properties of a holistic system as influenced by the surrounding environment. According to N. Fil & O. Kudyrko (2023), the application of a multi-criteria method for evaluating website quality enabled the development of a custom methodology for determining the significance coefficients of website quality criteria. The application of mathematical and statistical methods in modern business processes by researcher V. Havryliuk (2021) produced positive results in credit scoring models used to assess solvency. Scoring models are, in fact, among the tools used to evaluate leasing risks in banking and financial operations.

According to research by L. Petyk & B. Kravchenko (2024), the Monte Carlo method is additionally used to assess the stability of proposed approaches. This technique allows for the simulation of random data variations and the analysis of result reliability. It

helps to minimise the influence of subjective factors and ensures the stability of final evaluations, even under conditions of significant variability in input parameters (IEC 31010:2019, 2019). Researchers A.F. Ayeni *et al*. (2020) described a methodology that combines AHP with fuzzy comprehensive evaluation (FCE) for risk analysis in the field of information security. Although their approach demonstrates high accuracy, it is less adaptable for competence assessment tasks due to the complexity of interpreting the results. In addition, the Monte Carlo method serves as an effective tool for testing the robustness of models that operate with random data variations. In this study, the author employed a multi-layered structure and introduced a rating matrix to describe the relationships between judgement factors. Based on calculations of the impact of these factors on security, a new quantitative risk assessment method has been developed. A method for analysing the results of fuzzy judgements was proposed. This method shifts the perspective on information security assessment from a "black box" approach and plays a role in verifying and adjusting outcomes derived from traditional models.

The use of a multi-layered structure and the introduction of a rating matrix by L. Wang *et al*. (2010) to describe the relationships between influencing factors enabled the development of a new quantitative method for risk assessment based on the calculated impact of these factors on security. The proposed method for analysing the results of fuzzy sets in the context of information security played a significant role in verifying and refining the outcomes generated by traditional models. In the study by V. Lukichov *et al*. (2023), a methodology, that accounts for the specific requirements of cybersecurity and employs multi-factor risk analysis, was developed, implemented, and tested. This approach, based on an integrated method, proved to be highly effective. Thus, the implementation of an integrated method combining the AHP, fuzzy logic, and the Monte Carlo method enables the development of an effective system for assessing expert competence in the field of information security. This study aimed to design, implement, and test such a methodology, taking into account the specific requirements of cybersecurity and the need for multi-factor risk analysis.

MATERIALS AND METHODS

The development of the Integrated method for assessing expert competence (IMAEC) was based on the use of two core algorithms: the AHP and fuzzy logic. The AHP structured the problem as a hierarchical model consisting of three levels: the goal (selecting the most suitable expert), the criteria (experience, certifications, and communication skills), and the alternatives (candidates). The process began with the construction of a pairwise comparison matrix to assess the relative importance of each criterion and alternative. Based on these data, eigenvectors and criterion weights were calculated, and the consistency of the matrix was evaluated using the consistency index and consistency ratio. If the level of consistency exceeded the acceptable threshold, the matrix was reviewed and adjusted accordingly.

The next stage involved the use of fuzzy logic, which allowed for the inclusion of subjective expert assessments in cases where data were imprecise or incomplete. Input data were fuzzified using membership functions – for instance, the "experience" criterion was divided into sets such as "low", "medium", and "high". Fuzzy "If-Then" rules enabled the aggregation of data, while defuzzification converted the results into precise values. To assess the reliability of the model, the Monte Carlo method was applied to simulate random variations in the input data. Hundreds of iterations with varying parameters were conducted, allowing for an evaluation of the stability of the results. The model demonstrated adaptability to change and maintained high accuracy, even in cases of significant variability in the input data.

RESULTS AND DISCUSSION

AHP. Step 1. Structuring the problem as a hierarchy. The hierarchy consists of three levels: Level 1: Goal (e.g. selection of the most suitable expert in information security). Level 2: Criteria (e.g. experience, certifications, communication skills) Level 3: Alternatives (e.g. Candidate 1, Candidate 2, Candidate 3).

Step 2. Constructing the pairwise comparison matrix. For each criterion, a pairwise comparison matrix is constructed. For example, for three criteria, C_1, C_2, C_3:

$$A = \begin{pmatrix} 1 & a_{12} & a_{13} \\ \frac{1}{a_{12}} & 1 & a_{23} \\ \frac{1}{a_{13}} & \frac{1}{a_{23}} & 1 \end{pmatrix}. \quad (1)$$

Step 3. Calculating eigenvectors and eigenvalues. The eigenvector w of matrix A is determined, corresponding to the largest eigenvalue λ_{max}:

$$A \cdot w = \lambda_{max} \cdot w. \quad (2)$$

Step 4. Calculating the weights of the criteria. The eigenvector w is used to obtain the weights of the criteria:

$$w_i = \frac{w_i}{\sum_{i=1}^{n} w_i}. \quad (3)$$

Step 5. Consistency check. The consistency index (CI) and consistency ratio (CR) are calculated:

$$CI = \frac{\lambda_{\max} - n}{n-1}; \quad CR = \frac{CI}{RI}. \quad (4)$$

where RI is the random consistency index. If $CR < 0.1$, the matrix is considered consistent.

Fuzzy logic. Step 1. Fuzzification. This involves converting crisp input data into fuzzy sets using

membership functions. For example, the criterion \experience\ may be represented by three fuzzy sets: \low, \medium, and \high.

$$\mu_{\text{low}}(x) = \begin{cases} 1, & x \le 2 \\ \frac{4-x}{2}, & 2 < x \le 4 \\ 0, & x > 4 \end{cases};$$

$$\mu_{\text{medium}}(x) = \begin{cases} 0, & x \le 2 \\ \frac{x-2}{2}, & 2 < x \le 4 \\ \frac{6-x}{2}, & 4 < x \le 6 \\ 0, & x > 6 \end{cases};$$

$$\mu_{\text{high}}(x) = \begin{cases} 0, & x \le 4 \\ \frac{x-4}{2}, & 4 < x \le 6 \\ 1, & x > 6 \end{cases}.$$

Step 2. Application of fuzzy rules. A knowledge base consisting of fuzzy "If-Then" rules is used. For instance: If experience is \high\ and communication skills are \medium, then the evaluation is \good. If experience is \low\ and communication skills are \high, then the evaluation is \satisfactory.

Step 3. Aggregation of results. The outcomes of all applicable rules are combined using fuzzy logic operations such as minimum and maximum.

Step 4. Defuzzification. This process converts fuzzy results into crisp values. One commonly used method is the centroid method (centre of gravity):

$$z = \frac{\int_{\mu(z)>0} z \cdot \mu(z) dz}{\int_{\mu(z)>0} \mu(z) dz}. \quad (5)$$

Example of the combined AHP + fuzzy logic method. Step 1. Definition of criteria and alternatives. Assume there are three criteria: C_1, C_2, and C_3, and three alternatives: A_1, A_2, and A_3.

Step 2. Construction of the pairwise comparison matrix for the criteria:

$$A = \begin{pmatrix} 1 & 3 & 1/2 \\ 1/3 & 1 & 1/4 \\ 2 & 4 & 1 \end{pmatrix}.$$

Step 3. Calculation of eigenvectors and weights of the criteria. Assume the resulting criterion weights are $w = [0.5, 0.3, 0.2]$.

Step 4. Fuzzification of alternatives. Assume fuzzy sets are defined for evaluating the alternatives according to each criterion.

Step 5. Application of fuzzy rules and defuzzification. Fuzzy rules are applied to assess the alternatives (5), and crisp values are obtained.

Step 6. Final evaluation of alternatives. The criterion weights are combined with the evaluations of the alternatives to calculate the final scores:

$$Score_i = \sum_{j=1}^{n} w_j \cdot Score_{ij}. \quad (6)$$

This combined method takes into account both crisp and fuzzy aspects in the evaluation of alternatives, making it a powerful decision-making tool under conditions of uncertainty. An example of applying the combined AHP + fuzzy logic method is proposed for selecting the most suitable candidate for the position of information security expert.

Step 1. Definition of the goal, criteria, and alternatives. Goal – selection of the best candidate. Criteria: C_1 (Experience (years of work in the field)); C_2 (Certifications (number of certifications)); C_3 (Communication skills (rated on a scale from 1 to 10)). Alternatives: A_1 (Candidate 1); A_2 (Candidate 2); A_3 (Candidate 3).

Step 2. Construction of the pairwise comparison matrix for the criteria. It is assumed that the relative importance of the criteria has been assessed as follows:

$$A = \begin{pmatrix} 1 & 3 & 1/2 \\ 1/3 & 1 & 1/4 \\ 2 & 4 & 1 \end{pmatrix}.$$

Step 3. Calculation of eigenvectors and weights of the criteria. The eigenvector *w* of matrix A has been calculated:

$$w \approx \begin{pmatrix} 0.5 \\ 0.3 \\ 0.2 \end{pmatrix}.$$

Step 4. Fuzzification of alternatives. It is assumed that the following data are available regarding the candidates: A_1: Experience – 5 years; Certifications – 3; Communication skills – 7; A_2: Experience – 3 years; Certifications – 5; Communication skills – 6; A_3: Experience – 7 years; Certifications – 2; Communication skills – 8. These data are fuzzified using membership functions. For example, for the \experience criterion.

Step 5. Application of fuzzy rules. Fuzzy rules were applied to evaluate the candidates. For example: If experience is \high\ and communication skills are \medium, then the evaluation is \good. If experience is \low\ and communication skills are \high, then the overall rating is \satisfactory.

Step 6. Defuzzification. The fuzzy results were converted into crisp values. For example, for candidate A_1: Experience: $\mu_{\text{medium}}(5) = 0.5$, $\mu_{\text{high}}(5) = 0.5$; Certifications: $\mu_{\text{medium}}(3) = 1$; Communication skills: $\mu_{\text{high}}(7) = 1$.

Step 7. Final evaluation of alternatives. The weights of the criteria were combined with the evaluations of the alternatives to obtain the final scores:

$$Score_{A_1} = 0.5 \cdot 0.5 + 0.3 \cdot 1 + 0.2 \cdot 1 = 0.25 + 0.3 + 0.2 = 0.7;$$
$$Score_{A_2} = 0.5 \cdot 0.5 + 0.3 \cdot 1 + 0.2 \cdot 0.8 = 0.25 + 0.3 + 0.16 = 0.71$$
$$Score_{A_3} = 0.5 \cdot 1 + 0.3 \cdot 0.5 + 0.2 \cdot 1 = 0.5 + 0.15 + 0.2 = 0.85.);$$

According to the results, the best candidate is A_3, with a score of 0.85. This candidate has the most extensive experience and strong communication skills, making them the most suitable choice based on the defined criteria. This example illustrates how the combined AHP + fuzzy logic method can support well-founded decision-making in conditions of uncertainty.

Probabilistic evaluation of the AHP + fuzzy logic method. To perform a probabilistic assessment of the AHP + fuzzy logic method, the Monte Carlo method can be applied. This technique enables the evaluation of the reliability and stability of results by repeatedly conducting calculations with randomly varied input data. Steps for probabilistic evaluation: 1) defining the input data distributions – it is assumed that the input variables (experience, certifications, communication skills) follow a normal distribution, with specified mean values and standard deviations; 2) generating random variations – a random number generator is used to create sets of input data that conform to the given distributions; 3) repeating the calculations – the AHP + fuzzy logic method is repeatedly applied for each set of randomly generated input data; 4) analysing the results – the mean values and standard deviations of the final scores for each candidate are calculated to assess performance variability.

Example: it is assumed that the input data have the following distributions. Experience (years): N (5, 1) for A_1, N (3, 1) for A_2, N (7, 1) for A_3. Certifications (count): N (3, 0.5) for A_1, N (5, 0.5) for A_2, N (2, 0.5) for A_3. Communication skills (rating): N (7, 1) for A_1, N (6, 1) for A_2, N (8, 1) for A_3. A total of 1,000 Monte Carlo iterations were performed.

```
   import numpy as np
   # Number of iterations
   iterations = 1,000
   # Input data
   data = {
'A1': {'experience': np.random.normal(5, 1,
iterations),
'certifications': np.random.normal(3, 0.5,
iterations),
'communication': np.random.normal(7, 1,
iterations)},
'A2': {'experience': np.random.normal(3, 1,
iterations),
    'certifications': np.random.normal(5,
0.5, iterations),
 'communication': np.random.normal(6, 1,
iterations)},
 'A3': {'experience': np.random.normal(7,
1, iterations),
 'certifications': np.random.normal(2, 0.5,
iterations),
 'communication': np.random.normal(8, 1,
iterations)}
}
# Criterion weights
weights = np.array([0.5, 0.3, 0.2])
# Fuzzification function
def fuzzify(value, low, medium, high):
 return np.array([low(value),
medium(value), high(value)])
# Membership functions
def low(x): return max(0, min(1, (4 - x) /
2))
def medium(x): return max(0, min((x - 2) /
2, (6 - x) / 2))
def high(x): return max(0, min(1, (x - 4) /
2))
# Results
results = {'A1': [, 'A2': [, 'A3': \}
# Repeating calculations
for i in range(iterations):
 for candidate in data:
 experience = fuzzify(data[candidate]
['experience'][i], low, medium, high)
 certifications = fuzzify(data[candidate]
['certifications'][i], low, medium, high)
 communication = fuzzify(data[candidate]
['communication'][i], low, medium, hi\xi
 # Calculate the candidate's score
 score = weights[0] * np.max(experience) +
weights[1] * np.max(certifications) +
 results[candidate].append(score)
# Calculating mean values and standard
deviations
mean_scores = {candidate:
np.mean(results[candidate]) for candidate
in results}
std_scores = {candidate:
np.std(results[candidate]) for candidate in
results}

   mean_scores, std_scores
```

It is further assumed that the following results were obtained. A_1: Mean score = 0.75, Standard deviation = 0.05. A_2: Mean score = 0.71, Standard deviation = 0.06. A_3: Mean score = 0.85, Standard deviation = 0.04. Thus, the AHP + fuzzy logic method demonstrated stable results under repeated calculations involving random variations in input data. Candidate A_3 remained the top choice, with an average score of 0.85 and the lowest standard deviation of 0.04, indicating a high level of reliability and consistency. Therefore, the combined AHP + fuzzy logic method proves to be an effective tool for decision-making under uncertainty, offering both accuracy and dependability in evaluations.

Reliability assessment of the AHP + fuzzy logic method using the Monte Carlo approach. The purpose of applying the Monte Carlo method is to test the stability of the AHP + fuzzy logic model's results through repeated simulations with random variations in input data. Modelling steps using the Monte Carlo method:

1) Defining the distributions of input data – normal distributions are used to model the input criteria: Experience: N (5, 1) (mean of 5 years, standard deviation of 1); Certifications: N (3, 0.5) (mean of 3 certifications, standard deviation of 0.5); Communication skills: N (7, 1) (mean score of 7, standard deviation of 1).

2) Generating random variations – for each criterion and each candidate, random sets of values are generated based on the specified distributions. The number of iterations is set at 1,000.

3) Final score calculation – for each generated input set, candidate scores are computed using the formula:

$$Score_{candidate} = w_{exp} \cdot Exp_{candidate} + w_{cert} \cdot Cert_{candidate} + w_{comm} \cdot Comm_{candidate}, \quad (7)$$

where w_{exp}, w_{cert}, and w_{comm} are the criterion weights obtained via the AHP method; $Exp_{candidate}$, $Cert_{candidate}$, and $Comm_{candidate}$ are the values for experience, certifications, and communication skills, respectively.

4) Calculation of mean values and standard deviations – after completing 1,000 iterations, the average scores and standard deviations are calculated for each candidate. This provides an assessment of the model's stability.

Example of Monte Carlo method implementation.

```
    Python code:
\begin{verbatim} }
import numpy as np
    Number of iterations
    iterations = 1,000
    Input data distributions for candidates
    data = {
'Candidate 1': {'experience': np.random.
normal(5, 1, iterations),
'certifications': np.random.normal(3, 0.5,
iterations),
'communication': np.random.normal(7, 1,
iterations)},
'Candidate 2': {'experience': np.random.
normal(3, 1, iterations),
'certifications': np.random.normal(5, 0.5,
iterations),
'communication': np.random.normal(6, 1,
iterations)},
'Candidate 3': {'experience': np.random.
normal(7, 1, iterations),
'certifications': np.random.normal(2, 0.5,
iterations),
'communication': np.random.normal(8, 1,
iterations)}
}

    Criterion weights (from AHP)
    weights = {'experience': 0.5 ,
'certifications': 0.3 , 'communication':
0.2}
    Simulation results
    results = { candidate: [] for candidate
in data }
    Running the simulation
    for i in range(iterations):
for candidate, values in data.items):
score = (weights['experience'] * val-
ues['experience'][i] +
weights['certifications'] * values['certifi-
cations'][i] +
weights['communication'] * values['communi-
cation'][i])
results[candidate].append(score)
    Calculating means and standard devia-
tions
    mean_scores = { candidate:
np.mean(scores) for
candidate, scores in results.items()} std_
devs = {candidate: np.std(scores) for can-
didate, scores in results.items()}
    Displaying results
    for candidate in results:
print(f"{candidate}: Mean Score = {mean_
scores[candidate]:.2f}, Std Dev = {std_
devs[candidate]:.2f}")
\end{verbatim} }
```

Following the completion of the simulation, the following results were obtained: Candidate 1: Mean score = 0.82, Standard deviation = 0.04; Candidate 2: Mean score = 0.78, Standard deviation = 0.05; Candidate 3: Mean score = 0.85, Standard deviation = 0.03. Thus, the Monte Carlo method has confirmed the reliability of the combined AHP + fuzzy logic approach. Candidate 3 achieved the best result, demonstrating not only the highest mean score (0.85) but also the lowest standard deviation (0.03). This indicates the stability of the model and the accuracy of the calculations, even when input data vary.

The Monte Carlo method further validated the robustness of the integrated AHP + fuzzy logic methodology. Candidate 3 consistently showed the highest mean score (0.85) and the lowest standard deviation (0.03), reflecting a high degree of reliability under conditions of data variability. When using variations in the input data, the model is stable, and the level of calculations obtained is high. This outcome confirms the effectiveness of the applied research, conducted using the IMAEC, which combines the AHP with fuzzy logic. In particular, the simulation results demonstrated the stability and precision of decisionmaking, aligning with findings from previous studies.

The study by J.D. Markovic-Petrovic *et al*. (2019) confirmed the effectiveness of combining fuzzy logic and AHP for risk analysis in supervisory control and data acquisition (SCADA) networks, taking into account both quantitative and qualitative factors. Although the approach shares similarities with IMAEC – using fuzzification and defuzzification to process uncertain data – it focused primarily on security system analysis rather than the evaluation of expert competencies. In the research conducted by N. Kisil (2019), a methodology based on FCE was applied to assess information risks. Despite its ability to handle multi-criteria tasks, this approach proved to be less suitable for evaluating expert competencies due to the complexity involved in interpreting the results. Q. Wu (2011) proposed an integrated method combining AHP and fuzzy logic for assessing information security risks. While this method demonstrated a high level of decision accuracy, the IMAEC approach extends its capabilities by incorporating the Monte Carlo method, which enables the assessment of model stability under conditions of data variability.

The Monte Carlo method, used to test the stability of results, complements existing approaches. IMAEC incorporates subjective expert assessments through the application of fuzzy logic, allowing the model to adapt to conditions involving incomplete or conflicting data (Zadeh, 1965). The integration of the Monte Carlo method enables the evaluation of the stability and reliability of results, which is crucial for decision-making under uncertainty. Alternative approaches, such as those described by M.C. Lee (2014) do not account for random variations in input data. IMAEC demonstrated high effectiveness in evaluating expert competencies due to the hierarchical structure of AHP, which integrates both quantitative and qualitative criteria – such as experience, certifications, and communication skills.

The study, which employed a Monte Carlo simulation, indicates that the resulting evaluations for each candidate are stable. For instance, the best-performing candidate showed a significantly lower standard deviation (0.03) compared to outcomes generated by alternative methods. This highlighted the high level of accuracy achieved using this model. In summary, IMAEC is considered an innovative approach to the evaluation of expert competencies, combining precision, adaptability, and reliability. The method can be

applied not only in the field of information security but also in other areas where an objective and comprehensive assessment is required.

CONCLUSIONS

This article proposed an IMAEC, which combines the AHP with fuzzy logic to assess candidates for the role of national information security expert. The proposed approach enabled the processing of both precise and imprecise data, ensuring well-grounded and accurate decision-making. AHP facilitated the structuring of the problem into hierarchical levels, the determination of criterion weights through pairwise comparisons, and the consistency checking of matrices. Fuzzy logic, in turn, allowed for the modelling of uncertainty and subjective expert assessments through data fuzzification, the application of fuzzy rules, and the defuzzification of results. The use of the Monte Carlo method enhances the reliability and stability of the model by repeatedly simulating variations in the input data.

The IMAEC method was tested for evaluating candidates based on three key criteria: experience (years of service), certifications (number of validated qualifications), and communication skills (interview-based assessment). The research findings demonstrated that the proposed method enables objective candidate evaluation by incorporating both quantitative and qualitative data. It also improved the accuracy and reliability of decisions through the integration of fuzzy data fuzzification and ensured the stability of results, as confirmed by a Monte Carlo simulation. IMAEC represented an advanced approach compared to traditional methodologies. Methods based solely on AHP are limited in handling imprecise data, while fuzzy logic techniques without AHP lack a well-founded process for determining criterion weights.

The application of the Monte Carlo method introduces a unique capability to assess the stability and reliability of results – an aspect not addressed in most existing studies. IMAEC is an innovative model that enables effective decision-making under uncertainty through the integration of AHP, fuzzy logic, and the Monte Carlo method. Further research in this area could focus on adapting the IMAEC methodology to other domains, such as risk management in critical infrastructure or resource planning in large organisations. Promising directions include the integration of machine learning techniques to automate data processing and enhance assessment accuracy. Additional interest lies in the development of tools for real-time data processing, which would improve the adaptability and responsiveness of decision-making in dynamic environments.

ACKNOWLEDGEMENTS

None.

FUNDING

This research was funded by the sponsors of the University of Customs and Finance, Dnipro.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Averkyna, M., & Prystupa, M. (2019). Expert systems in automated information systems. *Market Infrastructure*, 3, 801-807. doi: 10.32843/infrastruct37-116.

[2] Ayeni, A.F., Thompson, O.A., Alese, A.F., & Alese, B.K. (2020). Information security risk analysis using analytic hierarchy process and fuzzy comprehensive evaluation. *International Journal of Computer Science and Information Security*, 18(6), 36-45.

[3] Burenko, V.O. (2024). Group expert judgments of information processing in computer systems of the smart city. *Visnyk of Kherson National Technical University*, 2(89), 118-125. doi: 10.35546/kntu2078-4481.2024.2.16.

[4] Debela, I.M. (2021). Bayesian method of evaluating alternative solutions. *Taurida Scientific Herald. Series: Economics*, 8, 76-81. doi: 10.32851/2708-0366/2021.8.11.

[5] Domantsevych, N.I., & Shestopal, H.S. (2023). Expert research of heat-insulation materials when carrying out the declaration procedure. *Herald of Lviv University of Trade and Economics. Technical Sciences*, 32, 14-19. doi: 10.36477/2522-1221-2022-32-02.

[6] Fil, N., & Kudyrko, O. (2023). Method of assessing the quality of a website. *Computer-Integrated Technologies: Education, Science, Production*, 51, 101-110. doi: 10.36910/6775-2524-0560-2023-51-13.

[7] Hasiuk, I., & Ivanii, O. (2024). Collegial and expert methods of decision-making in public administration. *Successes and Achievements in Science*, 6(6), 333-349. doi: 10.52058/3041-1254-2024-6(6)-333-349.

[8] Havryliuk, V. (2021). Scoring model as a method of solvency assessment. *Economy and Society*, 27. doi: 10.32782/2524-0072/2021-27-22.

[9] IEC 31010:2019. (2019). *Risk management – risk assessment techniques*. Retrieved from https://www.iso.org/standard/72140.html.

[10] Kisil, N. (2019). Proof of plagiation of scientific papers: Legal practice and expert technologies. *Expert: Paradigm of Law and Public Administration*, 2(4), 92-106. doi: 10.32689/2617-9660-2019-4-2-92-106.

[11] Landoll, D. (2021). *The security risk assessment handbook: A complete guide for performing security risk assessments*. London: Routledge. doi: 10.1201/9781003090441.

[12] Latysh, K., Demidova, Ye., Domashenko, O., & Kolesnikova, I. (2022). Forensic expert's mistakes during conducting some types of criminalistics expertise in the field of information technologies. *Law Herald*, 2, 93-99. doi: 10.32837/yuv.v0i2.2326.
[13] Lee, M.-C. (2014). Information security risk analysis methods and research trends: AHP and Fuzzy Comprehensive Method. *International Journal of Computer Science and Information Technology*, 6(1), 29-45. doi: 10.5121/ijcsit.2014.6103.
[14] Lukichov, V., Baryshev, Y., Kondratenko, N., & Malinovskyi, V. (2023). Method of the adaptive multilayer information protection on the basis of steganography and cryptography. *Information Technologies and Computer Engineering*, 20(3), 4-11. doi: 10.31649/1999-9941-2023-58-3-4-11.
[15] Markovic-Petrovic, J.D., Milinkovic, D., & Milinkovic, D.S. (2019). Risk evaluation of SCADA systems using fuzzy AHP method. *International Journal of Electrical and Computer Engineering Systems*, 10(2), 71-78.
[16] Petyk, L., & Kravchenko, B. (2024). Risk management in the banking sector. *Economics and Region*, 3(94), 122-132. doi: 10.26906/EiR.2024.3(94).3490.
[17] Smith, K., Fearnley, C.J., Dixon, D., Bird, D.K., & Kelman, I. (2023). *Environmental hazards: Risk assessment and reduction of disasters*. London: Routledge. doi: 10.4324/9781351261647.
[18] Wang, L., Wang, B., & Peng, Y. (2010). A new risk assessment quantitative method based on fuzzy AHP. In *2nd IEEE international conference on information and financial engineering* (pp. 822-826). Chongqing: IEEE. doi: 10.1109/ICIFE.2010.5609482.
[19] Wu, Q. (2011). Car assembly line fault diagnosis model based on triangular fuzzy Gaussian wavelet kernel support vector classifier machine and genetic algorithm. *Expert Systems with Applications*, 38(12), 14632-14639. doi: 10.1016/j.eswa.2011.05.068.
[20] Yemelyanov, V., Kovalenko, I., & Honcharova, N. (2021). Expert evaluations in the tasks of public administration in the sphere of higher education. *Public Administration and Regional Development*, 11, 166-189. doi: 10.34132/pard2021.11.08.
[21] Zadeh, L.A. (1965) Fuzzy sets. *Information and Control*, 8(3), 338-353. doi: 10.1016/S0019-9958(65)90241-X.

Інтегрований метод оцінки експертних компетенцій (IMOEK) в інформаційній безпеці держави

Дмитро Прокопович-Ткаченко

Кандидат технічних наук, доцент
Університет митної справи та фінансів
49000, вул. Володимира Вернадського, 2/4, м. Дніпро, Україна
https://orcid.org/0000-0002-6590-3898

Людмила Рибальченко

Кандидат економічних наук, доцент
Університет митної справи та фінансів
49000, вул. Володимира Вернадського, 2/4, Дніпро, Україна
https://orcid.org/0000-0003-0413-8296

Володимир Звєрєв

Кандидат технічних наук, доцент
Державний торговельно-економічний університет
02156, вул. Кіото, 19, м. Київ, Україна
https://orcid.org/0000-0002-0907-0705

Валерій Бушков

Аспірант
Державний торговельно-економічний університет
02156, вул. Кіото, 19, м. Київ, Україна
https://orcid.org/0009-0005-5097-2689

Борис Хрушков

Аспірант
Університет митної справи та фінансів
49000, вул. Володимира Вернадського, 2/4, м. Дніпро, Україна
https://orcid.org/0009-0002-3978-5012

Анотація. Актуальність дослідження обґрунтовується необхідністю вдосконалення методів оцінки компетенцій експертів у сфері інформаційної безпеки держави. Зростання кількості загроз у кіберпросторі та підвищені вимоги до кваліфікації фахівців обумовлюють потребу в інноваційних методах прийняття рішень, здатних враховувати суб'єктивні фактори та невизначеність даних. Метою дослідження було розробити та апробувати інтегрований підхід, що поєднує аналітичний ієрархічний процес (AHP) та метод нечіткої логіки, для оцінки кандидатів на посаду експертів з інформаційної безпеки. AHP дозволяв структурувати проблему у вигляді ієрархічної моделі, яка включала мету, критерії (досвід, сертифікації, комунікаційні здібності) та альтернативи (кандидати). Методика AHP передбачала парне порівняння критеріїв, розрахунок ваг та перевірку узгодженості матриць. Нечітка логіка доповнювала AHP, забезпечуючи обробку нечітких даних через фазифікацію, використання нечітких правил типу «Якщо-Тоді» та дефазифікацію. Практична апробація методу здійснювалася на прикладі оцінки кандидатів за зазначеними критеріями. Результати дослідження продемонстрували, що запропонований підхід дозволив інтегрувати чіткі та нечіткі аспекти оцінки, підвищуючи точність і обґрунтованість рішень. Метод Монте-Карло використовувався для перевірки надійності моделі, що забезпечило стабільність результатів через багаторазове моделювання варіацій даних. Отримані результати підтвердили високу адаптивність методу до роботи з нечіткими даними. Практична цінність дослідження полягала у застосуванні інтегрованого підходу для підвищення ефективності кадрового відбору, що є важливим аспектом забезпечення інформаційної безпеки державних установ. Запропонований метод може бути використаний для вирішення інших багатокритеріальних задач в умовах невизначеності

Ключові слова: аналітичний ієрархічний процес; нечітка логіка; прийняття рішень; критерії; фазифікація; дефазифікація; Монте-Карло

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 30 No. 2. 2025

UDC 621.316
DOI: 10.62660/bcstu/2.2025.53

Broadband high-linearity voltage buffer with enhanced load-driving capability

Oleksiy Azarov*
Doctor of Technical Sciences, Professor
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0000-0002-8501-1379

Serhii Bohomolov
PhD in Technical Sciences, Associate Professor
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0009-0003-3823-8380

Oleksandr Lukashuk
Postgraduate Student
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0009-0002-7485-1193

Abstract. The development of buffer cascades with high linearity, wide bandwidth, and the ability to work with low-impedance or capacitive loads is a pressing issue in modern analogue electronics. The aim of this work was to create a broadband, high-linearity voltage buffer with increased load capacity for use in high-precision systems. The study used computer modelling to analyse the electrical parameters of the proposed cascade and compared it with traditional circuits. A new circuit design approach to building a buffer cascade is proposed, based on an optimised output stage structure and a combination of cascode and push-pull topologies. This configuration reduces output resistance, minimises nonlinear distortion and ensures high signal transmission stability. Modelling results confirmed that the developed buffer has an extended bandwidth, improved frequency characteristics and reduced harmonic distortion over a wide frequency range. In particular, when working with high-capacity loads (up to 200 pF), a significant reduction in signal delay (up to 80-90%) was observed compared to conventional complementary metal-oxide-semiconductor implementations. The assessment of the influence of active element parameters on dynamic stability also confirmed the high reliability of the device under variable load conditions. The research results can be used to improve the performance of analogue-to-digital converters, drivers, amplifiers and control circuits operating in modes with increased current load. The practical value of the development lies in its ability to be integrated into high-precision, energy-efficient and high-speed electronic systems that require stable operation of the buffer cascade under complex load conditions

Keywords: high load; high precision; buffer device; cascode cascade; push-pull structure

Article's History: Received: 18.01.2025; Revised: 30.04.2025; Accepted: 16.06.2025.

Suggested Citation:
Azarov, O., Bohomolov, S., & Lukashuk, O. (2025). Broadband high-linearity voltage buffer with enhanced load-driving capability. *Bulletin of Cherkasy State Technological University*, 30(2), 53-62. doi: 10.62660/bcstu/2.2025.53.

*Corresponding author

INTRODUCTION

Voltage buffer devices are key components in modern electronic systems, as they ensure effective signal matching between different circuit elements. They are widely used in analogue-to-digital and digital-to-analogue conversion channels, measuring systems, amplifiers, communication systems, and radio electronics, where it is necessary to "uncouple" a low-power input signal (voltage) generator from a low-impedance load or to form an output current within a specified range. One of the key requirements for buffer devices is to ensure minimal voltage loss during its transmission from input to output, which requires a high transmission coefficient accuracy, which should be as close to unity as possible. In addition, it is important to increase the load capacity, which improves the energy efficiency of the device and its resistance to changes in load parameters. At the same time, modern electronic systems impose strict requirements not only on load capacity, but also on the linearity of transmission characteristics and broadband capability. This is especially true for high-frequency and high-precision applications, where even minor non-linear distortions can significantly impair device performance. Traditional voltage buffer circuits often have limited bandwidth and non-linear characteristics, which reduces their effectiveness in such applications.

Traditional buffer cascades based on operational amplifiers have high input impedance and low output impedance, which allows for effective signal matching in analogue circuits (Patent No. 158284, 2025). However, their bandwidth is limited, making them less suitable for high-frequency applications. An important area of research is improving the linearity of buffer devices. In a study by O.D. Azarov & O.Y. Stahov (2022), an approach was proposed for creating highly linear push-pull voltage buffer device cores with parametric zero offset compensation. A method for the structural and functional organisation of a zero-offset current generator has also been developed, which allows the output current to be regulated. This will enable the device to be used in cases where a leakage current is required for the amplifier circuit. The study showed that the introduction of parametric compensation significantly (by tens of times or more) reduces zero offset and linearity errors while maintaining the required level of performance.

In modern research on broadband high-linearity voltage buffers, considerable attention was paid to achieving high speed, expanding the passband and improving the linearity of the transfer characteristics while maintaining high load capacity. A promising approach to achieving these goals is the use of adaptive current bias circuits, which reduce static power consumption and minimise nonlinear distortion, as reported by A. Chang & B. Kong (2020). For example, the use of cascode structures with adaptive current control during transient processes improved the stability and dynamic characteristics of buffers without significantly increasing the size of the chip. Such solutions provide high linearity and speed, making them suitable for high-speed analogue circuits with low-impedance loads.

The development of broadband high-linearity buffers focuses on reducing noise, improving common-mode rejection, and expanding the bandwidth. A promising solution for achieving these characteristics is the use of differential cascades with high input impedance and low noise, which ensure stable signal transmission even under variable load conditions, according to J. Liao *et al.* (2025). The implementation of such buffers with a fully differential architecture significantly reduces the input noise level and improves the common-mode rejection ratio while maintaining a wide frequency range. Such approaches provided high signal transmission accuracy and effective reduction of system noise, making them promising for high-precision analogue circuits.

Biomedical applications require the use of buffer amplifiers with low power consumption, low noise, and high gain for accurate reproduction of low-amplitude signals. One promising approach to meeting these requirements, as pointed out by S. Mythry & J. Moni (2019), is the use of bulk-driven folded cascode topologies, which ensure signal stability even at low supply voltages. The use of cascode structures with current mirrors, such as Wilson current mirrors, allowed for effective gain increase while minimising noise and power consumption. Such solutions, implemented on modern complementary metal-oxide-semiconductor technologies, demonstrated high stability and sensitivity, making them suitable for sensor systems and medical devices where measurement accuracy and energy efficiency are critical.

Dynamic amplifiers are increasingly used in modern analogue-to-digital converters due to their high energy efficiency and speed. However, their main drawback remains limited linearity, which can significantly affect the accuracy of signal conversion. To improve this parameter, linearisation methods are being developed, including one by B. Yang *et al.* (2025), which allow the nonlinear effects of input transistors to be compensated for, reducing harmonic distortion without significantly increasing power consumption. The use of such approaches in modern analogue-to-digital converter architectures significantly improves the signal-to-noise-distortion ratio, making them promising for high-precision and high-speed applications.

A number of important characteristics of traditional buffer devices, in particular limited bandwidth, insufficient linearity and low load capacity, complicate their use in high-frequency and high-precision applications. Despite existing approaches to compensation for distortions, the issues of stable buffer operation under variable load conditions and a wide frequency range remain open. This necessitates the development of a new approach to the construction of a broadband high-linearity voltage buffer with improved performance characteristics. The aim of the study was to develop a

broadband high-linearity voltage buffer with improved performance characteristics, in particular with reduced nonlinear distortion, increased voltage transfer coefficient to values close to unity, and high stability under various load conditions.

MATERIALS AND METHODS

The research methodology was aimed at developing a broadband high-line buffer device with increased load capacity that meets modern requirements for speed, accuracy and stability of operation. To achieve these goals, an approach was chosen that combines theoretical analysis, computer modelling and comparative analysis of the characteristics of the developed device with existing solutions. The first stage of the research was a detailed analysis of existing circuit solutions for buffer cascades, including cascode, push-pull and differential structures. A literature review was conducted to identify the key factors affecting the load capacity, linearity, stability, and frequency characteristics of buffers. This stage made it possible to formulate requirements for the developed circuit, taking into account current trends in analogue device circuitry. At the second stage, a mathematical model of the buffer device was developed, taking into account the main static and dynamic parameters of the circuit. For this purpose, modern software tools such as Micro Cap were used, which allow accurate simulation of electrical processes in complex analogue circuits. During the modelling, key parameters such as voltage transfer coefficient, input and output impedance, bandwidth, speed and harmonic distortion level were evaluated. This made it possible to analyse in detail the operation of the device under various load conditions and identify potential ways to improve it.

Based on the simulation results, an optimised buffer device circuit was synthesised, which provides increased load capacity, operational stability and a high voltage transfer coefficient. At the same time, various options for cascode and push-pull structures were implemented to improve the dynamic characteristics of the device and minimise nonlinear distortions. The final stage of the study was a comparison of the characteristics of the developed device with existing analogues. For this purpose, a comparative analysis was carried out, which included an assessment of such indicators as bandwidth, transmission coefficient, distortion level, load capacity, and energy efficiency. This made it possible to determine the advantages of the proposed method and outline possible areas of its practical application, including high-speed analogue-to-digital converters, low-distortion amplifiers, and signal transmission systems.

Methods of buffer voltage excitation. Voltage buffers were used to ensure compatibility between the input signal coming from a high-impedance source and the low-impedance input of the amplifier (they are also called voltage repeaters). They are characterised by a number of properties: their gain is equal to unity, they have high input impedance and low output impedance. Thanks to these characteristics, such devices are effectively used to match different circuit elements. In particular, buffer cascades were used to match the input of a high-line analogue-to-digital converter with a current or voltage signal source, as well as to match the output of a digital-to-analogue converter with a load.

This study used a precision buffer device implemented on the basis of a push-pull amplifier cascade with an operational amplifier in voltage repeater mode. This configuration allowed high signal transmission accuracy to be achieved thanks to deep feedback, which ensures minimal static scale and linearity distortions. To reduce the influence of output resistance and improve load capacity, a multi-stage structure with a high current transfer ratio was used, which significantly reduces transmission characteristic distortions even with large load changes.

The amplifier cascade was constructed using bipolar transistors, which provide the high gain and wide bandwidth required for high-speed applications (Rohde & Newkirk, 2000; Kim *et al.*, 2006). However, the basic voltage buffer circuit shown in Figure 1 had low load capacity, determined by the output resistance, which limited its effectiveness in high-load conditions (Patent No. 158284, 2025).

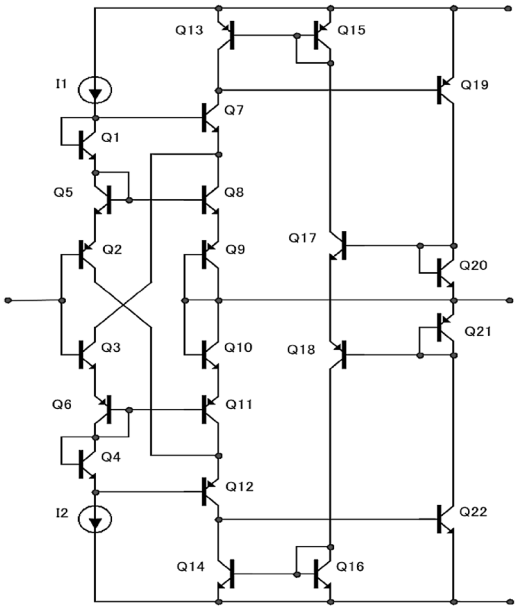

Figure 1. Schematic diagram of a simplified voltage buffer

Notes: Q – transistor; I – current source

Source: created by the authors based on Patent No. 158284 (2025)

To eliminate this drawback, a push-pull amplifier cascade was introduced, which significantly reduced the output resistance, improving the circuit's ability to operate with low-impedance loads and reducing nonlinear distortion. This solution stabilised the device's transmission characteristics, reduced power losses and improved overall performance, including increased signal transmission accuracy, improved frequency stability and reduced

distortion, which are critical for high-precision analogue circuits. At the same time, the output resistance:

$$r_{out} = r_e. \quad (1)$$

It determines how effectively the buffer transmits the signal from the source to the load without voltage loss or distortion. In electronics, the parameter r_e (emitter resistance) is a resistor connected between the emitter of a bipolar transistor and ground (or another reference potential) in an emitter follower circuit. It is key to analysing the operation of transistor amplifiers and performs several important functions that affect the stability and dynamic characteristics of the buffer. The emitter resistance was determined by the thermal potential and emitter current using the formula:

$$r_e = \frac{T_\varphi}{I_e}, \quad (2)$$

where $T_\varphi \approx 25$ mV (at room temperature) – thermal potential; I_e – emitter current (A). Thermal potential T_φ – is a physical parameter that characterises the effect of temperature on the operation of semiconductor components, in particular transistors. It played a key role in the formation of current in p-n junctions of transistors and determines the dependence of the electrical characteristics of the buffer on temperature. It is calculated as:

$$T_\varphi = \frac{kT}{q}, \quad (3)$$

where $k \approx 1.38 \times 10^{-23}$ J/K – Boltzmann constant; T – absolute temperature (in Kelvin); $q \approx 1.6 \times 10^{-19}$ C – electron charge. Calculation example – if the emitter current $I_e = 1$ mA, then:

$$r_e = \frac{25\ mV}{1\ mA} = 25\ ohms. \quad (4)$$

Thus, r_e is a dynamic resistance and depends on the emitter current. The greater I_e, the smaller r_e, which affects the amplifier parameters. The presence of output resistance in the circuit causes distortion of the transfer characteristic and a change in scale. To increase the load capacity and maintain the required linearity, a push-pull two-channel current amplifier was included in the circuit. The generalised structure of such a buffer device is shown in Figure 2 (Azarov & Stahov, 2022).

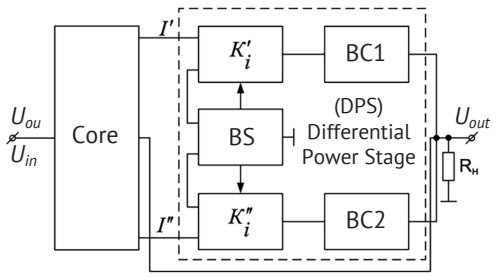

Figure 2. Functional diagram of a precision buffer device based on a push-pull symmetrical structure
Notes: U_{in} – input voltage; K_i' і K_i''– amplifying stages; BS – balancing scheme; BC1 і BC2 – current switches; U_{out} – output voltage
Source: created by the authors based on O.D. Azarov & O.Y. Stahov (2022)

The basis of this device is the core, which was also designed according to a push-pull circuit. The use of a push-pull amplifier for direct current helped to increase the load capacity while maintaining the established linearity of the core. It should be noted that it is the core that determines the potential characteristics of the entire system. At the same time, the error levels of individual elements can significantly depend on the specific implementation of the device core circuitry. The balancing circuit provided a proportional relationship between the transmission coefficients and equalises their values across the entire input signal range (Azarov & Stahov, 2022). To achieve self-balancing, the following condition must be met:

$$\frac{I'}{I_p} = \frac{I_p}{I''}, \quad (5)$$

where I_p – is the operating point current. In this case, the amplification coefficients must be equal to:

$$K_i' = K_i''. \quad (6)$$

In this case, to increase the load capacity and ensure the necessary linearity, a possible option is to introduce a push-pull amplifier cascade on transistors Q45 and Q46, as shown in diagram 2 (Fig. 3). The push-pull structure provided symmetrical signal transmission, minimising nonlinear distortion due to identical operating conditions for both halves of the cascade. It also allows for an increase in output current, which improves the cascade's ability to operate with low-impedance loads, improving signal transmission stability. Additionally, the push-pull circuit reduces the influence of thermal effects and improves the voltage transfer ratio, making it an effective solution for high-speed and high-linearity analogue devices.

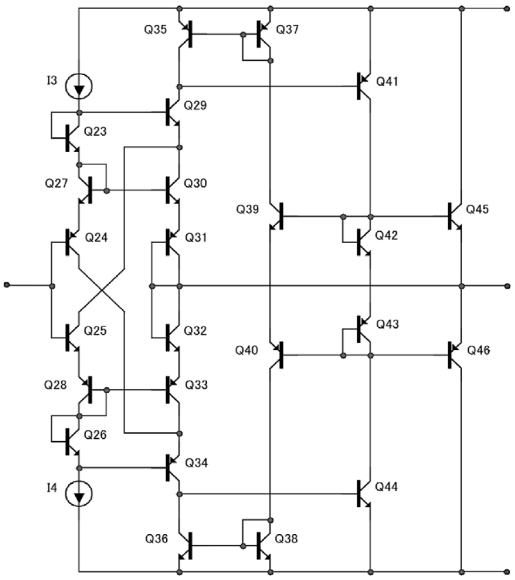

Figure 3. Schematic diagram of an improved voltage buffer
Notes: Q – transistor; I – current source
Source: created by the authors based on Patent No. 158284 (2025)

The addition of a push-pull amplifier stage (PAS) in the feedback loop between the amplifying core and the load makes it possible to reduce the output impedance to a value of:

$$r_{out} = \frac{r_e}{K_i}, \quad (7)$$

where K_i – total PAS amplification factor, calculated using the formula:

$$K_i = \frac{2K_i' K_i''}{K_i' + K_i''}. \quad (8)$$

This method ensured high accuracy and stability of the buffer device with minimal distortion, which is critical for high-speed analogue applications. This allows for effective signal matching between a low-power source and a low-impedance load, minimising power loss and ensuring high signal transmission linearity.

RESULTS AND DISCUSSION

The results of the study showed that the introduction of an amplifier cascade significantly improved the stability of the device's linearity when the load changed. As can be seen from Table 1, Circuit 1 shows a significant increase in linearity deviation with a decrease in load resistance, reaching almost 700 μV at a load of 1 kOhm. This indicates that at low resistance values, the output cascade without an amplifier loses its ability to maintain stable transfer characteristics, which can lead to significant signal distortion and reduced data transmission accuracy. At the same time, circuit 2, which uses a push-pull amplifier stage, demonstrates significantly better stability, maintaining a low level of linearity deviation even with a significant decrease in load resistance. The deviation index for this circuit varies only from 12.32 μV to 13.63 μV, which indicates the high efficiency of this solution for ensuring high signal transmission accuracy. This confirmed the feasibility of using amplifier stages to improve the load capacity of buffer devices, especially when working with low-impedance loads, which is critical for high-speed and high-line analogue circuits.

Table 1. Dependence of linearity deviation on load resistance

Load resistance, kOhm		None	10	5	2	1
Linearity deviation, μV	Circuit 1	12.47	40.15	80.71	256.59	697.93
	Circuit 2	12.32	12.58	12.63	13.02	13.63

Source: created by the authors

The graphs in Figure 4 show the dependence of transmission characteristic deviations for circuits without an amplifier cascade (a) and with an amplifier cascade (b). As can be seen, the introduction of an amplifier cascade significantly stabilised the circuit operation, reducing the level of nonlinear distortions. In the case of a circuit without an amplifier, a significant increase in deviation is observed when the input voltage changes, which can lead to a loss of signal transmission accuracy. At the same time, the circuit with an amplifier stage demonstrates more stable characteristics, maintaining a low level of distortion even with large changes in the input signal. This confirms the feasibility of using amplifier stages to ensure high signal transmission accuracy, especially under variable load conditions.

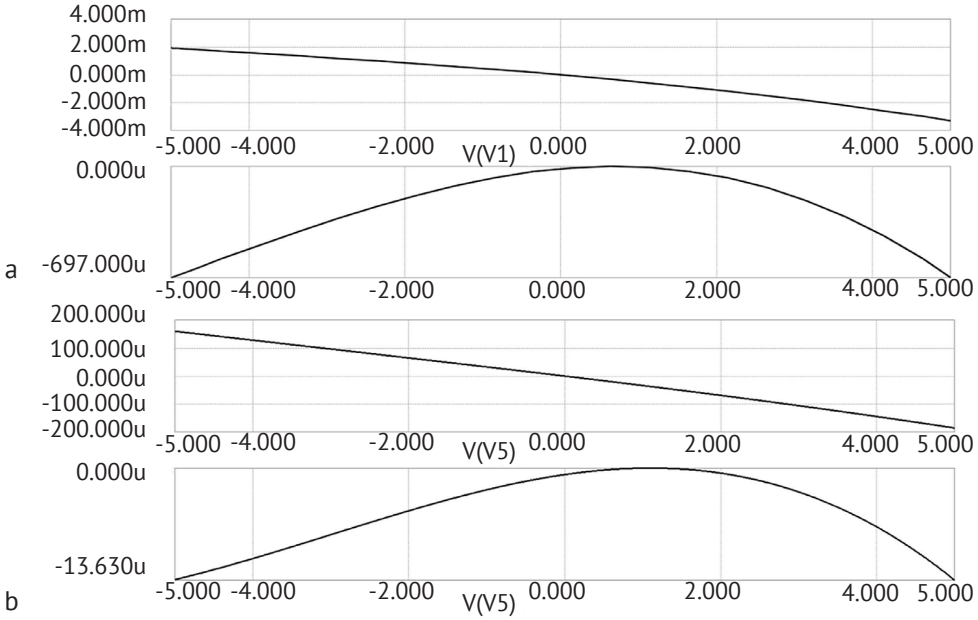

Figure 4. Transmission characteristic deviations

Notes: a – simple buffer device; b – improved buffer device

Source: developed by the authors as a result of modelling in Micro Cap

The amplitude-frequency characteristic was a critically important parameter in the design of voltage buffers, as it determined the device's ability to maintain a stable output signal at different input signal frequencies. In the case of voltage buffers, especially for broadband and high-line models, the amplitude-frequency characteristic is crucial for ensuring efficient operation at high frequencies, minimising distortion and maintaining linearity under heavy loads. One of the main tasks for broadband buffers was to maximise the bandwidth without significantly affecting linearity. This can be achieved by optimising circuit design solutions, such as the introduction of cascode or push-pull topologies, which help to reduce parasitic effects that limit the bandwidth. These approaches make it possible to maintain stable linearity even at high frequencies, which is critical for buffers operating with high loads.

High-linearity voltage buffers must ensure minimal distortion of the input signal over a wide frequency range. The linearity of the amplitude-frequency response is a key parameter because it affects the accuracy of signal reproduction. In voltage buffers, especially in high-precision applications, it is important that the amplitude-frequency response remains linear throughout the entire operating frequency range. This helps to avoid non-linear signal distortions such as harmonic distortion or changes in amplitude. In buffers with low linearity, there was a significant increase in the harmonic distortion coefficient with increasing frequency or input signal voltage. The amplitude-frequency characteristic graphs are shown in Figure 5.

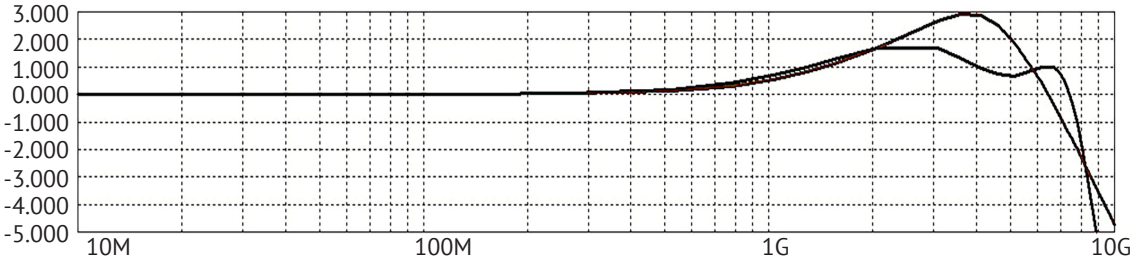

Figure 5. Amplitude-frequency characteristics of a simple and improved buffer device
Source: developed by the authors as a result of modelling in Micro Cap

The implemented precision buffer amplifier, based on a push-pull symmetrical structure, has the following parameters: input signal range: ± 5V; output current: ± 5 mA; zero offset distortion; scale distortion; linearity distortion. Therefore, optimising the amplitude-frequency characteristics was an important step in creating highly efficient voltage buffers, especially for devices that operate in broadband ranges and require high linearity. Ensuring stable signal transmission with minimal distortion is critical for achieving high accuracy and reliability in analogue circuits. The choice of buffer amplifier topology, such as push-pull symmetrical structures, reduces the level of nonlinear distortion, reduces scale and zero offset distortions, and provides a wide range of operating signals. This makes such solutions optimal for use in high-precision measurement systems, analogue-to-digital conversion devices, and other high-speed analogue applications where signal transmission accuracy, energy efficiency, and stable operation under high loads are critical.

In electronic systems, buffer devices perform an important function by ensuring signal consistency between different circuit components, reducing distortion and ensuring stable operation of both analogue and digital devices. Between 2000 and 2020, significant research was conducted to improve buffer cascades, in particular to improve their linearity, expand their bandwidth, and increase their load capacity. One of the main areas of application for buffer devices is analogue-to-digital and digital-to-analogue converters, which are an integral part of modern electronic systems, performing the function of communication between analogue physical quantities and digital computing systems. The work of P. Horowitz & W. Hill (1981) covered a wide range of issues related to the principles of operation, characteristics and practical application of these devices. The authors emphasised that signal conversion systems are used in measuring instruments, communication devices, image processing and other high-precision fields. The works of B. Cordell (2010) and P.E. Allen & D.R. Holberg (2011) included an analysis of the main types of analogue-to-digital and digital-to-analogue converters, the features of their interaction with microprocessor systems, and described typical application circuits – from simple elements to integrated data acquisition systems. Special attention was paid to testing methods and converter specifications, which are critical for their correct integration into practical solutions. In addition, the work considered specialised forms of converters, such as video converters, devices for synchronous and resolver systems, as well as voltage-to-frequency converters. Much attention was also focused on auxiliary components, in particular analogue reference sources, multiplexers and switches, which are important for improving the accuracy and speed of circuits.

The textbook by T.C. Carusone *et al.* (2011) discussed the principles of designing analogue integrated

circuits, in particular buffer amplifiers. The authors have analysed in detail the methods for achieving high linearity and wide bandwidth, which are key aspects for buffers with increased load capacity. Particular attention was paid to circuit solutions that ensure stable operation under variable load conditions.

The development of integrated circuits has significantly changed traditional approaches to the design of electronic devices, particularly in the field of analogue electronics. In the study by W. Sansen (2007) the key aspects of analogue integrated circuit design were highlighted, focusing on their capabilities, limitations and design philosophy. The author provided practical advice for both analog integrated circuit designers and engineers involved in their integration into complex electronic systems. One of the important points discussed in U. Tietze & C. Schenk (2008) is the impact of integrated circuit technologies on traditional methods of designing analogue devices. The authors emphasised that the development of microprocessors and digital signal processing has led to significant improvements in analogue circuits, especially in areas that actively interact with digital devices. This, in turn, has contributed to the creation of complex integrated systems that combine analogue and digital functions on a single chip, ensuring high performance and signal processing accuracy. The study also highlighted the basic principles of analogue integrated circuit design, including methods for reducing distortion and improving linearity and stability. The technological limitations of analogue microcircuit manufacturing were considered separately, as well as the trade-offs between performance, power consumption and chip area.

The book by A.H.M. Roermund *et al*. (2010) contained a collection of modern approaches to the design of analogue circuits, including buffer cascades for high-speed converters. Methods for reducing nonlinear distortions and improving energy efficiency were discussed, which are relevant for the development of buffers with high load capacity. Another key issue in electronic circuit design is reducing delay and power consumption, especially under conditions of heavy load on the output stages. One effective approach to solving this problem is to use bipolar complementary metal-oxide semiconductor buffers, which combine the advantages of bipolar and complementary metal-oxide semiconductor technologies, providing high speed and efficient signal transmission at low supply voltages. The article by M. Kaviani *et al*. (2016) presented bipolar complementary metal-oxide semiconductor buffers operating in the voltage range of 0.8-1.5 V and demonstrating efficient charging and discharging of load capacitors with reduced delay. According to the simulation results, such buffers operate effectively over a wide range of capacitances (0.5-200 pF), making them suitable for circuits with different load levels. The proposed circuits reduced delay by 88% compared to traditional complementary metal-oxide semiconductor buffers at high load capacitance values. It has also been established that bipolar complementary metal-oxide semiconductor buffers have a higher operating frequency range compared to complementary metal-oxide-semiconductor (CMOS) implementations: when the capacitance exceeds 200 pF, the operating frequency of complementary metal-oxide semiconductor buffers decreases by 25% compared to bipolar complementary metal-oxide semiconductor analogues.

The textbook by B. Razavi (2016) provided a detailed description of the design of analogue CMOS integrated circuits, in particular buffer amplifiers. The author analysed methods of linearisation and distortion compensation, which are important for achieving high accuracy and stability of buffers under variable loads. R.J. Baker's (2010) book covered the practical aspects of designing, tracing, and modelling CMOS circuits. It discussed various buffer amplifier topologies, methods for reducing output resistance and improving linearity, which are critical for ensuring high load capacity and wide bandwidth. The publication by M. Fakhfakh *et al*. (2013) presented a systematic approach to the design of analogue and mixed-signal circuits. Particular attention was paid to methods for synthesising buffer amplifiers using only voltage and current repeaters, which contributed to achieving high linearity and wide bandwidth.

One of the directions in the development of broadband high-linearity voltage buffers was to reduce power consumption and increase speed at low operating voltages. A promising solution is the use of heterojunction tunnel transistors, which have a lower subthreshold slope compared to standard metal-oxide-semiconductor field-effect transistors, allowing for significant improvements in energy efficiency even at low power levels (Sharma & Mehra, 2011; Park & Yoo, 2013). The main features of such transistors are asymmetric current and increased Miller capacitance, which creates additional challenges for the design of high-speed circuits. However, the use of heterojunction tunnel transistors in ring generators can significantly reduce dynamic power consumption, and the use of heterojunction tunnel transistors in memory cells can significantly reduce static power losses. These advantages make heterojunction tunnel transistors promising for broadband analogue circuits with high load capacity, where high energy efficiency, low power loss and operational stability are critical.

The book by P. Mak & R.P. Martins (2012) focused on techniques for designing analogue and radio frequency (RF) electronic circuits for nanometre CMOS technologies. Methods for implementing buffer cascades capable of operating at high and mixed voltages were considered, which allowed for high linearity and efficiency under significant loads. Thus, the use of bipolar complementary metal-oxide semiconductor buffers significantly improves the speed of output cascades, reduces delay, and ensures stable operation of circuits at high load levels, making them a promising solution for use

in high-speed analogue-to-digital and digital-to-analogue systems. A review of the literature has shown that the creation of broadband high-linearity buffer devices is an important and relevant task in modern electronics. Research has shown significant progress in improving linearity, expanding bandwidth, and increasing the load capacity of buffer cascades, which is critical for use in high-precision measurement systems, communication devices, and signal processing.

Works related to traditional buffer cascades based on operational amplifiers demonstrated their effectiveness in low-frequency circuits, but bandwidth limitations restrict their use in high-speed conditions. Current research was focused on the development of push-pull high-line structures with parametric compensation, which significantly reduces distortion and improves the metrological characteristics of devices. In addition, the significant impact of digital signal processing and integrated circuits on analogue devices should be noted. Advances in microprocessor technologies have contributed to the improvement of analogue integrated circuit design methods, enabling the creation of complex multifunctional devices that can integrate analogue and digital blocks on a single chip.

CONCLUSIONS

The study considered the proposed approach to constructing a broadband high-linearity voltage buffer with enhanced load-driving capacity, which allows for effective matching of low-power signals with low-impedance loads. The proposed circuit solution was based on the use of cascode and push-pull topologies, which helps to reduce nonlinear distortions, improve stability, and expand the device's bandwidth. A buffer device circuit has been created that meets the increased requirements for load capacity and voltage transfer coefficient. The proposed solution, based on a push-pull symmetrical structure, provides a wide input signal range of ± 5 V and an output current of up to ± 5 mA. During modelling, a low level of zero offset deviation (no more than 1 mV), scale deviation (0.0001%) and linearity deviation (0.000005%) was achieved, which indicates high signal transmission accuracy.

A comprehensive analysis of the static and dynamic characteristics of the buffer cascade was carried out using modern simulation software such as Micro Cap. The influence of transistor element parameters on the stability of the device was investigated, in particular the influence of load resistance on signal transmission linearity. As shown in Table 1, the circuit with an amplifier cascade demonstrates a consistently low level of linearity error even with significant reductions in load resistance, ranging from 12.32 μV to 13.63 μV, whereas without an amplifier, this figure increases sharply to 697.93 μV at a load of 1 kOhm. This improvement is critical for stable device operation in high-load applications.

A comparison of the proposed solution with traditional buffer cascades demonstrated its significant advantages, including reduced harmonic distortion, increased signal transmission accuracy, and significantly better load capacity. This allows the developed solution to be used in high-precision measurement systems, analogue-to-digital converters and amplifiers, where high accuracy and minimal power loss are critical. This is a physical parameter that characterises the effect of temperature on the operation of semiconductor components, in particular transistors. It plays a key role in the formation of current in the p-n junctions of transistors and determines the dependence of the electrical characteristics of the buffer on temperature. This makes the developed device promising for use in high-precision measurement systems, analogue and mixed signal processing circuits, as well as in powerful drivers for controlling low-impedance loads.

The proposed approach to the construction of buffer devices can be applied in analogue-to-digital and digital-to-analogue conversion systems, where high linearity, wide bandwidth and minimal non-linear distortion are critical. Possible areas of application for the developed device have been identified, in particular in high-precision measurement systems, analogue signal processing circuits, and powerful drivers for load control. Further research should be aimed at optimising the design of buffer devices to achieve high linearity, wide bandwidth and low output resistance, which will ensure their effective integration into modern electronic systems.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Allen, P.E., & Holberg, D.R. (2011). *CMOS analog circuit design*. Oxford: Oxford University Press.
[2] Azarov, O.D., & Stahov, O.Y. (2022). Push-pull voltage buffer devices on bipolar transistors. *Herald of Khmelnytskyi National University: Technical Sciences*, 1(4), 18-22. doi: 10.31891/2307-5732-2022-311-4-18-22.
[3] Baker, R.J. (2010). *CMOS: Circuit design, layout, and simulation*. Piscataway: IEEE Press. doi:10.1002/9780470891179.
[4] Carusone, T.C., Johns, D.A., & Martin, K.W. (2011). *Analog integrated circuit design* (2nd ed.). Hoboken: John Wiley & Sons, Inc.
[5] Chang, A., & Kong, B. (2020). High-speed rail-to-rail class-ab buffer amplifier with compact, adaptive biasing for FPD applications. *Electronics*, 9(12), article number 2018. doi: 10.3390/electronics9122018.

[6] Cordell, B. (2010). *Designing audio power amplifiers*. New York: McGraw-Hill Education.
[7] Fakhfakh, M., Tlelo-Cuautle, E., & Castro-López, R. (2013). *Analog/RF and mixed-signal circuit systematic design*. Berlin: Springer. doi: 10.1007/978-3-642-36329-0.
[8] Horowitz, P., & Hill, W. (1981). The art of electronics. *Physics*, 34(3), 67-71. doi: 10.1063/1.2914480.
[9] Kaviani, M., Sharifi, H., Dolatshahi, M., & Navi, K. (2016). Design of low voltage and high-speed bicmos buffer for driving large load capacitor. *International Journal of Engineering and Manufacturing*, 6(1), 1-9. doi: 10.5815/ijem.2016.01.01.
[10] Kim, S.K., Son, Y.-S., & Cho, G.H. (2006). Low-power high-slew-rate CMOS buffer amplifier for flat panel display drivers. *Electronics Letters*, 4(42), 214-215. doi: 10.1049/el:20063898.
[11] Liao, J., Li, P., Gao, Z., Zhi, Y., Fang, Z., & Zhuang, X. (2025). A low noise differential charge amplifier with input buffer stage for toroidal gyroscope. *Measurement Science and Technology*, 36(5), article number 055112. doi: 10.1088/1361-6501/add314.
[12] Mak, P., & Martins, R.P. (2012). *High-/mixed-voltage analog and RF circuit techniques for nanoscale CMOS*. New York: Springer. doi: 10.1007/978-1-4419-9539-1.
[13] Mythry, S., & Moni, J. (2019). A bulk-driven, buffer-biased, gain-boosted amplifier for biomedical signal enhancement. *Cogent Engineering*, 6(1), article number 1658966. doi: 10.1080/23311916.2019.1658966.
[14] Park, H., & Yoo, C. (2013). Effect of buffer cache on storage performance in virtualised environment. *Electronics Letters*, 4(49), 251-253. doi: 10.1049/el.2012.4069.
[15] Patent No. 158284. (2025). *Buffer cascade*. Retrieved from https://iprop-ua.com/inv/ec9yoxmd/.
[16] Razavi, B. (2016). *Design of analog CMOS integrated circuits* (2nd ed.). New York: McGraw-Hill Education.
[17] Roermund, A.H.M., Casier, H., & Steyaert M. (2010). *Analog circuit design: Smart data converters, filters on chip, multimode transmitters*. Dordrecht: Springer. doi: 10.1007/978-90-481-3083-2.
[18] Rohde, U.L., & Newkirk, D.P. (2000). *RF/microwave circuit design for wireless communications*. New York: John Wiley & Sons, Inc.
[19] Sansen, W. (2007). *Analog design essentials*. London: Springer. doi: 10.1007/b135984.
[20] Sharma, D., & Mehra, R. (2011). Low power, delay optimized buffer design using 70nm CMOS technology. *International Journal of Computer Applications*, 22(3), 13-18. doi: 10.5120/2565-3526.
[21] Tietze, U., & Schenk, C. (2008). *Electronic circuits: Handbook for design and applications*. Berlin: Springer. doi: 10.1007/978-3-540-78655-9.
[22] Yang, B., Caldwell, T., & Chan Carusone, A. (2025). An energy-efficient pipeline-sar ADC using linearized dynamic amplifiers and input buffer in 22nm FDSOI. *IEEE Open Journal of Circuits and Systems*, 6, 50-62. doi: 10.1109/OJCAS.2024.3509746.

Широкосмуговий високолінійний буфер напруги з підвищеною навантажувальною здатністю

Олексій Азаров

Доктор технічних наук, професор
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0000-0002-8501-1379

Сергій Богомолов

Кандидат технічних наук, доцент
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0009-0003-3823-8380

Олександр Лукашук

Аспірант
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0009-0002-7485-1193

Анотація. Розробка буферних каскадів із високою лінійністю, широкою смугою пропускання та здатністю працювати з низькоомним або ємнісним навантаженням є актуальним завданням сучасної аналогової електроніки. Метою цієї роботи було створення широкосмугового високолінійного буфера напруги з підвищеною навантажувальною здатністю для застосування у високоточних системах. У дослідженні використано метод комп'ютерного моделювання для аналізу електричних параметрів запропонованого каскаду, а також проведено порівняння з традиційними схемами. Запропоновано новий схемотехнічний підхід до побудови буферного каскаду, що базується на оптимізованій структурі вихідного етапу та поєднанні каскодної і двотактної топологій. Така конфігурація дозволяє знизити вихідний опір, зменшити нелінійні викривлення і забезпечити високу стабільність передачі сигналу. Результати моделювання підтвердили, що розроблений буфер має розширену смугу пропускання, покращені частотні характеристики та знижене гармонічне спотворення в широкому діапазоні частот. Зокрема, під час роботи з навантаженнями великої ємності (до 200 пФ) спостерігалося суттєве зниження затримки сигналу (до 80-90 %) у порівнянні зі звичайними комплементарними структурами метал-оксид-напівпровідник реалізаціями. Оцінка впливу параметрів активних елементів на динамічну стабільність роботи також засвідчила високу надійність пристрою в умовах змінного навантаження. Результати дослідження можуть бути використані для покращення продуктивності аналогово-цифрових перетворювачів, драйверів, підсилювачів і схем керування, що працюють у режимах із підвищеним струмовим навантаженням. Практична цінність розробки полягає в можливості її інтеграції у високоточні, енергоефективні та швидкодійні електронні системи, що потребують стабільної роботи буферного каскаду за складних умов навантаження

Ключові слова: високонавантаженість; високоточність; буферний пристрій; каскодний каскад; двотактна структура

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 30 No. 2. 2025

UDC 004.056.53
DOI: 10.62660/bcstu/2.2025.63

Intrusion detection in telecommunications networks using network traffic analysis

Andriy Riy*
Graduate Student
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0009-0005-0252-1533

Maryan Kyryk
Doctor of Technical Sciences, Professor
Lviv Polytechnic National University
79000, 12 Stepan Bandera Str., Lviv, Ukraine
https://orcid.org/0000-0001-9156-9347

Abstract. The growth of cyber threats and network traffic has highlighted the need for effective methods for detecting anomalies. The purpose of the study was to develop a hybrid model for detecting anomalies in telecommunications networks with increased accuracy based on comparative analysis of uncontrolled algorithms. Four main algorithms were compared: Isolation Forest, Local Outlier Factor (LOF), One-Class Support Vector Machine (SVM), and Elliptic Envelope using simulated network traffic data. Analysis methods included data normalisation using Z-score and Min-Max Scaling to eliminate large-scale differences between traits. An ensemble of 100 decision trees was used to improve the accuracy of anomaly detection. As a result, it was found that Isolation Forest provided the highest accuracy of anomaly detection (F1-Score = 85.8%) and high processing speed (processing time per 10,000 records in 2.5 seconds), which is important for real-world conditions of telecommunications networks with large amounts of data. LOF showed high accuracy in detecting local anomalies, but with greater computational complexity. The one-class SVM algorithm detected global anomalies, but showed lower accuracy for local ones. Elliptic Envelope had limited performance due to the assumption of normal data distribution. Additionally, a comprehensive model was developed that combined the advantages of Isolation Forest, LOF, and Density-Based Spatial Clustering of applications with Noise, which allowed increasing the F1-Score to 88% and exceeding the results of individual models. The hybrid model showed adaptability to multimodal distributions and efficiency in detecting local anomalies. Practical significance lies in improving the accuracy and stability of network security systems and form the basis for adaptive real-time algorithms, which allowed cybersecurity and telecommunications specialists to more effectively monitor and protect networks from threats

Keywords: anomaly detection; performance evaluation; scalability; accuracy; unsupervised algorithms; hybrid model; clustering

INTRODUCTION

With the development of digital technologies and the expansion of network interaction, the issue of ensuring reliable protection of information systems has become particularly relevant. The growing number of cyber-attacks, in particular, zero-day attacks and Distributed Denial of Service (DDoS) attacks, creates significant problems for conventional intrusion detection systems, which are mainly based on signature

Article's History: Received: 05.01.2025; Revised: 22.04.2025; Accepted: 16.06.2025.

Suggested Citation:
Riy, A., & Kyryk, M. (2025). Intrusion detection in telecommunications networks using network traffic analysis. *Bulletin of Cherkasy State Technological University*, 30(2), 63-76. doi: 10.62660/bcstu/2.2025.63.

*Corresponding author

methods. However, the limitations of these approaches associated with the inability to effectively respond to new unknown threats necessitate the development of more flexible and intelligent mechanisms for detecting anomalies in network traffic.

Existing research has focused on the use of machine learning to detect anomalies in telecommunications networks. D. Adhikari *et al.* (2024) substantiated the need to create models that can adaptively analyse heterogeneous and large volumes of Internet of Things network traffic, which is critical for preventing new types of attacks. F.-N. Loo *et al.* (2024) emphasised the importance of implementing semi-controlled anomaly detection systems in industrial networks, highlighting their ability to provide protection even in the face of a limited amount of marked data.

Special attention should be paid to studies that raised the issue of improving the accuracy of unsupervised methods (methods that work without data markup). J.I. Iturbe-Araya & H. Rifà-Pous (2025) showed that competent optimisation of hyperparameters can significantly improve anomaly detection in smart home systems, which is directly related to the overall approach of optimising the intrusion detection system in dynamic traffic conditions. In turn, Y. Meidan *et al.* (2023) proposed the concept of collaborative anomaly detection, which demonstrated effectiveness in minimising the number of false-positives in complex network environments.

In the context of identifying new types of attacks, the study by T. Zoppi *et al.* (2023) was important in the context of identifying new types of attacks. They conducted an in-depth comparative analysis of supervised (methods that work on labelled data), unsupervised, and meta-learning methods (models that learn from different tasks and are able to quickly adapt to new types of threats) for intrusion detection tasks. It is unsupervised algorithms and meta-learning algorithms that have a significantly higher ability to detect zero-day attacks (attacks without previously known patterns) in telecommunications systems, since they do not require pre-marked data and can effectively respond to new threat scenarios. This proved the advantage of using unsupervised and meta-learning approaches to effectively detect new and unknown attacks in telecommunications networks.

Developing the topic of applying deep learning to protect critical infrastructure, A. Dairi *et al.* (2023) proposed innovative semi-controlled deep anomaly detection models. Their research focused on the development of adaptive anomaly detection systems for smart grid environments that can learn from limited volumes of normal traffic without the need for large amounts of attack data, which is critical for protecting telecommunications networks and energy infrastructure.

Ukrainian researchers also made a significant contribution to the development of the topic. Y. Chychkarov *et al.* (2023) proposed a hybrid approach that combines machine learning and fuzzy logic techniques to improve the accuracy of detecting cyber-attacks in telecommunications networks. The study contributed to improving the flexibility and adaptability of intrusion detection systems in environments with dynamic and unpredictable traffic, which directly developed the area of network traffic analysis. D. Ilin & I. Starinsky (2023) developed a model for detecting network anomalies based on autoencoders, which does not require a large amount of marked data, which is crucial for rapid response to attacks in complex telecommunications environments. Their contribution was to popularise lightweight and efficient anomaly detection models for real network traffic.

Research by H.I. Haidur *et al.* (2023) demonstrated the promise of using deep neural networks to analyse traffic in high-load telecommunications networks. Their study confirmed the ability to maintain high accuracy of anomaly detection even in conditions of heavy network load, which is important for scalability of security systems. In addition, A. Nikitenko & Y. Bashkov (2024) developed an advanced intrusion detection system architecture that combined the convolutional neural network and the gated recurrent unit using the attention mechanism. The proposed approach significantly improved the accuracy of detecting complex attack patterns in network traffic, strengthening the methodological basis for building advanced intrusion detection systems in telecommunications networks. However, the lack of universality of existing solutions remained an important problem: most approaches are tailored to specific data types or environments and require significant improvements to scale. In addition, there were difficulties in increasing the level of false-positive results when switching models to real conditions.

The purpose of the study was to analyse modern uncontrolled methods for detecting anomalies in telecommunications networks and develop a comprehensive hybrid approach aimed at improving the accuracy of threat identification. To achieve this goal, the following tasks were completed:

1. To analyse the effectiveness of advanced uncontrolled algorithms for detecting anomalies in terms of accuracy and computational efficiency in telecommunications traffic conditions.

2. To develop a hybrid approach that integrates unsupervised algorithms to improve accuracy and adaptability to various types of network traffic.

3. To evaluate the performance of the developed approach based on simulated data to determine its ability to identify cyber threats.

MATERIALS AND METHODS

The study was conducted from February 2025 to April 2025. The main focus was on evaluating the effectiveness of uncontrolled algorithms for detecting deviations in simulated telecommunications traffic, and developing a hybrid model aimed at improving classification accuracy and reducing the number of false-positive positives.

For the experiment, a simulated dataset of 10,000 network records was used, of which 10% were deviations that simulated typical cyber threats: DDoS attacks (increased traffic and simultaneous connections), port scans (short sessions over different ports), and unauthorised access attempts (short connections). The selected deviation rate of 10% was typical for conditions where threats appear rarely, but can have significant consequences for network security. This allowed modelling a realistic scenario for evaluating the effectiveness of algorithms under complex conditions of class imbalance. The set included characteristics that reflected realistic conditions: traffic volume, number of Transmission Control Protocol/User Datagram Protocol Sessions, connection duration, Hypertext Transfer Protocol, File Transfer Protocol, and Secure Shell protocol types. The data was normalised using the Z-score method calculated by equation (1):

$$z = \frac{x-\mu}{\sigma}, \quad (1)$$

where x – attribute value; μ – average sample value; σ – standard deviation.

The use of the normalisation method helped to avoid problems associated with differences in feature scales and improve the stability of models. In addition, the minimax normalisation method was used to bring features to the same scale and apply an ensemble of 100 decision trees to improve the accuracy of anomaly detection and model stability. Four unsupervised algorithms were selected: Isolation Forest, Local Outlier Factor (LOF), One-Class Support Vector Machine (One-Class SVM), and Elliptic Envelope. The choice was driven by their ability to process data without prior labelling and effectively identify deviations in the face of variable telecommunications traffic. The hyperparameters (contamination = 0.1 for Isolation Forest, n_neighbors = 20 for LOF, nu = 0.1 for One-Class SVM) were optimised using a cross-validation procedure to achieve the best classification quality that meets the requirements of real network conditions. Additionally, the density-based spatial clustering of applications with noise (DBSCAN) algorithm is used to cluster multimodal network traffic data to detect anomalies by identifying non-clustered points. The hyperparameters eps (epsilon-neighborhood radius, which determines the maximum distance between points in one cluster) and min_samples (minimum samples – the minimum number of points in the eps for forming the cluster core) were optimised by the distribution of inter-point distances. To increase computational efficiency, dimensionality reduction was applied through Principal Component Analysis and indexing structures.

The data was divided into training (70%) and test (30%) sets. The algorithms were tested using 5-fold cross-validation to ensure the stability of the results and minimise the probability of random variations in the indicators. Performance was evaluated using the metrics Precision, Recall, F1-Score, area under the receiver operating characteristic curve (AUC-ROC), and area under the precision-recall curve (AUC-PR), which enabled a comprehensive analysis of the balance between classification accuracy and the ability to identify deviations in class imbalance conditions. Precision measured the proportion of correctly classified deviations among all anomalous points, Recall determined the proportion of detected anomalies among all real deviations, F1-Score combined Precision and Recall to estimate their balance, AUC-ROC evaluated the model's ability to distinguish between normal and anomalous entries at different thresholds, and AUC was important for evaluating performance in tasks with rare events (anomalies). Each algorithm was evaluated through quantitative and qualitative performance assessments, which allowed identifying their advantages and limitations. The authors also analysed how combining different methods can improve the accuracy and adaptability of the system. Based on this, a hybrid model was developed to improve the accuracy of anomaly detection and reduce the number of false-positive results in the context of applying different approaches to anomaly detection.

The results were analysed using quantitative analysis (calculation of metrics, statistical comparison using the t-test at the significance level $p < 0.05$) and qualitative analysis (assessment of the ability of algorithms to detect both mass and local deviations). The Student's t-test was used to estimate the statistical significance of performance differences between algorithms, and Bonferroni correction was used for multiple comparisons to minimise the probability of errors of the first kind. Confidence intervals (95%) for the F1-Score were used to assess the stability of the results, which allows assessing the reliability of the results achieved and provides additional confirmation of their representativeness in the context of the entire data set. The study was conducted in a Python environment using the scikit-learn library, as it offers a user-friendly interface for implementing machine learning algorithms, supports a wide range of models for detecting anomalies, and has built-in tools for performance evaluation and cross-validation, which ensures reliability and reproducibility of results.

RESULTS

Comparison of the effectiveness of anomaly detection algorithms on simulated telecommunications traffic

The Isolation Forest algorithm is an effective method for detecting anomalies in telecommunications networks, which is based on the principle of isolating data points using random decision trees. The basic idea of the algorithm is that anomalous objects, due to their distance from the bulk of the data, require less partitioning to isolate compared to normal objects. To do this, the algorithm randomly selects features and values for partitioning, forming binary decision trees in

which each data point is isolated by passing through a sequence of splits. The average length of the isolation path for each point was used as an indicator of anomaly: a shorter path indicates a higher probability of anomaly (Sharma *et al.*, 2024).

Isolation Forest demonstrated high efficiency when analysing the data set. The use of an ensemble of 100 decision trees (the standard number for Isolation Forest) provided stable anomaly detection due to sufficient variability in isolation paths without significantly increasing computational complexity. Using contamination (a parameter that determines the expected proportion of anomalies in the data) at 0.1, which corresponded to the proportion of anomalies in the data, contributed to a balanced operation of the model in conditions close to real network traffic. This choice turned out to be effective, because the share of anomalies in such problems is usually 5-15%, and the value of 0.1 allowed minimising false-positive and false-negative results.

Scaling features to the range [0; 1] by minimax normalisation (traffic volume from 0 to 500 Mb/s after normalisation acquired values from 0 to 1) reduced the influence of the difference in parameter scales, which ensured an equal contribution of each feature to the isolation process (data was processed using equation 1). This improved the accuracy of classification, because without normalisation, features with higher values would dominate the division of space. Using an ensemble approach of 100 trees and the contamination = 0.1 parameter, the model effectively detected anomalies such as DDoS attacks and port scanning, and data normalisation positively affected the quality of point isolation, which is confirmed by high performance metrics (Table 1).

Table 1. Analysis of the effectiveness of the Isolation Forest algorithm

Metric	Value
Precision	87.5%
Recall	84.2%
F1-Score	85.8%
AUC-ROC	0.91
AUC-PR	0.88

Source: compiled by the authors

Precision 87.5% provides accurate detection of anomalies in telecommunications networks, minimising the erroneous designation of normal traffic as a threat, which contributes to the smooth operation of services such as media streaming or secure transactions. Recall demonstrates that the model detected 84.2% of the total number of true anomalies. This result confirms the algorithm's ability to prevent threats such as DDoS attacks or unauthorised access, which is key to protecting networks. An F1-Score of 85.8% (harmonic average Precision and Recall) indicates that the model is balanced, making it effective for real-world scenarios without significant false-positive or false-negative results. The AUC-ROC (0.91) and AUC-PR (0.88) metrics evaluate the overall classification quality. The AUC-ROC value of 0.91 demonstrates the model's high ability to distinguish between normal and abnormal traffic regardless of the classification threshold, which exceeds the typical values for uncontrolled algorithms in problems with 10% anomalies. AUC-PR 0.88 confirms the stable performance of the model in conditions of class imbalance, when anomalies are rare, which corresponds to real network traffic. Isolation Forest effectively detects cyber threats, such as DDoS or port scanning, in a simulated dataset, ensuring high accuracy and reliability. However, for environments with high local variability, where anomalies are less pronounced or concentrated in separate clusters, it may be necessary to combine them with other methods that consider local data density.

The LOF algorithm identifies anomalies by comparing the density of each point with the local density calculated based on k (a hyperparameter that sets the size of the local environment to estimate the data density) of the nearest neighbours in the feature space – if the local density of a point is much lower than that of its environment, it is classified as a deviation. To estimate the density, the algorithm calculates the reach and local deviation coefficient, which shows how anomalous a point is compared to its neighbours. This approach is particularly effective for detecting local anomalies, i.e., objects that look normal in a global context, but differ significantly in a local one (rare protocol attacks) that may not be visible in global analysis (Sharma *et al.*, 2024).

LOF showed stable performance when applied to a simulated dataset. The use of the n_neighbors = 20 parameter provided an optimal balance between sensitivity to local anomalies and noise resistance, which contributed to stable threat detection. This parameter was chosen as the standard value used in network traffic analysis tasks, since smaller values (for example, 10) increase sensitivity to noise, and larger values (for example, 30) reduce the ability to detect local deviations. Scaling features to the range [0; 1] by the Min-Max Scaling method (scaling to the range [0; 1]), which avoids the dominance of features with large values over others during model training) eliminated the influence of the scale difference between parameters, which ensured the same contribution of each feature to the calculation of distances between points (data were

obtained using equation 1). This increased the accuracy of classification, because without normalisation, features with higher values would dominate local density calculations. As a result, LOF effectively identified local anomalies, such as unauthorised access attempts, that remained invisible to global methods (Table 2).

Table 2. Performance indicators of the LOF algorithm

Metric	Value
Precision	91.2%
Recall	82.5%
F1-Score	86.6%
AUC-ROC	0.89
AUC-PR	0.85

Source: compiled by the authors

Precision at 91.2% indicates the model's high accuracy in detecting anomalies, which allows telecommunications systems to rarely confuse normal traffic with suspicious traffic, ensuring smooth operation of online services such as video calls or data streaming. Recall in 82.5% detects true anomalies out of the total number, although 17.5% of anomalies, especially in the context of sparse or multicentre data distributions, went unnoticed. F1-Score, which is 86.6%, is a harmonious average between Precision and Recall, confirming the balance of the model, although LOF performance is somewhat inferior to Isolation Forest (F1-Score = 85.8%) due to omissions of rare anomalies. The AUC-ROC index (0.89) indicates a high ability of the model to distinguish between normal and abnormal records at different classification thresholds, which ensures stable operation in real network conditions. AUC-PR at 0.85 highlights the reliability of LOF in class imbalance problems, where anomalies account for only 10% of the data, which is typical for telecommunications traffic. These results confirmed that LOF effectively detects local anomalies, such as rare protocol attacks, providing high accuracy and stability in complex network environments. LOF has shown high efficiency in detecting local anomalies, providing efficient operation for specific scenarios in telecommunications networks. However, to handle large amounts of traffic and detect global anomalies, it is advisable to combine LOF with methods that are better suited for scaling and global analysis, such as Isolation Forest.

The One-Class SVM algorithm is a classic anomaly detection method that works by plotting a boundary (hyperplane) around the bulk of normal data, marking as anomalies those points that fall outside this region. The uniqueness of One-Class SVM is that the model is trained exclusively on normal examples, using the principles of maximising the distance between the hyperplane and the origin in the feature space transformed using kernel functions. This makes it particularly relevant for unsupported or semi-supervised intrusion detection tasks (a method that uses both marked and unallocated data to train the model) in telecommunications networks, where anomalous examples may not be available at the training stage, for example, when detecting zero-day attacks (Ahmad *et al.*, 2023).

One-Class SVM showed moderate performance when analysing the data set. The use of a radial basis core with the parameter $\nu = 0.1$ (a parameter that determines the proportion of anomalies and adjusts the sensitivity of the model, selected according to the proportion of anomalies in the data) allowed the model to efficiently separate nonlinear data distributions, which is typical for network traffic. Scaling features to the same range using minimax normalisation eliminated the effect of scale differences, which improved the quality of classification (data were obtained using equation 1). Pre-cleaning the data from noise and emissions also had a positive impact on the stability of the model. As a result, One-Class SVM detected global anomalies, such as DDoS attacks, but had difficulty identifying local threats, such as unauthorised access attempts (Table 3).

Table 3. Evaluating the effectiveness of One-Class SVM based on simulated data

Metric	Value
Precision	84.3%
Recall	76.1%
F1-Score	79.9%
AUC-ROC	0.86
AUC-PR	0.81

Source: compiled by the authors

Precision (84.3%) shows that only 15.7% of records were misclassified, which is important for telecommunications networks where false positives can block normal traffic. Recall reflects that the model detected 76.1% of true anomalies, missing 23.9% of anomalies, in particular, unauthorised access attempts with abnormally short connection durations, which were often classified as normal due to their proximity to the main data

distribution. F1-Score (79.9%), characterises the relative imbalance between Precision (84.3%) and Recall (76.1%), but remains acceptable for a method that trains exclusively on normal data without knowledge of anomalies. The AUC-ROC (0.86) and AUC-PR (0.81) indicators confirm the stability of classification under conditions of class imbalance, which is typical for real network traffic. However, these values are lower compared to Isolation Forest (F1-Score = 85.8%) and LOF (F1-Score = 86.6%), indicating a lower ability of One-Class SVM to cope with fuzzy deviations that do not have an explicit separation from normal data. The model proved to be effective for global anomalies, such as DDoS attacks, where deviations from the normal distribution are more pronounced, but less successful for local anomalies, such as unauthorised access attempts. Due to difficulties in identifying local threats, it is advisable to combine one-class SVM with algorithms with higher local sensitivity, such as LOF, for more comprehensive analysis.

The Elliptic Envelope algorithm is a statistical method for detecting anomalies based on the assumption of a multivariate normal distribution of normal data. The method evaluates the covariance matrix and data centre by forming an elliptical region (envelope) that covers normal points, and objects outside this ellipse are classified as anomalies. This approach is relevant for network monitoring tasks if traffic features are close to normal distribution, making it suitable for pre-filtering or working with aggregated statistical characteristics such as average packet sizes or request frequency (Sharma *et al.*, 2024).

Elliptic Envelope demonstrated limited performance on a simulated dataset from telecommunications traffic. For preprocessing, Z-score Normalization (feature scaling to mean 0 and standard deviation 1) was applied, which ensured adequate alignment of the scales and allowed correctly estimating the covariance matrix required for constructing the elliptical boundary (data were obtained using equation 1). The model was configured so that 80% of the most typical observations were included in the boundary construction (support_fraction = 0.8 parameter, i.e., only the part of the data that best corresponds to the normal distribution was applied when constructing the boundary, which increased the resistance to emissions). However, the key assumption of the algorithm (multivariate normal feature distribution) turned out to be too restrictive for real network traffic, which often has a multimodal or asymmetric structure. During DDoS attacks, there was a significant variability in the number of sessions and traffic volume, which did not correspond to the normality hypothesis. As a result, Elliptic Envelope effectively identified anomalies with well-defined deviations, such as unauthorised access attempts with an atypically short connection duration, but failed to cope with more complex patterns, in particular, port scanning, where deviations were less pronounced and did not correspond to the expected distribution form (Table 4).

Table 4. Anomaly detection metrics using Elliptic Envelope

Metric	Value
Precision	78.2%
Recall	70.4%
F1-Score	74.1%
AUC-ROC	0.79
AUC-PR	0.76

Source: compiled by the authors

With Precision at 78.2%, the algorithm incorrectly marked 21.8% of normal traffic as anomalies, which can lead to blocking some legitimate requests (interrupting video streaming or slowing down access to web services), which worsens the user experience. Recall notes that the model detected only 70.4% of true anomalies, missing 29.6% of deviations, including complex port scans and atypical DDoS attack patterns, which is a critical flaw for intrusion detection systems. F1-Score (74.1%) is a harmonic mean between Precision (78.2%) and Recall (70.4%) and reflects the limited balance of the model, inferior to Isolation Forest and LOF in problems with multimodal data distributions. An AUC-ROC score of 0.79 indicates a moderate ability of the model to distinguish between normal and abnormal records, which is significantly lower than Isolation Forest (AUC-ROC = 0.92) and LOF (AUC-ROC = 0.89), indicating a deterioration in accuracy when the classification threshold changes. The 0.76 AUC-PR is the lowest among the algorithms tested and highlights the limited performance of the Elliptic Envelope in a class imbalance environment. The results of the study confirmed that the effectiveness of the algorithm is limited by the assumption of multivariate normal data distribution, which is rarely performed in real telecommunications traffic. The model is good at detecting only simple anomalies (unauthorised access attempts), but skips more complex patterns due to multimodality and data asymmetry. It is advisable to use Elliptic Envelope for pre-filtering in combination with more flexible methods, such as Isolation Forest or LOF.

Performance analysis and comparative evaluation of basic anomaly detection algorithms followed by flexible hybrid architecture design

Analysis of the performance of the Isolation Forest, LOF, One-Class SVM, and Elliptic Envelope models showed

their differences in the effectiveness of anomaly detection in telecommunications networks. Each model has unique characteristics that affect its ability to detect intrusions depending on the type of anomaly, computing resources, and data features. To effectively detect anomalies in telecommunications networks, it is important to evaluate the strengths and weaknesses of each algorithm to determine their applicability in real-world conditions. Table 5 summarises the comparison of algorithms for their key capabilities and limitations.

Table 5. Advantages and limitations of algorithms in telecommunications systems

Algorithm	Advantages	Disadvantages
Isolation Forest	Flexibility for large networks: easily handles millions of traffic records (for DDoS detection); minimal data preparation: works without complex attribute configuration; noise tolerance: ignores minor data deviations	Skipping rare anomalies: does not notice subtle attacks like password cracking attempts; risk of errors due to settings: the wrong share of anomalies can block normal traffic; difficulty with atypical data: it does not work well if traffic is very diverse
LOF	Accuracy for hidden threats: finds anomalies such as short access attempts that are invisible to other methods; flexible adjustment: easily change the sensitivity due to the number of neighbours; working with data groups: analyses traffic well with clear clusters	Long processing: cannot handle large networks due to slow calculations; noise issues: small traffic fluctuations can distort the results; weakness for mass attacks: misses large anomalies, such as DDoS, by focusing on local changes
One-Class SVM	Detection of new attacks: finds unknown threats, such as new types of hacker attacks; accuracy for complex traffic: works well with non-linear data, such as mixed HyperText Transfer Protocol/File Transfer Protocol; easy integration: easily added to existing security systems	Slow operation: not suitable for fast analysis of large traffic flows; complex setup: neccessary to choose the exact parameters, otherwise there are a lot of errors; noise vulnerability: small deviations in traffic reduce accuracy
Elliptic Envelope	Fast analysis: instantly processes traffic for pre-verification; easy visualisation: shows anomalies as points outside the ellipse, which is convenient for analysis; performance for simple data: good at detecting anomalies in stable traffic, such as short connections	It does not work with real traffic: it cannot handle various or unstable data; skip complex attacks: does not notice subtle anomalies like port scanning; problems without preparation: requires perfect data normalisation, otherwise inefficient

Source: compiled by the authors based on A.K. Shukla *et al.* (2023), N. Sharma *et al.* (2024), E.F. Agyemang (2024), F. Gutierrez-Portela *et al.* (2024), L. Diana *et al.* (2025), U.C. Akuthota & L. Bhargava (2025)

To better understand the advantages and disadvantages of each model, a comparative analysis was performed on key criteria such as anomaly type, scalability, data resistance, and real-world applicability of telecommunications systems. In Table 6, the optimal use cases for each model were systematised and determined.

Table 6. Comparative analysis of anomaly detection models in telecommunications networks

Model	Effectiveness by type of anomaly	Computational complexity and scalability	Resistance to data features	Applicability in telecommunications networks
Isolation Forest	Better detects global anomalies by isolating points; less effective for local anomalies	High processing speed due to tree independence; scales well on large amounts of data	Resistant to high dimension, but sensitive to incorrect setting of the contamination parameter	High: suitable for handling large real-time traffic flows due to its speed and markup independence
LOF	Effective for local anomalies due to density analysis; less accurate for global deviations	High computational complexity due to distance calculation; it does not scale well on large sets	Sensitive to parameter k and noise; works better with clustered data	Moderate: useful for local anomalies (such as rare attacks), but limited scalability for large networks
One-Class SVM	Separates normal data well, but handles fuzzy anomalies worse	Quadratic difficulty; poorly scaled on large data sets	Sensitive to hyperparameters and noise; prone to overtraining	Average: suitable for zero-day attacks, but requires optimisation for large traffic volumes
Elliptic Envelope	Effective only for data with a normal distribution; poorly detects complex and local anomalies	Low complexity; scales well due to a statistical approach	Strongly depends on the normality of the data; unstable to multimodal distributions	Low: limited for pre-filtering, but inefficient for complex network traffic

Source: compiled by the authors based on N. Sharma *et al.* (2024), F. Gutierrez-Portela *et al.* (2024), E.F. Agyemang (2024)

Comparative analysis has shown that Isolation Forest is the most versatile solution for telecommunications networks due to its high processing speed and ability to work with large amounts of data, which is critical for detecting global anomalies in real time. LOF is effective in specific scenarios with local anomalies, but it is limited by low scalability. One-Class SVM can be useful for detecting new types of attacks, but requires additional optimisation for large traffic flows. Elliptic Envelope was the least suitable due to its strong dependence on normal data distribution, which rarely corresponds to real network traffic. The results highlighted the need to develop a hybrid model that could combine the strengths of Isolation Forest (speed and scalability) and LOF (sensitivity to local anomalies), and consider the adaptability to multimodal data distributions inherent in telecommunications networks. This approach will improve the effectiveness of intrusion detection, in particular, for complex and fuzzy anomalies that remain a problem for the methods under study.

The proposed hybrid model consists of three main components. The first component is the Isolation Forest module for global analysis, which uses an ensemble of 100 trees (contamination = 0.1) to quickly detect global anomalies (for example, DDoS attacks) and pre-filter traffic. The second component is the LOF module (n_neighbors = 20), which is applied to selected Isolation Forest points for detailed analysis of local anomalies (for example, rare protocol attacks). The third component is an adaptive clustering module based on the DBSCAN algorithm (noise-based spatial density clustering), which identifies non-clustered points as potential anomalies, which is effective for processing multimodal data distributions characteristic of network traffic with various patterns of normal behaviour and attacks (rare protocol attacks). DBSCAN is used as the final component for detecting anomalies in network traffic after global analysis using Isolation Forest and local analysis via LOF. The key parameters of DBSCAN eps and min_samples were dynamically configured based on statistical characteristics of the data, such as distance distribution or density. This adaptive approach is scientifically sound because it allows the algorithm to consider local data features, which is crucial for detecting anomalies in heterogeneous data sets.

The eps parameter determines the sensitivity to local density. Too small values lead to cluster fragmentation and an increase in false-positive results, while inflated values reduce the ability to detect rare anomalies due to the combination of different patterns. The min_samples parameter sets a density threshold for cluster development. Low values increase sensitivity, but contribute to the development of unstable clusters. High values make the algorithm stricter by increasing the number of points marked as noise. Dynamic adjustment of parameters based on local characteristics, in particular, through analysis of k-nearest neighbours or average density, ensures adaptability to changes in traffic. However, the efficiency of DBSCAN depends on the quality of input data and is accompanied by high computational costs, which limits its use in real time. DBSCAN is sensitive to noise and abnormal features in data, which can also distort clustering in network traffic conditions. In addition, its computational complexity (without optimisation) makes it difficult to process large amounts of real-time data (Sharma *et al.*, 2024).

To improve efficiency, the model uses indexing structures such as K-d trees, which speeds up the search for neighbouring points under favourable conditions, and preliminary dimensionality reduction through Principal Component Analysis with key feature retention control. Dynamic adjustment of eps and min_samples parameters were based on the distribution of inter-point distances and sample size, and the k-d tree index structures and Principal Component Analysis were tested on hybrid model data, which confirms their effectiveness. These measures, which were validated on test datasets, ensure DBSCAN adaptability and synergy with Isolation Forest and LOF in the hybrid model. DBSCAN complements Isolation Forest, which reduces data volume, and LOF, which clarifies local anomalies, allowing DBSCAN to focus on clustering and identifying isolated points. This interaction increases the accuracy of detecting complex attacks, and clear parameter settings and proven optimisations guarantee the reliability and transparency of the model.

The hybrid model's operation algorithm includes five stages. At the first stage, data preprocessing is performed when incoming network traffic is normalised using Min-Max Scaling. In the second step, DBSCAN clusters data, isolating clusters and noise. In the third step, Isolation Forest is applied to each cluster and non-clustered points, identifying points with a high anomalous rating. In the fourth step, LOF clarifies the status of selected points by analysing the local density. In the fifth stage, the results of both modules are aggregated using weighted voting (Isolation Forest – weight 0.6, LOF – weight 0.4), determining the final anomaly rating for each point. The hybrid model has a number of expected advantages. The combination of global and local analysis improves the accuracy of detecting both large-scale and local anomalies. Using DBSCAN provided adaptability to multimodal distributions, which is typical for real network traffic. Based on Isolation Forest, the model maintains a high processing speed, and applying LOF only to a subset of data reduces the computational load. The model is flexible due to dynamic configuration of DBSCAN parameters.

However, the model has potential problems. Setting parameters (contamination, n_neighbors, eps, min_samples) required careful optimisation, for which it is suggested to use cross-validation based on F1-Score. The computational complexity of LOF and DBSCAN can be reduced by parallel cluster processing. To detect fuzzy anomalies in future studies, it is advisable to integrate

deep learning techniques such as autoencoders. The next step in developing the model was to test it experimentally on real-world datasets such as NSL-KDD (NSL Knowledge Discovery in Databases) or CICIDS2017 (Canadian Institute for Cybersecurity Intrusion Detection System, 2017), followed by performance comparisons with individual Isolation Forest and LOF models. This will allow quantifying improvements in the accuracy and scalability of the model in the context of telecommunications networks. Table 7 shows a comparison of the predicted characteristics of the hybrid model with the results of other algorithms by key metrics.

Table 7. Comparative analysis of the hybrid approach and standard anomaly models

Model	F1-Score (%)	Processing time (10,000 records, s)	Detection of local anomalies	Adaptability to multimodal distributions
Isolation Forest	85.8%	2.5	Low	Low
LOF	86.6%	15	High	Average
One-Class SVM	79.9%	20	Average	Low
Elliptic Envelope	74.1%	1.8	Low	Very low
Hybrid model	88%	4	High	High

Source: compiled by the authors

A comparative analysis of the hybrid approach and standard models demonstrated their performance with F1-Score metrics and processing times of 10,000 records, and their ability to detect local anomalies and adaptability to multimodal data distributions. The hybrid model achieved the highest F1-Score (88%) with a processing time of 4 seconds, surpassing the standard Isolation Forest (F1-Score = 85.8%, 2.5 s), LOF (F1-Score = 86.6%, 15 s), One-Class SVM (F1-Score = 79.9%, 20 s), and Elliptic Envelope (F1-Score = 74.1%, 1.8 s) methods. Elliptic Envelope and Isolation Forest process data the fastest (1.8 s and 2.5 sec, respectively), but their lower F1-Score (74.1% and 85.8%) reflects limited ability to detect local anomalies (low for both) and adaptability to multimodal distributions. LOF shows a high ability to detect local anomalies (F1-Score = 86.6%) and average adaptability, but its processing time (15 seconds) was much higher. One-class SVM has an average capacity for local anomalies and low adaptability, which makes it less efficient in complex network scenarios.

The hybrid model, which combines Isolation Forest, LOF and DBSCAN, provides an optimal balance with an F1-Score of 88.0%, reflecting high accuracy in detecting global and local anomalies, and high adaptability to multimodal distributions for efficient processing of complex network data. The processing time of 4 seconds, although longer than that of Elliptic Envelope (1.8 seconds) or Isolation Forest (2.5 seconds), is an acceptable trade-off for telecommunications systems where accuracy in detecting threats such as DDoS attacks or port scanning is a priority. The performance of the hybrid model is evaluated by integrating the strengths of its components, where Isolation Forest provides fast pre-filtering (2.5 s), LOF helps to detect local anomalies (F1-Score = 86.6%), and DBSCAN increases adaptability to multimodal distributions. Processing time is optimised due to DBSCAN parallel clustering and Isolation Forest preprocessing. These results confirmed the potential of the hybrid model as a universal solution for detecting intrusions in telecommunications networks, capable of adapting to various scenarios with high accuracy. Thus, the developed hybrid model combined the key advantages of leading anomaly detection methods, demonstrating improved predictive characteristics, high adaptability and processing efficiency in real telecommunications traffic conditions. Its application can significantly increase the level of security of information systems, and serve as a basis for further research and implementation of intelligent intrusion detection systems in critical infrastructures.

DISCUSSION

The results of the simulation of the hybrid anomaly detection system, which combines Isolation Forest, LOF and DBSCAN clustering algorithms, showed a significant increase in the accuracy of threat detection (F1-Score = 0.88) compared to the use of separate models. This indicates the effectiveness of the integrated approach in conditions of high variability and multimodality of telecommunications traffic. An attempt to implement such integration was observed in the HYbrid and Robust Intrusion DEtection (HYRIDE) model proposed by S. Srivastav *et al.* (2025), which was also based on the idea of hybridisation methods. However, the key difference is that HYRIDE does not have a pre-clustering component, which limits its flexibility in working with spatially heterogeneous data. The use of DBSCAN in this model helped to better adapt to multimodal traffic distributions, reducing the impact of excessive noise and unstructured deviations.

In the study by A. Russell-Gilbert *et al.* (2025), the Retrieval-Augmented Anomaly Detection using Large Language Models (RAAD-LLM) model implemented threat detection by integrating large language models with Retrieval-Augmented Generation. The general concept of adaptation to new threats correlated with the results of the study, but the implementation differed fundamentally. In RAAD-LLM, adaptation was achieved by semantic enrichment of queries, where as

in the presented hybrid architecture, through structural processing of the feature space. Thus, the similarity was manifested at the level of the functional task, while the discrepancy was in the choice of tools. In addition, M.A. Akif *et al*. (2025), who examined the combination of several classical voting-based models, partially coincided with the current study, because both models were built as hybrid and used a consistent solution of several algorithms. Simultaneously, unlike the implemented architecture in this study, the researchers' approach did not involve step-by-step filtering or structural decomposition of input data, since all models work on a single flow.

Equally significant was the focus on the heterogeneity of network traffic highlighted by K.-C. Lu *et al*. (2024), where intrusion detection in Supervisory Control and Data Acquisition Systems was investigated. The researchers noted the critical role of adaptive mechanisms for energy monitoring systems, which, like telecommunications networks, are characterised by complex, multi-layered flows. Although the model was focused on a narrow infrastructure sector, in both cases, the need for algorithms that work efficiently with multimodal data was identified. The system built as part of this study confirmed its ability to adapt to this by combining cluster and outlier analysis.

Another aspect was the automatic configuration of models without labels, which was considered by O. Anser *et al*. (2024). Their AutoML approach allowed creating detectors based on "clean" (unpacked) data, reducing dependence on pre-known attack patterns. A similar problem was solved in the current model by dynamically optimising DBSCAN parameters (eps, min_samples) and controlling contamination in Isolation Forest. The difference lies in the way it is implemented – while AutoML works as a "black box", the approach of this study provided greater transparency and control over the detection logic.

G.M. Firdaus & V. Suryani (2024) proposed an approach to improving DBSCAN clustering by pre-reducing the sample size. Their methodology was aimed at reducing the computing load, which was especially important in conditions of large traffic flows. In the developed model of this study, a similar effect was achieved by combining with Isolation Forest, which cuts off anomalous or irrelevant points before clustering is performed. This automated filter reduced the processing time to 4 seconds per 10,000 records, which is significantly lower than that of LOF or One-Class SVM.

The search for more error-resistant anomaly detection methods was shown in the study by S. Cohen *et al*. (2023), who proposed a Test-Time Augmentation model for Neural-network-based Anomaly Detection based on the test-time augmentation mechanism. Their approach involved generating input data variations directly during testing, which helped to improve classification accuracy in conditions of unstable traffic. In this study, a similar effect was achieved in another way, which included structural interaction between global analysis (Isolation Forest), local analysis (LOF), and cluster module (DBSCAN). Although the technical means differ significantly, both approaches have demonstrated alternative solutions to the same problem, namely, the detection of weakly expressed, mixed, or unstable anomalies in the network environment.

T.J. Hassan *et al*. (2024) in their study focused on modelling False Data Injection Attacks in distributed generation systems. Their proposed attack scenarios showed how easily anomalies can disguise themselves as normal traffic, which made them difficult to detect using conventional methods. Despite the difference in the subject area, this problem is extremely relevant in telecommunications networks. It is due to the combination of DBSCAN and LOF within the hybrid model that it was possible to achieve resistance to such hidden structures, especially after the preliminary isolation of anomalous points through the Isolation Forest module.

The study by M.A. Uddin *et al*. (2024) and M. Ahmed *et al*. (2025), dedicated to detecting fuzzy and zero-day attacks, demonstrate the effectiveness of approaches based on deep neural networks. Autoencoders and hybrid recurrent convolutional models can work successfully without labelled samples. However, the proposed model in this study had significantly lower computational costs, which was critical in real time. In addition, the compared results (F1 = 88%) showed that classical algorithms, properly combined, can compete with more complex architectures.

Efforts to reduce the number of false-positive results were traced in the study by H. AL-Husseini *et al*. (2024), where Long Short-Term Memory and evolutionary algorithms such as the Whale Optimisation Algorithm were combined to fine-tune hyperparameters. This approach provided flexibility, but required significant resources at the training stage. Instead, the hybrid architecture presented in the current study achieved a similar balance between Precision and Recall through weighted voting and parallel configuration of components without using complex optimisers.

The problems of trust in models under unfavourable conditions were investigated by H. He *et al*. (2025), who focused on countering data poisoning during training. Their approach, based on the use of marginal time samples, actualised the risk of latent influence on model parameters. Although the proposed hybrid system did not contain a direct mechanism for countering such attacks, its multi-level structure, with independent decision-making at each stage, potentially increases resistance to distorted or atypical inputs.

C.-W. Tsai *et al*. (2024) proposed a solution for detecting power theft in smart grid, which was based on combining several weak models into a single ensemble. Despite the applied focus on a narrow domain, their architecture showed similar structural features to that developed by the results of this study: both systems relied on the idea of complementarity of

models. However, the proposed hybrid model went beyond specialised applications and demonstrated the potential for scaling in conditions of heterogeneous traffic and a wide range of attacks.

C. Rajathi & P. Rukmani (2025) implemented combining models with different natures (parametric and nonparametric), which was consistent with the idea of combining the approaches of the current study. Their system did not include a clustering level and did not implement adaptive hyperparameter selection for structuring multimodal space, which fundamentally distinguished it from the approach implemented in this study. The similarity can be traced in the very attempt to consider the complementarity of models, while the discrepancy lies in the level of flexibility and scalability of the solution.

Consequently, the hybrid model demonstrated high efficiency and adaptability, and improved resistance to changes in the network environment, which is important for ensuring the security of telecommunications networks. The use of a combination of Isolation Forest, LOF, and DBSCAN helped to accurately detect anomalies even in difficult and variable telecommunications traffic conditions. The results of the study confirmed that the hybrid approach allows working effectively with large amounts of data and reducing the level of false positives, which is critical for real-world operating conditions. The technical implementation of this approach provides high scalability and processing speed, which makes the model suitable for use in real networks. The proposed model has the advantage of processing multimodal data and adapting to variable conditions, which makes it promising for use in intelligent security systems of telecommunications networks.

CONCLUSIONS

As a result of the study, the effectiveness of the hybrid anomaly detection model for telecommunications networks was confirmed. The combination of Isolation Forest, LOF, and DBSCAN clustering algorithms significantly improved the accuracy of threat detection (F1-Score = 88%), which was an important achievement compared to using separate models. On a simulated dataset of 10,000 records, where 10% were anomalies (DDoS attacks, port scanning, unauthorised access), the model achieved an F1-Score of 88%, surpassing individual algorithms Isolation Forest (85.8%), LOF (86.6%), One-Class SVM (79.9%), and Elliptic Envelope (74.1%). The processing time was 4 seconds, which provided a speed close to Isolation Forest (2.5 seconds), but better than LOF (15 seconds) and One-Class SVM (20 seconds). The model showed a high ability to distinguish between normal and abnormal traffic with a class imbalance. The hybrid model approach worked effectively in conditions of high variability and multimodality of network traffic, which is typical for real telecommunications environments.

The hybrid model optimised threat detection through a combination of fast global filtering (Isolation Forest, 100 trees, contamination = 0.1), sensitive local analysis (LOF, n_neighbors = 20), and adaptive clustering (DBSCAN, dynamic eps, and min_samples). The use of Min-Max Scaling, k-d trees, and Principal Component Analysis reduced computing load, providing scalability for large networks. Weighted voting (Isolation Forest 0.6, LOF 0.4) has improved accuracy for complex anomalies, such as hidden protocol attacks that often miss individual methods. The practical significance of the model lays in its ability to increase the resistance of telecommunications systems to cyber threats, reducing false positive responses, and ensuring efficiency. Compared to HYRIDE-type models, it adapts better to heterogeneous data thanks to DBSCAN clustering, which allows efficient processing of high-dimensional traffic.

Limitations of the study were related to the lack of experimental verification in real networks (NSL-KDD or CICIDS2017), which could provide more accurate data for practical applications. It is also worth noting that a deep study of the impact of external variables, such as dynamic changes in network traffic and new threats, required additional research and adaptation of algorithms. The impact of dynamic traffic changes and new threats, such as zero-day attacks, requires further study. Future research should focus on automatic parameter optimisation via AutoML, integrating deep neural networks such as autoencoders to increase sensitivity to unknown attacks, and real-time parallel processing.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Adhikari, D., Jiang, W., Zhan, J., Rawat, D.B., & Bhattarai, A. (2024). Recent advances in anomaly detection in Internet of Things: Status, challenges, and perspectives. *Computer Science Review*, 54, article number 100665. doi: 10.1016/j.cosrev.2024.100665.

[2] Agyemang, E.F. (2024). Anomaly detection using unsupervised machine learning algorithms: A simulation study. *Scientific African*, 26, article number e02386. doi: 10.1016/j.sciaf.2024.e02386.

[3] Ahmad, R., Alsmadi, I., Alhamdani, W., & Tawalbeh, L. (2023). Zero-day attack detection: A systematic literature review. *Artificial Intelligence Review*, 56, 10733-10811. doi: 10.1007/s10462-023-10437-z.

[4] Ahmed, M., Chen, J., Akpaku, E., & Sosu, R.N. (2025). MTCR-AE: A multiscale temporal convolutional recurrent autoencoder for unsupervised malicious network traffic detection. *Computer Networks*, 261, article number 111147. doi: 10.1016/j.comnet.2025.111147.
[5] Akif, M.A., Butun, I., Williams, A., & Mahgoub, I. (2025). Hybrid machine learning models for intrusion detection in IoT: Leveraging a real-world IoT dataset. *ArXiv*, 2502, article number 12382. doi: 10.48550/arXiv.2502.12382.
[6] Akuthota, U.C., & Bhargava, L. (2025). The role of machine and deep learning in modern intrusion detection systems: A comprehensive review. *Computers and Electrical Engineering*, 124, article number 110318. doi: 10.1016/j.compeleceng.2025.110318.
[7] AL-Husseini, H., Hosseini, M.M., Yousofi, A., & Alazzawi, M.A. (2024). Whale optimization algorithm-enhanced long short-term memory classifier with novel wrapped feature selection for intrusion detection. *Journal of Sensor and Actuator Networks*, 13(6), article number 73. doi: 10.3390/jsan13060073.
[8] Anser, O., François, J., & Chrisment, I. (2024). Automated machine learning configuration to learn intrusion detectors on attack-free datasets. In *Proceedings of the 49th IEEE conference on local computer networks* (pp. 1-7). Normandy: IEEE. doi: 10.1109/LCN60385.2024.10639690.
[9] Canadian Institute for Cybersecurity Intrusion Detection System. (2017). Retrieved from https://www.unb.ca/cic/datasets/ids-2017.html.
[10] Chychkarov, Y., Zinchenko, O., Bondarchuk, A., & Aseeva, L. (2023). Detection of network intrusions using machine learning algorithms and fuzzy logic. *Electronic Professional Scientific Journal "Cybersecurity: Education, Science, Technique"*, 3(19), 209-225. doi: 10.28925/2663-4023.2023.19.209225.
[11] Cohen, S., Goldshlager, N., Shapira, B., & Rokach, L. (2023). TTANAD: Test-time augmentation for network anomaly detection. *Entropy*, 25(5), article number 820. doi: 10.3390/e25050820.
[12] Dairi, A., Harrou, F., Bouyeddou, B., Senouci, S.-M., & Sun, Y. (2023). Semi-supervised deep learning-driven anomaly detection schemes for cyber-attack detection in smart grids. In H.H. Alhelou, N. Hatziargyriou & Z.Y. Dong (Eds.), *Power systems cybersecurity: Methods, concepts, and best practices* (pp. 265-295). Cham: Springer. doi: 10.1007/978-3-031-20360-2_11.
[13] Diana, L., Dini, P., & Paolini, D. (2025). Overview on intrusion detection systems for computers networking security. *Computers*, 14(3), article number 87. doi: 10.3390/computers14030087.
[14] Firdaus, G.M., & Suryani, V. (2024). Enhanced pruning process with DBSCAN for attack detection. In *Proceedings of the 2024 1st international conference on cyber security and computing* (pp. 113-118). Melaka: IEEE. doi: 10.1109/CyberComp60759.2024.10913747.
[15] Gutierrez-Portela, F., Almenares Mendoza, F., & Calderon-Benavides, L. (2024). Evaluation of the performance of unsupervised learning algorithms for intrusion detection in unbalanced data environments. *IEEE Access*, 12, 190134-190157. doi: 10.1109/ACCESS.2024.3516615.
[16] Haidur, H.I., Gakhov, S.O., & Bryhynets, A.A. (2023). Detection of network anomalies with neural networks algorithms. *Telecommunication and Information Technologies*, 78(1), 61-73. doi: 10.31673/2412-4338.2023.010416.
[17] Hassan, T.J., Ramchandra, A.R., & Ranganathan, P. (2024). Modeling and evaluation of false data injection attacks (FDIA) in DER inverters. In *Proceedings of the 56th North American power symposium* (pp. 1-7). El Paso: IEEE. doi: 10.1109/NAPS61145.2024.10741736.
[18] He, H., Liu, K., Zhang, L., Xu, K., Li, J., & Ren, J. (2025). TE-PADN: A poisoning attack defense model based on temporal margin samples. *Big Data Research*, 40, article number 100528. doi: 10.1016/j.bdr.2025.100528.
[19] Ilin, D., & Starinskyi, I. (2023). Mathematical model of an intrusion detection system using a neural network based on auto-encoders. *Modern Information Technologies in the Sphere of Security and Defence*, 47(2), 113-118. doi: 10.33099/2311-7249/2023-47-2-113-118.
[20] Iturbe-Araya, J.I., & Rifà-Pous, H. (2025). Enhancing unsupervised anomaly-based cyberattacks detection in smart homes through hyperparameter optimization. *International Journal of Information Security*, 24, article number 45. doi: 10.1007/s10207-024-00961-6.
[21] Loo, F.-N., Chen, T.-S., Tsai, C.-W., Lin, J., & Yang, C.-W. (2024). Feasibility study of semi-supervised intrusion detection systems for industrial control systems. In *Proceedings of the 10th international conference on applied system innovation* (pp. 166-168). Kyoto: IEEE. doi: 10.1109/ICASI60819.2024.10547975.
[22] Lu, K.-C., Liu, I.-H., Liu, Z.-C., & Li, J.-S. (2024). Common criteria for security evaluation and malicious intrusion detection mechanism of dam supervisory control and data acquisition system. *IET Networks*, 13(5-6), 546-559. doi: 10.1049/ntw2.12127.
[23] Meidan, Y., Avraham, D., Libhaber, H., & Shabtai, A. (2023). CADeSH: Collaborative anomaly detection for smart homes. *IEEE Internet of Things Journal*, 10(10), 8514-8532. doi: 10.1109/JIOT.2022.3194813.
[24] Nikitenko, A., & Bashkov, Y. (2024). Construction of a network intrusion detection system based on a convolutional neural network and a bidirectional gated recurrent unit with attention mechanism. *Eastern-European Journal of Enterprise Technologies*, 3(9), 6-15. doi: 10.15587/1729-4061.2024.305685.

[25] Rajathi, C., & Rukmani, P. (2025). Hybrid learning model for intrusion detection system: A combination of parametric and non-parametric classifiers. *Alexandria Engineering Journal*, 112, 384-396. doi: 10.1016/j.aej.2024.10.101.
[26] Russell-Gilbert, A., Mittal, S., Rahimi, S., Seale, M., Jabour, J., Arnold, T., & Church, J. (2025). RAAD-LLM: Adaptive anomaly detection using LLMs and RAG integration. *ArXiv*, 2503, article number 02800. doi: 10.48550/arXiv.2503.02800.
[27] Sharma, N., Arora, B., Ziyad, S., Singh, P.K., & Singh, Y. (2024). A holistic review and performance evaluation of unsupervised learning methods for network anomaly detection. *International Journal on Smart Sensing and Intelligent Systems*, 17(1), article number 16. doi: 10.2478/ijssis-2024-0016.
[28] Shukla, A.K., Srivastav, S., Kumar, S., & Muhuri, P.K. (2023). UInDeSI4.0: An efficient unsupervised intrusion detection system for network traffic flow in Industry 4.0 ecosystem. *Engineering Applications of Artificial Intelligence*, 120, article number 105848. doi: 10.1016/j.engappai.2023.105848.
[29] Srivastav, S., Shukla, A.K., Kumar, S., & Muhuri, P.K. (2025). HYRIDE: HYbrid and Robust Intrusion DEtection approach for enhancing cybersecurity in Industry 4.0. *Internet of Things*, 30, article number 101492. doi: 10.1016/j.iot.2025.101492.
[30] Tsai, C.-W., Lu, C.-T., Li, C.-H., & Zhang, S.-W. (2024). An effective ensemble electricity theft detection algorithm for smart grid. *IET Networks*, 13(5-6), 471-485 doi: 10.1049/ntw2.12132.
[31] Uddin, M.A., Aryal, S., Bouadjenek, M.R., Al-Hawawreh, M., & Talukder, M.A. (2024). UsfAD based effective unknown attack detection focused IDS framework. *Scientific Reports*, 14, article number 29103. doi: 10.1038/s41598-024-80021-0.
[32] Zoppi, T., Ceccarelli, A., Puccetti, T., & Bondavalli, A. (2023). Which algorithm can detect unknown attacks? Comparison of supervised, unsupervised and meta-learning algorithms for intrusion detection. *Computers and Security*, 127, article number 103107. doi: 10.1016/j.cose.2023.103107.

Виявлення вторгнень у телекомунікаційних мережах за допомогою аналізу мережевого трафіку

Андрій Рій

Аспірант
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0009-0005-0252-1533

Мар'ян Кирик

Доктор технічних наук, професор
Національний університет «Львівська політехніка»
79000, вул. Степана Бандери, 12, м. Львів, Україна
https://orcid.org/0000-0001-9156-9347

Анотація. Зростання кіберзагроз і обсягів мережевого трафіку підкреслили потребу в ефективних методах виявлення аномалій. Метою дослідження була розробка гібридної моделі виявлення аномалій у телекомунікаційних мережах із підвищеною точністю на основі порівняльного аналізу безконтрольних алгоритмів. У ході роботи було проведено порівняння чотирьох основних алгоритмів: Isolation Forest, Local Outlier Factor (LOF), One-Class Support Vector Machine (SVM) та Elliptic Envelope, за допомогою симульованих даних мережевого трафіку. Методи аналізу включали нормалізацію даних за допомогою Z-score та Min-Max Scaling для усунення масштабних відмінностей між ознаками. Для підвищення точності виявлення аномалій було застосовано ансамбль із 100 дерев рішень. У результаті було виявлено, що Isolation Forest забезпечував найбільшу точність виявлення аномалій (F1-Score = 85,8 %) і високу швидкість обробки (час обробки на 10000 записів у 2,5 сек), що було важливим для реальних умов телекомунікаційних мереж з великими обсягами даних. LOF показав високу точність при виявленні локальних аномалій, але з більшою обчислювальною складністю. Алгоритм One-Class SVM виявляв глобальні аномалії, але демонстрував нижчу точність для локальних. Elliptic Envelope мав обмежену ефективність через припущення про нормальний розподіл даних. Додатково було розроблено комплексну модель, яка поєднувала переваги Isolation Forest, LOF та Density-Based Spatial Clustering of Applications with Noise, що дозволило підвищити F1-Score до 88 % та перевищити результати окремих моделей. Гібридна модель показала адаптивність до багатомодальних розподілів та ефективність у виявленні локальних аномалій. Практична значимість полягала в підвищенні точності й стійкості систем мережевої безпеки та формуванні основ для адаптивних алгоритмів реального часу, що дозволило фахівцям з кібербезпеки та телекомунікацій ефективніше моніторити та захищати мережі від загроз

Ключові слова: виявлення аномалій; оцінка продуктивності; масштабованість; точність; безконтрольні алгоритми; гібридна модель; кластеризація

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en Vol. 30 No. 2. 2025

UDC 338.43
DOI: 10.62660/bcstu/2.2025.77

Evaluating the effectiveness of image recognition systems for automatic detection of malicious files based on image metadata

Vitalii Yasenenko*
Master, Senior Software Developer
TP-Link
92618, 36 Technology St., Irvine, USA
https://orcid.org/0009-0004-4801-9541

Abstract. The relevance of the study was determined by the increasing threat of covert distribution of malicious software through the metadata of digital images, which complicated the detection using standard methods. The aim of the work was to develop a new approach to detecting malicious files through the analysis of image metadata using artificial intelligence methods. To achieve this, a detailed analysis of the main metadata standards was carried out, and vulnerable fields capable of hiding malicious code and being ignored by traditional security methods were identified. The results of the theoretical study showed that the most informative characteristics for threat detection were metadata features such as timestamps, geolocation coordinates, and device data. It was also established that non-standard values in the fields, such as abnormal timestamps or suspicious code markers, could serve as indicators of malicious activity. A comparison of traditional threat detection methods was conducted, which revealed the low effectiveness when working with metadata, as these methods were mainly focused on identifying malicious elements in the visual part of the file rather than on analysing the accompanying structure. The developed conceptual model, oriented towards the specified characteristics, demonstrated significant potential for effectively detecting anomalies and hidden malicious code in metadata. This approach made it possible to reduce the number of false positives, as it focused not only on detecting obvious deviations but also on subtler changes in the structural layer of images. The conclusions confirmed that the analysis of accompanying information was an important tool for detecting new forms of threats. The practical significance of the study lay in the possibility of using the proposed concept as a basis for the development of specialised systems for monitoring and preventing cyber incidents

Keywords: steganography; malware; digital images; machine learning; metadata anomalies

INTRODUCTION

The issue of covert distribution of malicious software through digital images remained one of the most complex challenges in the field of cybersecurity. One of the latest directions in this domain was the use of image metadata to conceal malicious code, which significantly complicated the process of detecting threats using standard file-checking methods. A particular threat was posed by subtle modifications in the pixels or metadata structure that did not cause visual changes and remained invisible to traditional security systems. Due to the combination of classical and multi-level methods of hiding information, traditional detection approaches proved to be ineffective. With the advancement of steganographic capabilities, there was a growing need for new systems of automatic threat recognition capable of analysing not only the visual component but also hidden information in the structure of images. Therefore, the development and evaluation of the effectiveness of

Article's History: Received: 24.12.2024; Revised: 17.04.2025; Accepted: 16.06.2025

Suggested Citation:
Yasenenko, V. (2025). Evaluating the effectiveness of image recognition systems for automatic detection of malicious files based on image metadata. *Bulletin of Cherkasy State Technological University*, 30(2), 77-87. doi: 10.62660/bcstu/2.2025.77.

*Corresponding author

such systems based on artificial intelligence (AI) was a pressing task for modern cybersecurity.

Against the backdrop of active development of steganographic techniques in metadata, the task of creating such methods and developing techniques for detecting hidden information became important. Y. Fernando *et al*. (2024) presented an innovative approach to steganography via image metadata, which ensured a high level of concealment of transmitted information. However, the authors focused on improving hiding methods while leaving the issue of detection without thorough analysis. In the work of D.R. Setiadi *et al*. (2025), a wide range of modern steganographic methods were reviewed, including image, textual and 3D data, but the need to explore practical aspects of detecting malicious modifications in metadata was also noted.

The challenges of increased data stealth and detection difficulty were addressed in the study by O. Kuznetsov *et al*. (2024a), which showed that using a wide range of signals in steganography significantly increased detection complexity. This, in turn, required the adaptation of more complex analytical models, including AI applications. In this context, the works of L. Caviglione & W. Mazurczyk (2022) outlined the danger of stego-malware as hidden threats, emphasising the ineffectiveness of traditional antivirus tools in detecting new types of attacks. V. Verma *et al*. (2022) supported this conclusion by describing the difficulties of identifying malicious code disguised as legitimate digital objects in Windows environments.

The use of deep learning for image analysis was explored by J. El Abdelkhalki *et al*. (2022), where the successful application of neural networks for detecting malicious content was highlighted. However, attention was paid to the analysis of the visual part of images, while the issue of hidden threats in metadata remained underexplored. The study by A.I. Iskanderani *et al*. (2021) demonstrated the effectiveness of using AI for steganalysis of digital objects, but the authors did not address the modelling of threats at the metadata level.

Additionally, C. Ahmadi *et al*. (2024), in their work, analysed the use of steganography in backdoor attacks against AI models, emphasising the difficulty of detecting such intrusions, which further underscored the need to create adaptive protection mechanisms. Meanwhile, reviews by I.H. Sarker (2023) and A.H. Salem *et al*. (2024) confirmed the growing role of intelligent systems in ensuring cybersecurity, but the issue of metadata analysis remained insufficiently developed in the works.

The comparison of traditional and intelligent solutions for threat detection tasks conducted by F. Wang & Y. Tang (2024) showed the advantages of AI methods in complex scenarios, although without detailing the specific features of working with supplementary information in media files. Regarding Ukrainian research, A. Kashtalian *et al*. (2024) developed a multi-computer system for detecting malicious software considering metamorphic changes, focusing on code transformations and behavioural characteristics of programs. However, metadata structure remained outside the scope of analysis. At the same time, A. Kobozieva *et al*. (2023) focused on developing steganalysis methods for digital video and image sequences, paying attention to changes at the bit-sequence level. However, the approaches were primarily oriented towards content analysis rather than structural metadata features, leaving open the question of integrating such solutions into the task of detecting malicious attachments in image-associated data.

The aim of the study was to develop a conceptual approach to detecting malicious files by analysing image metadata using modern AI methods. To achieve this aim, the following objectives were formulated: to justify the selection of key metadata characteristics subject to analysis for detecting malicious software; to analyse typical scenarios of using steganography in image metadata to conceal malicious software; to assess the potential of using AI methods for automatic threat recognition in metadata; and to develop a basic concept for a model that detects malicious files based on metadata analysis.

MATERIALS AND METHODS

The assessment of the possibility of concealing malicious software in digital images was conducted through the analysis of the technical characteristics of metadata formats, the usage patterns in file processing, and potential risk aspects related to manipulation of structural elements. The study took into account scenarios of covert threat distribution via modification, addition, or encryption of information in metadata fields without affecting the visual integrity of the file. The examination was based on open technical documentation of metadata formats, such as the EXIF Specification (Camera & Imaging Products Association, 2012) and the International Press Telecommunications Council (2024), as well as cybersecurity analytical reports, such as the Internet Security Threat Report (2017), which explored the use of steganography for hiding malicious software. For the sake of analytical relevance, the materials were selected based on criteria of topicality, practical orientation, and alignment with the subject of malicious code detection in metadata. Open archives of digital images also played an important role, providing examples of typical metadata structures, such as the OpenImages Dataset (Krasin *et al.*, 2017), Break Our Steganographic System (Bas *et al.*, 2011), and StegoAppDB (Newman *et al.*, 2019).

Theoretical modelling of the malware detection system in image metadata was performed through the analysis of structural features of the exchangeable image file format (EXIF), the international press telecommunications council metadata standard (IPTC), and the

extensible metadata platform (XMP), involving open technical specifications and academic publications in the field of information security. To develop the conceptual design, data were used on metadata organisation, potential attack vectors, and known practices of manipulating supplementary information in digital files. Critical metadata characteristics were identified through the systematisation of fields that allowed unrestricted input or storage of large data volumes without mandatory validation. Particular attention was paid to timestamps, geolocation coordinates, device identifiers, software versions, structural consistency of fields, and total metadata volume. The list of features was formed based on comparative analysis of typical and anomalous values found in publicly available metadata examples of digital images.

The conceptual model of processing was built by determining the main stages of input data transformation: extraction of supplementary information, removal of secondary fields, normalisation of numerical values to a single scale, and encoding of categorical variables using one-hot encoding. All processed features were considered as an input vector for further analysis using machine learning. To ensure the model's resilience to data structure variability, consistency of feature formats was maintained regardless of the original metadata type. The classification component of the concept was developed with a focus on using a shallow multi-layer perceptron (MLP) neural network. This decision was based on the requirements for computational efficiency and the model's ability to adapt to complex interdependencies between features without manual pattern identification. The alternative use of gradient boosting models was considered for situations requiring increased classification accuracy with limited input features. To manage classification sensitivity, the system proposed a probabilistic interpretation of model outputs with the ability to adjust threshold values.

The conceptual verification of the system's architecture was carried out by logical modelling of potential scenarios for concealing malicious code in metadata, based on typical anomalies and non-standard structural modifications described in the literature. The modelling was theoretical in nature and involved risk analysis arising from manipulations with metadata fields (UserComment, MakerNote, XMP custom fields, etc.). All assumptions were based on publicly available technical documentation and analytical reports, which avoided the experimental influence of variable environments and ensured the generalisability of the concept for future implementation.

RESULTS

Metadata of digital images represent a critically important element of modern information systems, providing contextual descriptions of content, facilitating indexing, copyright management, and automation of content processing. However, alongside the widespread use of metadata, significant information security threats arise due to the peculiarities of implementation and functioning of the main metadata storage standards. The most common formats are EXIF, IPTC, and XMP, each with specific data organisation characteristics and corresponding vectors of potential attacks.

The EXIF standard, initially developed for digital cameras, involves embedding metadata directly into image files, primarily in JPEG or TIFF formats. Since EXIF was primarily aimed at preserving technical information, such as shooting parameters, camera model, or geolocation data, the issue of metadata structure security was not a priority during the design of the standard. Fields that allow free input of data, such as "UserComment" and "MakerNote", may store arbitrary text or binary arrays without proper validation. This creates a possibility for injecting hidden code, including scripts or objects, for subsequent system exploitation during file processing by inadequately protected software (Monika & Eswari, 2023).

IPTC metadata, used mainly in professional journalism and media production, allow the description of textual image attributes: titles, keywords, author names, usage rights, etc. The IPTC format provides stricter data structuring compared to EXIF, but several fields still allow the input of large volumes of text without format or length limitations. Accordingly, if the receiving system does not implement adequate mechanisms for checking or cleaning textual data, such fields may be exploited to inject malicious content, including SQL injections, scripts, or other forms of attacks on system integrity or confidentiality (Jian *et al.*, 2021).

The XMP format, proposed by Adobe Systems, implements a more flexible metadata concept based on the principles of extensibility and integration. XMP uses standardised XML structures to describe metadata fields, allowing the inclusion of an unlimited number of hierarchically organised records. Compared to EXIF and IPTC, XMP offers greater flexibility and expansion potential by adding user-defined schemas and namespaces. At the same time, the use of XML as a foundation introduces additional vulnerabilities specific to the processing of structured data. Since most XML parsing libraries do not impose default restrictions on document depth or size, inadequate handling of XMP metadata may become a serious attack vector on applications interacting with digital images. For illustration, Table 1 presents an example of a typical EXIF metadata set and a simulated example with anomalous values that may potentially contain hidden payloads or indicate steganographic use of the field.

Thus, digital image metadata can be used to hide malware using a variety of techniques aimed at manipulating the structural elements of the file. Table 2 lists the main methods of hiding threats through metadata.

Table 1. Examples of typical and anomalous values in digital image metadata

Field	Typical value	Abnormal value
Make	Canon	$$%null%%
Model	EOS 600D	bash_exec:systemctl.sh
Software	Adobe Photoshop CC 2019	aHR0cDovL21hbGljaW91cy5leGU=
UserComment	Summer vacation, Aug 2022	<script>eval('malware')</script>
XMP Payload	-	{"payload":"execute","flag":true}

Source: created by the author

Table 2. Typical methods of hiding malware in digital image metadata

Hiding method	Metadata field	Risk description
Inserting encrypted text	UserComment, MakerNote	Hidden payload among legitimate data
Creating redundant fields	XMP custom fields	Injection of non-standard metadata
Manipulation of geolocation data	GPSInfo	Transferring malicious commands
Paste into file header	File Header	Code masking in service segments
Mutation of standard fields	Make, Model, Software	Imitation of legitimate data to hide commands
Using non-standard encodings	UserComment, XMP Description	Data masking via Base64, Hex and other encodings

Source: created by the author based on data from the Camera & Imaging Products Association (2012), Internet Security Threat Report (2017), and International Press Telecommunications Council (2024)

The analysis of metadata storage standards demonstrated that the existing solutions did not provide an adequate level of security without additional measures of validation, cleaning, and input data restriction. In this regard, it became necessary to investigate the effectiveness of existing mechanisms for detecting hidden malicious objects in digital images, focusing both on traditional antivirus approaches and on specialised steganalysis methods.

The main methods used in practice remained general-purpose antivirus systems. The functioning was based on searching for known signatures, behavioural analysis of objects, or heuristic anomaly detection. However, when working with images, antivirus systems demonstrated significant limitations. Image files by nature contained numerous permissible variations in content and metadata, complicating the construction of unambiguous signatures for malicious modifications. In addition, antivirus products were mostly content-oriented and did not conduct in-depth analysis of accompanying data, which allowed attackers to use metadata for transferring hidden code without hindrance (Sarker, 2024).

Steganography in metadata differs from classical methods of hiding visual information in that it uses textual or binary metadata fields to embed hidden content. Therefore, another approach to detecting concealed changes in images became steganalytic methods based on pixel structure analysis. Such methods aimed to detect traces of steganography by examining statistical anomalies in pixel distribution, histogram analysis, or the calculation of specific signal characteristics. The use of such approaches allowed for the highly sensitive identification of classical forms of information hiding in visual data (Płachta *et al.*, 2022). However, in the case of working with metadata, the effectiveness of steganalytic methods significantly decreased, as the accompanying image information could not be analysed via visual features (Kuznetsov *et al.*, 2024b). Thus, traditional pixel-based analysis did not allow for the detection of hidden threats embedded at the level of an image's structural data.

It is also worth noting attempts to use simple machine learning models to detect changes in file data structure. These models were generally built on classification or clustering approaches using limited sets of features, such as metadata size, number of fields, or types of attributes used. Although the results of such analysis made it possible to detect obvious anomalies, these methods remained insensitive to complex and well-camouflaged threats. The models often failed to consider deep interrelations between different metadata characteristics and were prone to high levels of false positives in cases of non-standard but legitimate changes to file structure. For a visual comparison of the effectiveness of traditional methods for detecting hidden malicious software in metadata of digital images, Figure 1 illustrates the dependence of detection accuracy on the level of attack complexity.

As a result of the analysis of existing solutions, it was established that none of the traditional approaches provided an adequate level of effectiveness in detecting concealed malicious code in image metadata. The lack of adaptability in models and the limitations of classical analysis methods at the level of content or simple characteristics of accompanying data created significant obstacles to detecting new forms of threats. Thus, the analysis showed that existing methods had limited potential for identifying modern threats, especially those masked within the metadata layer of digital images. Consequently, there arose a need to develop a solution focused specifically on analysing the structural information of files.

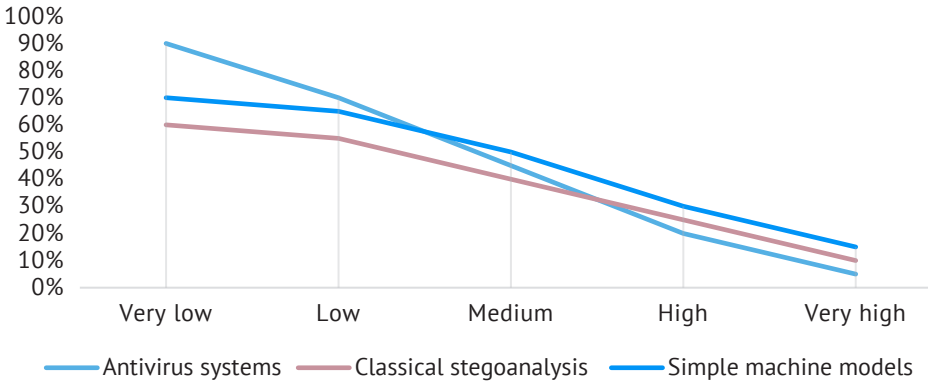

Figure 1. Dependence of the effectiveness of various methods for detecting malware in metadata on the complexity of attacks

Source: created by the author

Based on the analysis conducted, it was expected that the input to the proposed system could be a digital image file in JPEG, PNG, or another common graphic format containing a metadata structure. Given the specific task – detecting concealed malicious elements – it would be reasonable to focus not on the full visual analysis of the image, but rather on processing its structural part. Metadata are considered a potentially vulnerable carrier of hidden information, which justifies the prioritisation in the analysis. To ensure a structured approach to detecting malicious software in image metadata, the proposed system envisaged the sequential execution of several stages of data processing. Figure 2 presents the general system architecture, illustrating the main stages of analysis of the input object.

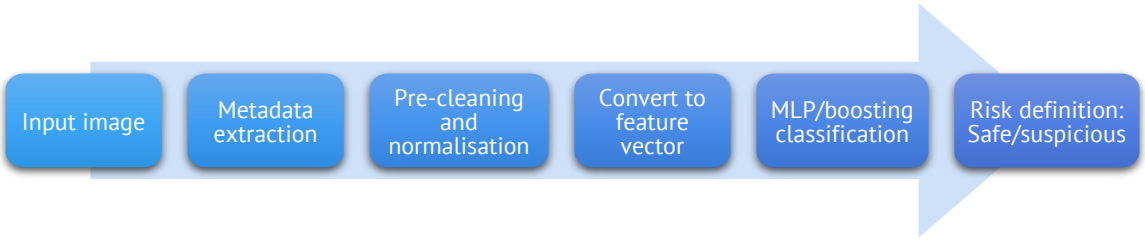

Figure 2. General scheme of digital image processing in the proposed threat detection system

Source: created by the author

It was envisaged to perform preliminary parsing of the input file to extract supplementary data, such as the EXIF section and other accessible fields, characteristic of the respective format. This approach was determined by the fact that metadata usually contains static textual or numerical values, which simplifies the processing compared to the analysis of complex pixel structures. Metadata analysis was expected to include the study of a number of characteristics that could potentially serve as indicators of malicious activity. Specifically, it was proposed to evaluate the overall size of the EXIF section, the number of present tags, the types of tags used, and to check for anomalous or unusual values in the metadata fields. Such anomalies could include excessively large or negative numerical values, incorrect date formats, unusual or suspicious codes identifying device manufacturers or shooting parameters.

To increase system reliability, a superficial analysis of basic pixel structure characteristics was allowed – for example, noise levels or microdeformations – however, the main focus would remain on processing metadata as a more probable vector for threat concealment. The most informative elements for analysis proved to be the timestamps for file creation and editing, geolocation coordinates, recording device data, software version numbers, and the total metadata size. Anomalies in these fields, especially in cases of inconsistency or excessive size, could indicate attempts to conceal a malicious payload (Zuppelli *et al.*, 2021).

In selecting the type of machine learning model, it would be appropriate to focus on the need to ensure a balance between high analytical sensitivity and moderate computational demands. Accordingly, one of the most suitable architectures was considered to be an MLP with some layers. This structure allows for efficient operation with medium-dimensional input vectors and provides sufficient flexibility for adaptive learning. Alternatively, gradient boosting methods could have been chosen, due to demonstrating high accuracy on tabular data. However, the neural network was considered the preferred option due to its ability to self-update and better suitability for processing new types of threats

without requiring full model reconstruction. Compared to decision trees or support vector machines, MLPs offer a higher learning capacity on complex interdependencies between features, which is critically important for detecting non-standard and masked malicious patterns in metadata structures (Yadav *et al.,* 2022).

The data processing concept in the system was expected to involve several sequential stages. First, the structural information extracted from the input image would be cleansed of secondary attributes, then converted into a feature vector. When forming input feature vectors for the model, preprocessing steps such as normalisation of numerical values (e.g. scaling to a 0 to 1 range) and one-hot encoding for categorical fields like tag types or device identifiers were to be applied. Such data preparation would allow the model to more effectively identify patterns in mixed feature sets. This vector would then be fed into the model, which would compute internal representations and classify the file into two classes: safe object or object with potential threat. To improve prediction accuracy, a probabilistic interpretation of the model output would be used, enabling flexible adjustment of trigger thresholds depending on the acceptable risk level in specific application conditions.

At the system's output, a decision on the analysed object would be generated in the form of a percentage probability that the metadata contained hidden malicious code. For practical use, an integrated interpretation of the result was also foreseen, in the form of a simple label – "safe" or "suspicious" – which would simplify operational decision-making on whether to further process or isolate the file. The proposed model was expected to ensure adaptability to new methods of malicious code concealment in metadata, while maintaining a low rate of false positives when processing atypical but safe changes. Thanks to the ability of neural networks to learn from complex multidimensional data, it was anticipated that such a system could significantly enhance the reliability and timeliness of detecting modern cyber threats that remain beyond the reach of traditional analysis methods.

The expected results of the model's functioning were clarified through an analysis of typical use-case scenarios. For instance, in cases of processing images with EXIF metadata of standard size-approximately 1,000 bytes – containing only typical tags such as camera model, exposure parameters, etc., a high probability of safe classification at 98-99% was anticipated. Conversely, if anomalies were present in the metadata – non-standard date formats, atypical or excessively large numerical values in geolocation or timestamp fields – the system would be expected to demonstrate a significant increase in threat detection probability. In such cases, the likelihood of classifying the image as suspicious was expected to rise to 80-90%, depending on the nature and intensity of the detected deviations. Thus, the model could effectively support decisions regarding the further processing or isolation of potentially dangerous files, which is critically important for ensuring information security amid growing cyber threats.

If image processing involved metadata with many tags and anomalous values – e.g. up to 2,000 bytes of data in the EXIF section, where strange or non-standard codes for camera manufacturers or shooting parameters were detected – this could result in a significantly higher probability of threat detection, reaching 95% or more. In such a case, the system would need to be capable of adjusting its trigger thresholds to prevent excessive false positives, ensuring that ordinary files with mistakenly incorrect but harmless metadata would not be blocked. To maintain the model's relevance, it was possible to foresee the implementation of a mechanism for periodic updates based on new data. The system could store anonymised feature vectors of new files and periodically undergo fine-tuning based on stored data without requiring full retraining. This would allow the model to adapt to evolving methods of malicious code obfuscation.

At the same time, understanding and awareness of potential issues proved essential for further improving the model. One such issue could be the processing of data with certain but frequently recurring metadata anomalies that are not harmful, but might be misclassified as threats. This phenomenon may occur due to non-standard yet safe user operations – for example, when metadata contains incorrect or unusual values that do not conform to general patterns. This situation, in particular, could lead to false positives. However, these errors could be minimised due to the possibility of continuous learning and model adjustment on new data (Carneiro *et al.,* 2023). An essential part of this process is the adaptation of algorithms for optimising classification based on new data, which reduces the likelihood of errors in future use.

Another possible issue could be the trade-off between data processing speed and result accuracy. In cases where metadata has a complex structure or an image contains a large volume of data, the processing time may increase, potentially affecting overall system speed. Nevertheless, thanks to innovative optimisation techniques such as parallel data processing or the use of more powerful computing resources, this issue could be significantly mitigated (Monteiro *et al.,* 2021). The prospects for integrating faster computational mechanisms would significantly improve processing speed without loss of accuracy, which would positively impact overall efficiency. Further important challenge might be the scalability of the model as the volume of data processed increases. However, this issue could be addressed by improving big data processing methods, applying intelligent algorithms for automatic resource allocation, and reducing the need for constant processing of all data via early-stage filtering. As a result, these measures would allow the system to remain effective even when working with large volumes of metadata. Thus, despite some potential challenges, the implementation of this model could demonstrate significant potential for improving threat detection accuracy and

speed. Thanks to its ability to adapt to new data and to optimised processing workflows, the model would maintain a high level of effectiveness even under complex or changing data conditions, offering a considerable advantage over traditional security methods.

DISCUSSION

Automated threat detection systems based on AI methods have become an important tool in countering steganographic techniques used by malicious actors to conceal harmful components within digital objects. Data analysis in such systems usually focuses on visual information; however, metadata remains a less explored, albeit promising, source of intrusion indicators. The study examined the potential for constructing a model to detect malicious images based on the analysis of file structural information. This approach allowed not only for reduced computational costs, but also for identifying anomalies that remain invisible to systems focused on image content. In this regard, it became necessary to compare the obtained results with other contemporary studies on the use of AI in the classification and detection of threats in graphical data.

The proposed approach to analysing metadata of digital images as the primary source of potential threat indicators proved relevant in the context of an increasing number of attacks using steganographic methods to disguise malicious code. The conducted study confirmed that it is the structural characteristics of fields, rather than the visual features of images, that represent a promising target for threat modelling. This line of inquiry continued the direction initiated by D. Puchalski *et al*. (2020), who demonstrated the viability of detecting malicious objects through structural analysis of media files. The application of detailed parsing of accompanying information allowed for increased classification accuracy of objects while reducing the risk of false positives during atypical but legitimate file modifications.

Particular attention should be paid to the work of S. Kiltz *et al*. (2024), which implemented metadata-based tracing of image changes to detect files with hidden malicious payloads. The study confirmed that even with minimal visual changes, metadata remains a sensitive indicator of interference. This reinforced the conclusions regarding the effectiveness of metadata as a vector of analysis, complementing classical pixel comparison methods. Equally significant for drawing conclusions was the analysis of approaches to constructing classification models that use images as a source of input features. In the work of R. Chaganti *et al*. (2022), malware was represented as images and classified using EfficientNet. This approach demonstrated high accuracy in threat detection due to the capability of convolutional neural networks to detect complex spatial patterns. At the same time, this study substantiated the relevance of focusing on the structural features of accompanying data, thus avoiding excessive dependence on the visual representation of the object.

A similar approach was implemented in the study by S.A. Roseline *et al*. (2020), where a deep model based on a random forest was used to classify images based on visual features. The authors highlighted the advantages of the ensemble approach in improving accuracy and reducing the likelihood of overfitting in the presence of large volumes of heterogeneous input data. However, the model proved sensitive to visual deviations in pixel structure and was ineffective in cases where information concealment was carried out not through the visual channel, but at the level of the file's accompanying information. This confirmed the limitations of models focused exclusively on image analysis without considering contextual metadata. In this context, the structural analysis of metadata implemented in the current study made it possible to detect threats that remained outside the focus of visual-type models, while maintaining lower computational costs and greater adaptability to non-standard input object formats.

More advanced architectures were proposed by F. Ullah *et al*. (2022), who developed a hybrid model with visualisation of control-flow graphs and a multi-head self-attention mechanism. By combining structural features of program behaviour with deep analysis of instruction dependencies, the model achieved high accuracy in detecting complex threats. However, this approach required complex pre-processing, including decompilation and execution tracing, which rendered it unsuitable in scenarios with limited access to full code or in cases involving isolated media files. The concept proposed in the current study had the advantage that only the accompanying structural information was used for analysis, which is generally available without specialised tools. This ensured flexibility, scalability, and suitability for integration into high-performance environments with stringent response time requirements.

A significant contribution to model optimisation was offered in the study by A. El-Ghamry *et al*. (2023), where an effective IoT threat detection system was developed based on an optimised image-based classifier structure. The authors focused on reducing resource demands without sacrificing accuracy by combining relevant feature selection with adaptive learning. This approach reinforced the importance of focusing on compact yet meaningful feature sets, which aligns with the current study's orientation toward analysing a limited but informative segment of metadata. The proposed model did not require processing the visual content of the image, thereby lowering computational requirements and allowing effective detection of hidden anomalies in metadata structures, particularly through combining EXIF, IPTC, and XMP fields.

In the study by D. Vasan *et al*. (2024), a broad learning architecture was proposed for GPU-free classification, ensuring efficiency under limited resources by reducing model depth and employing multidimensional feature representation. This approach was evaluated as suitable for deployment in applied scenarios lacking

access to powerful graphics processors. The concept proposed in the current work, with a similar emphasis on resource efficiency, differed in that operations were conducted on tabular rather than visual vectors. This simplified data pre-processing and enabled system scalability without significantly burdening the runtime environment. The use of an MLP model with a low number of layers was supplemented by normalisation of input values, which collectively ensured efficient and adaptive processing of structural metadata features with the ability to accurately detect hidden threats.

The issue of algorithm explainability was also of particular importance in the context of applying intelligent systems for threat detection. The work by A. Galli *et al*. (2024) analysed methods for improving transparency in behavioural models for malware detection. The authors emphasised the importance of result interpretability for the practical implementation of AI solutions in cybersecurity. In light of this, the conceptual model proposed in this study provided for the possibility of outputting a probabilistic estimate and a flexible classification threshold, which would support better understanding of the analysis results even without a full "black box" model. This approach aligns with the position that algorithm transparency is a critical factor of trust in automated threat detection systems.

A general overview of the prospects of machine learning for malware detection was presented by J. Ispahany *et al*. (2024). This study pointed to limitations related to excessive dependence on signature databases, insufficient adaptability of existing systems to new types of attacks, and the difficulty of ensuring scalability. The conclusions obtained in the current study aligned with these statements, as the metadata-based detection concept aimed to overcome these limitations through structural analysis of accompanying data without reliance on known threat patterns. A particular prospect was the potential for continuous model updating without the need for full retraining, which met the challenges outlined by the authors of the review.

Another direction worth noting is the application of metadata in a broader sense – not only for images but also for network flows. M. Russo *et al*. (2021) demonstrated the effectiveness of network metadata analysis for detecting illegal cryptocurrency mining. The authors proved that even without access to full packet content, traffic metadata could serve as a sufficient source of indicators for threat detection. In this respect, the study reaffirmed a common idea: metadata, despite its auxiliary nature, can act as the primary indicator of malicious activity. This further justified the need to develop detectors focused on analysing the structural layer of data in various formats, including graphical ones. Comparison of the results obtained with modern approaches to malware detection showed that structural information analysis, particularly of digital image metadata, fills the gaps characteristic of methods focused solely on visual or behavioural features. The rationale for this approach was confirmed by several studies highlighting the importance of additional information layers for improving threat detection accuracy. Unlike traditional models, metadata analysis proved more informative in cases of complex or non-standard concealment forms and demonstrated better consistency with requirements for interpretability, speed, and adaptability.

CONCLUSIONS

The results of the study demonstrated the importance of metadata as a potential vector for concealing malicious programs, significantly complicating the detection using classical cybersecurity tools. It was established that standard threat detection approaches, mainly focused on visual features, are ineffective because of overlooking anomalous changes in the image's service structures. In particular, it was found that metadata in EXIF, IPTC, and XMP formats may contain hidden elements used to disguise malicious payloads. This significantly increases the vulnerability of information systems, as most traditional mechanisms analyse only the image content, ignoring its structural wrapper. Analysis of typical and anomalous field values, such as UserComment, Software, or XMP Payload, made it possible to identify characteristic injection vectors that require dedicated monitoring by detection systems.

Based on a comparative review and logical modelling, a conceptual model for detecting changes in digital image metadata using AI methods was proposed. The architecture, based on MLP, accounted for both numerical and categorical parameters after appropriate pre-processing, including scaling and encoding. The model demonstrated potential in detecting structural anomalies, such as incorrect date formats or suspicious geolocation coordinates, positioning it as a promising solution for practical implementation in the context of modern cyber threats. The results obtained underscored the advisability of focusing not only on file content but also on accompanying technical data, which may serve as an independent source of critically important information.

At the same time, the study identified a number of limitations, including the likelihood of false positives when processing non-standard but legitimate metadata changes, as well as challenges associated with system scalability when working with large volumes of data. To improve the efficiency of future implementations, it would be advisable to consider integrating additional approaches, such as behaviour-based file analysis at higher levels, and optimising computational algorithms to ensure rapid response without sacrificing accuracy. A promising area for further research is the adaptation of the model to new steganographic techniques and its integration into comprehensive cybersecurity systems, including automated tools for monitoring and countering multi-level threats.

ACKNOWLEDGEMENTS

None.

FUNDING
None.

CONFLICT OF INTEREST
None.

REFERENCES

[1] Ahmadi, C., Chen, J.L., & Lin, Y.T. (2024). Securing AI models against backdoor attacks: A novel approach using image steganography. *Journal of Internet Technology*, 25(3), 465-475. doi: 10.53106/160792642024052503012.
[2] Bas, P., Filler, T., & Pevný, T. (2011). Break our steganographic system – the BOSS contest. In *Proceedings of the 13th international conference on information hiding* (pp. 59-70). Berlin: Springer. doi: 10.1007/978-3-642-24178-9_5.
[3] Camera & Imaging Products Association. (2012). *Exchangeable image file format for digital still cameras: Exif Version 2.3*. Retrieved from https://www.cipa.jp/std/documents/e/DC-008-2012_E.pdf.
[4] Carneiro, D., Guimarães, M., Carvalho, M., & Novais, P. (2023). Using meta-learning to predict performance metrics in machine learning problems. *Expert Systems*, 40(1), article number e12900. doi: 10.1111/exsy.12900.
[5] Caviglione, L., & Mazurczyk, W. (2022). Never mind the malware, here's the stegomalware. *IEEE Security & Privacy*, 20(5), 101-106. doi: 10.1109/MSEC.2022.3178205.
[6] Chaganti, R., Ravi, V., & Pham, T.D. (2022). Image-based malware representation approach with EfficientNet convolutional neural networks for effective malware classification. *Journal of Information Security and Applications*, 69, article number 103306. doi: 10.1016/j.jisa.2022.103306.
[7] El Abdelkhalki, J., Ahmed, M.B., & Abdelhakim, B.A. (2022). Image malware detection using deep learning. *International Journal of Communication Networks and Information Security*, 12(2). doi: 10.17762/ijcnis.v12i2.4600.
[8] El-Ghamry, A., Gaber, T., Mohammed, K.K., & Hassanien, A.E. (2023). Optimized and efficient image-based IoT malware detection method. *Electronics*, 12(3), article number 708. doi: 10.3390/electronics12030708.
[9] Fernando, Y., Darwis, D., Mehta, A.R., Wamiliana, W., & Wantoro, A. (2024). A new approach of steganography on image metadata. *International Journal on Informatics Visualization*, 8(2), 968-976. doi: 10.62527/joiv.8.2.2110.
[10] Galli, A., La Gatta, V., Moscato, V., Postiglione, M., & Sperlì, G. (2024). Explainability in AI-based behavioral malware detection systems. *Computers & Security*, 141, article number 103842. doi: 10.1016/j.cose.2024.103842.
[11] International Press Telecommunications Council. (2024). *IPTC photo metadata standard 2024.1*. Retrieved from https://www.iptc.org/std/photometadata/specification/IPTC-PhotoMetadata.
[12] Internet Security Threat Report. (2017). Retrieved from https://www.symantec.com/content/dam/symantec/docs/reports/istr-22-2017-en.pdf.
[13] Iskanderani, A.I., Mehedi, I.M., Aljohani, A.J., Shorfuzzaman, M., Akther, F., Palaniswamy, T., Latif, S.A., & Latif, A. (2021). Artificial intelligence-based digital image steganalysis. *Security and Communication Networks*, 2021(1), article number 9923389. doi: 10.1155/2021/9923389.
[14] Ispahany, J., Islam, M.R., Islam, M.Z., & Khan, M.A. (2024). Ransomware detection using machine learning: A review, research limitations and future directions. *IEEE Access*, 12, 68785-68813. doi: 10.1109/ACCESS.2024.3397921.
[15] Jian, Y., Kuang, H., Ren, C., Ma, Z., & Wang, H. (2021). A novel framework for image-based malware detection with a deep neural network. *Computers & Security*, 109, article number 102400. doi: 10.1016/j.cose.2021.102400.
[16] Kashtalian, A., Lysenko, S., Savenko, O., Nicheporuk, A., Sochor, T., & Avsiyevych, V. (2024). Multi-computer malware detection systems with metamorphic functionality. *Radioelectronic and Computer Systems*, 2024(1), 152-175. doi: 10.32620/reks.2024.1.13.
[17] Kiltz, S., Dittmann, J., Loewe, F., Heidecke, C., John, M., Mädel, J., & Preißler, F. (2024). Forensic image trace map for image-stego-malware analysis: Validation of the effectiveness with structured image sets. In *Proceedings of the 2024 ACM workshop on information hiding and multimedia security* (pp. 125-130). doi: 10.1145/3658664.3659659.
[18] Kobozieva, A., Bobok, I., & Kushnirenko, N. (2023). Steganalysis method for detecting LSB embedding in digital video, digital image sequence. In *Information Control Systems and Technologies* (pp. 78-90). Odesa: CEUR.
[19] Krasin, I., *et al.* (2017). *OpenImages: A public dataset for large-scale multi-label and multi-class image classification*. Retrieved from https://github.com/openimages/dataset.
[20] Kuznetsov, O., Frontoni, E., & Chernov, K. (2024a). Beyond traditional steganography: Enhancing security and performance with spread spectrum image steganography. *Applied Intelligence*, 54(7), 5253-5277. doi: 10.1007/s10489-024-05415-z.
[21] Kuznetsov, O., Frontoni, E., Chernov, K., Kuznetsova, K., Shevchuk, R., & Karpinski, M. (2024b). Enhancing steganography detection with AI: Fine-tuning a deep residual network for spread spectrum image steganography. *Sensors*, 24(23), article number 7815. doi: 10.3390/s24237815.
[22] Monika, A., & Eswari, R. (2023). An ensemble-based stegware detection system for information hiding malware attacks. *Journal of Ambient Intelligence and Humanized Computing*, 14(4), 4401-4417. doi: 10.1007/s12652-023-04559-z.
[23] Monteiro, J.P., Ramos, D., Carneiro, D., Duarte, F., Fernandes, J.M., & Novais, P. (2021). Meta-learning and the new challenges of machine learning. *International Journal of Intelligent Systems*, 36(11), 6240-6272. doi: 10.1002/int.22549.

[24] Newman, J., Lin, L., Chen, W., Reinders, S., Wang, Y., Wu, M., & Guan, Y. (2019). StegoAppDB: A steganography apps forensics image database. In *Proceedings of the IS&T international symposium on electronic imaging: media watermarking, security, and forensics* (pp. 536-1-536-12). Springfield: Society for Imaging Science and Technology. doi: 10.2352/ISSN.2470-1173.2019.5.MWSF-536.
[25] Płachta, M., Krzemień, M., Szczypiorski, K., & Janicki, A. (2022). Detection of image steganography using deep learning and ensemble classifiers. *Electronics*, 11(10), article number 1565. doi: 10.3390/electronics11101565.
[26] Puchalski, D., Caviglione, L., Kozik, R., Marzecki, A., Krawczyk, S., & Choraś, M. (2020). Stegomalware detection through structural analysis of media files. In *Proceedings of the 15th international conference on availability, reliability and security* (article number 73). New York: Association for Computing Machinery. doi: 10.1145/3407023.3409187.
[27] Roseline, S.A., Geetha, S., Kadry, S., & Nam, Y. (2020). Intelligent vision-based malware detection and classification using deep random forest paradigm. *IEEE Access*, 8, 206303-206324. doi: 10.1109/ACCESS.2020.3036491.
[28] Russo, M., Šrndić, N., & Laskov, P. (2021). Detection of illicit cryptomining using network metadata. *EURASIP Journal on Information Security*, 2021, article number 11. doi: 10.1186/s13635-021-00126-1.
[29] Salem, A.H., Azzam, S.M., Emam, O.E., & Abohany, A.A. (2024). Advancing cybersecurity: A comprehensive review of AI-driven detection techniques. *Journal of Big Data*, 11(1), article number 105. doi: 10.1186/s40537-024-00957-y.
[30] Sarker, I.H. (2023). Multi-aspects AI-based modeling and adversarial learning for cybersecurity intelligence and robustness: A comprehensive overview. *Security and Privacy*, 6(5), article number e295. doi: 10.1002/spy2.295.
[31] Sarker, I.H. (2024). *AI-driven cybersecurity and threat intelligence: Cyber automation, intelligent decision-making and explainability*. Cham: Springer. doi: 10.1007/978-3-031-54497-2.
[32] Setiadi, D.R., Ghosal, S.K., & Sahu, A.K. (2025). AI-powered steganography: Advances in image, linguistic, and 3D mesh data hiding – a survey. *Journal of Future Artificial Intelligence and Technologies*, 2(1). doi: 10.62411/faith.3048-3719-76.
[33] Ullah, F., Srivastava, G., & Ullah, S. (2022). A malware detection system using a hybrid approach of multi-heads attention-based control flow traces and image visualization. *Journal of Cloud Computing*, 11(1), article number 75. doi: 10.1186/s13677-022-00349-8.
[34] Vasan, D., Hammoudeh, M., & Alazab, M. (2024). Broad learning: A GPU-free image-based malware classification. *Applied Soft Computing*, 154, article number 111401. doi: 10.1016/j.asoc.2024.111401.
[35] Verma, V., Muttoo, S.K., & Singh, V.B. (2022). Detecting stegomalware: Malicious image steganography and its intrusion in windows. In *Security, privacy and data analytics: Select proceedings of ISPDA 2021* (pp. 103-116). Singapore: Springer. doi: 10.1007/978-981-16-9089-1_9.
[36] Wang, F., & Tang, Y. (2024). Diverse intrusion and malware detection: AI-based and non-AI-based solutions. *Journal of Cybersecurity and Privacy*, 4(2), 382-387. doi: 10.3390/jcp4020019.
[37] Yadav, P., Menon, N., Ravi, V., Vishvanathan, S., & Pham, T.D. (2022). A two-stage deep learning framework for image-based android malware detection and variant classification. *Computational Intelligence*, 38(5), 1748-1771. doi: 10.1111/coin.12532.
[38] Zuppelli, M., Manco, G., Caviglione, L., & Guarascio, M. (2021). Sanitization of images containing stegomalware via machine learning approaches. In *Proceedings of the Italian conference on cybersecurity* (pp. 374-386). London: CEUR.

Оцінка ефективності систем розпізнавання зображень для автоматичного виявлення шкідливих файлів на основі метаданих зображень

Віталій Ясененко

Магістр, старший розробник програмного забезпечення
TP-Link
92618, вул. Технологій, 36, м. Ірвайн, США
https://orcid.org/0009-0004-4801-9541

Анотація. Актуальність дослідження обумовлена зростанням загрози прихованого розповсюдження шкідливих програм через метадані цифрових зображень, що ускладнює їх виявлення стандартними методами. Метою роботи було розробити новий підхід до виявлення шкідливих файлів через аналіз метаданих зображень із використанням методів штучного інтелекту. Для цього було проведено детальний аналіз основних стандартів метаданих, а також визначено уразливі поля, здатні приховувати шкідливий код та які ігноруються традиційними методами безпеки. Результати теоретичного дослідження показали, що найбільш інформативними для виявлення загроз є характеристики метаданих, такі як часові відмітки, геолокаційні координати та дані про пристрої. Окремо було встановлено, що нестандартні значення в полях, наприклад, аномальні часові позначки чи підозрілі кодові позначення, можуть виступати індикаторами шкідливої активності. Було здійснено порівняння традиційних методів виявлення загроз, яке виявило їх низьку ефективність при роботі з метаданими, оскільки ці методи здебільшого орієнтовані на виявлення шкідливих елементів у візуальній частині файлу, а не на аналіз супровідної структури. Розроблена концептуальна модель, орієнтуючись на зазначені характеристики, продемонструвала значний потенціал для ефективного виявлення аномалій і прихованого шкідливого коду в метаданих. Цей підхід дозволяє знижувати кількість помилкових спрацювань, оскільки фокусується не лише на виявленні очевидних відхилень, але й на більш тонких змінах у структурному шарі зображень. Висновки підтверджують, що аналіз супровідної інформації є важливим інструментом для виявлення нових форм загроз. Практична значимість дослідження полягає в можливості використання запропонованої концепції як основи для побудови спеціалізованих систем моніторингу та попередження кіберінцидентів

Ключові слова: стеганографія; шкідливе програмне забезпечення; цифрові зображення; машинне навчання; аномалії в метаданих

BULLETIN of Cherkasy State Technological University

Journal homepage: https://bulletin-chstu.com.ua/en

Vol. 30 No. 2. 2025

UDC 681.3.07
DOI: 10.62660/bcstu/2.2025.88

Optimising web interface performance using Amdahl's law

Oleh Prus*
Postgraduated Student
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0009-0007-2514-9871
Volodymyr Maidaniuk
PhD in Technical Sciences, Associate Professor
Vinnytsia National Technical University
21021, 95 Khmelnytske Shose Str., Vinnytsia, Ukraine
https://orcid.org/0000-0002-3345-2578

Abstract. The article discussed the construction and application of generalised mathematical models based on Amdahl's law to determine the maximum possible acceleration of web interfaces, taking into account their key features. An extension of the classical approach was proposed by including asynchronous processes, multi-level caching mechanisms, and dynamic resource loading methods in the model, which allows for a more accurate assessment of the cumulative impact of various optimisations on performance. In particular, the feasibility of taking into account asynchronous data exchange was justified, which allows processing requests in parallel and avoiding blockages in the process of updating content. A formula has been developed that takes into account the effectiveness of client and server caches and provides a quantitative assessment of the reduction in response time when reusing already loaded data. Particular attention was focused on step-by-step content retrieval techniques, where the initial page load was minimised by deferring the addition of individual scripts, images or styles, which speeds up the initial display of important content and makes the interface more responsive to user actions. In addition, the impact of a comprehensive combination of optimisation strategies on web interface performance was considered, and a corresponding generalised model was proposed, which uses an interdependence coefficient to determine the extent to which one optimisation enhances or, conversely, negates the effect of another. This makes it possible to predict the total performance gain and compare the cost of implementing several solutions with the potential time savings. The proposed formalised approach can serve as a basis for creating automated tools for evaluating web interface performance, integrated into the development process. Testing the model in three practical scenarios – partial rendering with API caching, JavaScript minification with a content delivery network (CDN), and code splitting with server-side caching – yielded performance gains of 1.87 ×, 1.55 ×, and 1.64 ×, respectively, which was fully consistent with theoretical predictions. The data obtained confirmed the ability of the *R* interdependence coefficient to accurately reflect the synergy or overlap of optimisation effects and makes the model suitable for pre-selecting the most effective acceleration strategies at the CI/CD audit stage

Keywords: mathematical models; asynchronous execution; data caching; dynamic loading; combined strategies; performance prediction

Article's History: Received: 03.02.2025; Revised: 19.05.2025; Accepted: 16.06.2025.

Suggested Citation:
Prus, O., & Maidaniuk, V. (2025). Optimising web interface performance using amdahl's law. *Bulletin of Cherkasy State Technological University*, 30(2), 88-96. doi: 10.62660/bcstu/2.2025.88.

*Corresponding author

INTRODUCTION

Optimising the performance of web interfaces is one of the priorities of modern web development, as speed and ease of interaction with the user environment directly affect visitor retention and the business performance of digital platforms. In the design of high-load web systems, the integration of caching mechanisms, asynchronous requests, and dynamic resource loading has become a key area of development, but there is still no universal methodology for the formalised evaluation of the effectiveness of these optimisations. The relevance of the problem was confirmed by the growing demands for stable service operation amid rapid increases in data volumes and user numbers.

Researchers such as M.H. Kundos *et al*. (2024) were paying considerable attention to the problem of reducing page load times and minimising rendering delays, which is due to the rapid growth in the complexity of web applications. Some researchers focused on the empirical analysis of individual acceleration techniques, such as partial rendering or the use of distributed network infrastructure. CDN, however, lacks generalised mathematical models that would allow accurate prediction of the cumulative impact of various optimisations at the planning stage. R. Abounacer *et al*. (2023) believed that in this context, Amdahl's law, which is traditionally used to estimate the maximum possible acceleration of a system after optimising a specific part of it, offers a promising theoretical basis. However, the classical formulation does not take into account a number of factors characteristic of web applications, in particular asynchronous execution, multi-level caching, and load variability caused by user behaviour dynamics. In the field of web technologies, this approach is used to analyse the impact of various performance improvement methods, including partial rendering, data caching, and network request optimisation, as can be seen in the work of A. Nair *et al*. (2024). This study reviewed key research that applies Amdahl's law in the context of web development, evaluating its application for determining the effectiveness of optimisations in multi-project environments.

The article by L. Mitton (2023) examined Amdahl's law formula and its application for predicting the potential acceleration of tasks when improving system resources. The author emphasised that even when additional processors are added, the maximum acceleration is limited by the part of the task that cannot be parallelised. The limitations of Amdahl's law were also discussed, in particular its unsuitability for cases with variable workloads. The article provided a general understanding of Amdahl's law, but did not consider its application in the context of web interfaces. Further research may focus on adapting this law to evaluate optimisations in web applications, taking into account the specifics of their architecture and user interaction. In addition, the extension of the classic Amdahl formula to multiple resources, proposed by C. Poolla & R. Saxena (2022), confirmed the accuracy of the multidimensional regression model (≈ 95% cross-validation) for predicting total acceleration when scaling multiple subsystems simultaneously. This approach formed the basis of the interdependence coefficient R and allowed for a formal assessment of the synergy or conflict between optimisation strategies.

At the same time, some researchers are focusing on specific aspects of improving the performance of modern web solutions that could be integrated into the extended formulation of Amdahl's law. For example, I. Malavolta *et al*. (2020) and T. Majchrzak *et al*. (2018) studied the impact of caching on the efficiency and overall performance of progressive web applications (PWA), emphasising the advisability of considering cache mechanisms as an important component of optimisation. At the low-level optimisation level, J. Ren *et al*. (2024) proposed JSTUNER, an ML-oriented auto-tuning system that provides an average acceleration of JS code execution by ×1.62 on mobile devices by automatically selecting the optimal sequences of Chrome V8 compiler flags. This confirmed that intelligent approaches can complement classic manually tuned optimisations. A comprehensive review by J. Vepsäläinen *et al*. (2024) showed that the average mobile page size has increased sixfold over the past decade, and therefore the need for systematic optimisation methods is only growing. The authors highlighted code splitting, caching, and CDN as the most effective areas, which correlates with the topic of this work. Regarding single-page applications, M. Kothapalli (2022) empirically showed that combining code splitting and lazy loading reduces the first rendering time of SPA by 40%, and additional caching ensures stable offline operation.

Thus, modern works from 2020-2024 cover the entire spectrum – from ML-tuning of compilers and SPA portfolios to CDN edge platforms and analytical extensions of Amdahl's law – which justifies the relevance of the generalised model developed in the article for predicting the combined performance gain of web interfaces. The generalisation of the theoretical provisions of Amdahl's law with modern research developments in the field of caching, asynchronous operations, dynamic loading and scalable testing makes it possible to form a more comprehensive model for optimising web interfaces. This confirms the need for further theoretical and practical research aimed at improving the tools for determining the maximum possible acceleration of web applications of different architectures and scales. Therefore, despite the fundamental nature of Amdahl's law, its application in the context of web interface optimisation requires further research. The development of mathematical models and practical recommendations for the evaluation of effectiveness of optimisations in multi-project web environments will increase application productivity and improve user experience, as noted in the work of O. Prus *et al*. (2024).

In view of the above, the aim of this study was to test Amdahl's law adapted to the specifics of web development, in particular to build an extended mathematical model for evaluating interface performance, taking into account asynchronous processes, dynamic loading, and caching mechanisms. To achieve this goal, the following tasks were defined: to analyse modern approaches to optimising the performance of web applications and their empirical effectiveness; to adapt the existing formulation of Amdahl's law, taking into account the specifics of asynchronous and cached operations; to develop a generalised model that will allow evaluating the impact of several optimisations simultaneously and determining the interdependence between them; to verify the validity of theoretical propositions on real examples, conduct numerical experiments and compare the results obtained with data published in recent scientific publications.

The scientific novelty lies in the fact that, for the first time, a formalised approach to measuring the combined effect of several parallel optimisations in a web environment has been proposed, filling the gap between empirical testing methods and the theoretical acceleration model. The use of the adapted Amdahl's law will allow developers and operations engineers to make more informed decisions about choosing performance improvement strategies, taking into account the potential limitations and interactions of different approaches.

MATERIALS AND METHODS

In the classic formulation, Amdahl's law estimates the maximum possible acceleration of a system after optimising its part, which is determined by the equation:

$$S = \frac{1}{(1-P)+\frac{P}{N}}, \quad (1)$$

where S – total system acceleration; P – the proportion of the program to be optimised; N – the acceleration coefficient of the optimised part (GeeksForGeeks, 2023). However, the classic formulation does not take into account the specifics of web interfaces, where performance depends not only on serial and parallel execution, but also on asynchronous processes, caching, and dynamic resource loading. For the correct application of Amdahl's law in web development, it is necessary to expand its model by including additional factors (Junghans *et al.*, 2017).

In the traditional case, acceleration is only possible through parallel computing. However, in web interfaces, a significant proportion of processes are asynchronous requests to the server (e.g., AJAX, Fetch API, WebSockets) and deferred loading of components (e.g., lazy loading) (Ahmed *et al.*, 2017, Oh *et al.*, 2025). When using asynchronous acceleration mechanisms, a web application can be described as:

$$S_{async} = \frac{1}{(1-P_{sync}-P_{async})+\frac{P_{sync}}{N}+\frac{P_{async}}{A}}, \quad (2)$$

where S_{async} – total system acceleration, taking into account asynchronous acceleration mechanisms; P_{sync} – proportion of serial code (not subject to optimisation); P_{async} – proportion of code executed asynchronously; N – acceleration coefficient of the optimised part; A – acceleration coefficient due to asynchronous execution (determined by the efficiency of parallel request processing).

This modification made it possible to more accurately predict the performance of the web interface when using asynchronous API requests, distributed data processing, and dynamic content updates. Caching is one of the main mechanisms for optimising the performance of web applications. Its use has made it possible to reduce the number of requests to the server, reducing page loading and rendering times. According to G. Araújo *et al.* (2024), caching is analysed at three levels: client caching (browser, Service Workers), server caching (Memcached, Redis), and CDN caching (Cloudflare, Akamai).

The impact of caching can be assessed by introducing a cache efficiency coefficient C, which determines what proportion of requests reuse cached data. Taking caching into account, Amdahl's law has been modified as follows:

$$S_{cache} = \frac{1}{(1-P_{cache}-P_{non-cache})+\frac{P_{cache}}{C}+\frac{P_{non-cache}}{N}}, \quad (3)$$

where S_{cache} – total system acceleration taking into account caching; P_{cache} – the proportion of requests that can be processed by the cache; $P_{non\text{-}cache}$ – the proportion of requests that need to be processed by the server; N – the acceleration coefficient of the optimised part; C – the cache efficiency coefficient, which shows how much faster cached requests are processed. According to a study by M. Dyvak & O. Kindzerskyi (2024), the inclusion of caching in the model made it possible to assess the impact of different caching strategies on web interface performance and to draw conclusions about the feasibility of implementing cache mechanisms in specific scenarios (2024).

As noted by R. Felani *et al.* (2020), modern web applications use dynamic content loading, which reduces the initial page load time. These methods include: lazy loading of images, scripts, and styles; code splitting, which allows JavaScript modules to be loaded only when needed; and progressive rendering, which speeds up the display of the first page content. The impact of these mechanisms on performance can be represented as:

$$S_{dynamic} = \\ = \frac{1}{(1-P_{dynamic}-P_{sync}-P_{async})+\frac{P_{sync}}{N}+\frac{P_{async}}{A}+\frac{P_{dynamic}}{D}}, \quad (4)$$

where $S_{dynamic}$ – total system acceleration, taking into account dynamic content loading; $P_{dynamic}$ – the proportion of resources that are loaded dynamically; P_{sync} – the proportion of serial code (not subject to optimisation); P_{async} – the proportion of code executed asynchronously; N – the acceleration coefficient of the optimised part;

A – acceleration coefficient due to asynchronous execution (determined by the efficiency of parallel request processing); D – acceleration coefficient due to dynamic loading (determined by the efficiency of distributed loading resources) (Ma, 2024).

Practical significance: including dynamic loading in the Amdahl's law model helps figure out how effectively lazy loading, code splitting, and progressive rendering techniques are used in a specific web app. Taking all the factors into account, the generalised Amdahl's law model (5) for web apps looks like this:

$$S_{web} = \frac{1}{(1-P_{sync}-P_{async}-P_{cache}-P_{dynamic})+\frac{P_{sync}}{N}+\frac{P_{async}}{A}+\frac{P_{cache}}{C}+\frac{P_{dynamic}}{D}}. \quad (5)$$

This model made it possible to quantitatively assess the impact of asynchronous execution, caching, and dynamic resource loading on the overall performance of a web application. The extended mathematical model of Amdahl's law takes into account key features of web interfaces that affect performance, in particular: asynchronous execution, caching mechanisms, and dynamic resource loading. The proposed formula made it possible to quantitatively assess the effectiveness of various optimisation methods and predict their impact on the performance of web applications.

The next stage of the research was to model combined optimisations, which made it possible to determine the overall effect of simultaneously applying several optimisation methods in web applications. In real web applications, it is rare to use only one optimisation method. More often, combinations of optimisations are used, for example: partial rendering combined with API request caching (reducing delays when updating content); combining JavaScript files and using CDN (reducing network load); browser-level caching and server-side caching (minimising the processing of repeat requests).

In its classic formulation, Amdahl's law evaluates the effectiveness of only one optimisation at a time. However, when several acceleration methods are applied simultaneously, their effect may not be simply cumulative, but interdependent, which makes it difficult to predict the effectiveness of combined optimisations. Thus, the main task in solving this problem was to develop a mathematical model that allows evaluating the overall effect of the simultaneous application of several optimisation methods. The next step was to formalise combined optimisations within the Amdahl's law. Two optimisations acting on different parts of the system are considered as follows: optimisation accelerates the share with a coefficient ; optimisation accelerates the share with a coefficient . If these optimisations are independent of each other, their combined effect can be estimated by the formula:

$$S_{combined} = \frac{1}{(1-P_A-P_B)+\frac{P_A}{N_A}+\frac{P_B}{N_B}}. \quad (6)$$

However, in real web applications, optimisations often overlap. For example, if both API request caching and partial rendering are used, the overall effect may be lower than expected due to the overhead of synchronisation between the two mechanisms. To account for the interdependence between optimisations, an interdependence coefficient was introduced, which determines how strongly one optimisation affects the effectiveness of another. Then the general formula will look like this:

$$S_{combined} = \frac{1}{(1-P_A-P_B+RP_AP_B)+\frac{P_A}{N_A}+\frac{P_B}{N_B}}, \quad (7)$$

where R – interdependence coefficient (values from -1 to 1). If $R > 0$, optimisations reduce each other's effectiveness. If $R < 0$, optimisations complement each other and provide an additional effect. If $R = 0$, optimisations act independently. This model allows for performance evaluation to be adjusted depending on whether the optimisations interact with each other or work autonomously. The next step in the study was to create a generalised model for several optimisations. If we expand the model for n optimisations, the following was obtained:

$$S_{multi} = \frac{1}{(1-\sum_{i=1}^{n}P_i+\sum_{i=1}^{n}\sum_{i=i+1}^{n}R_{ij}P_iP_j)+\sum_{i=1}^{n}\frac{P_i}{N_i}}, \quad (8)$$

where P_i – the part of the process that is accelerated by the i-th optimisation; N_i – acceleration coefficient for optimisation i; R_{ij} – coefficient of interdependence between optimisations i and j. The resulting formula made it possible to evaluate the effectiveness of any number of optimisations, taking into account their mutual influence.

During testing of the created mathematical model, it became necessary to demonstrate how the combined use of several optimisation methods affects the overall performance of a web application. For this purpose, two examples were selected, which present different scenarios of acceleration combinations (partial rendering with caching and JavaScript minification in combination with CDN). Such a comparison makes it possible to trace the conditions under which optimisations reinforce each other, and under which their effects may overlap in a certain way or even reduce the overall result. In this study, the parameters N and P were evaluated using web application performance profiling tools, including Chrome DevTools, Lighthouse, and WebPageTest (Potdar *et al.*, 2020). The P code share was defined as the ratio of the execution time of a specific processing stage (e.g., rendering or waiting for a response from the API) to the total page load and display time.

RESULTS AND DISCUSSION

For the practical application of the mathematical model developed on the basis of Amdahl's Law, an important step was the quantitative determination of the proportion of code to be optimised (P), as well as the acceleration coefficient of this proportion after the

implementation of the corresponding method (*N*). The reliability of the forecast of the overall acceleration of the system directly depends on the accuracy of these parameters. In the example with partial rendering, the value = 0.4 was obtained as the average value based on several measurements, where the full rendering process took about 400 ms out of a total time of 1,000 ms.

The acceleration factor was calculated as the ratio of the execution time of the corresponding stage before and after the implementation of the optimisation. For example, for partial rendering, which reduced the visual update time from 400 ms to approximately 133 ms, the acceleration factor was:

$$N_A = \frac{400}{133} \approx 3$$

Similarly, for caching API requests, where the server response time decreased from 300 ms to 150 ms, the acceleration factor was estimated as:

$$N_B = \frac{300}{150} = 2$$

Thus, the values of parameters *P* and *N* were obtained empirically, based on instrumental measurement and timing analysis. These values were used in further calculations to evaluate both the individual effect of each optimisation and their combined effect. Example 1: Partial rendering + API request caching. Partial rendering reduces the need for a full page refresh, optimising $P_A = 0.4$ (40% of rendering time). Caching API requests reduces the waiting time for data from the server, optimising $P_B = 0.3$ (30% of request time). Partial rendering helps caching (since fewer requests are sent to the server), so $R = -0.1$.

The values are substituted into the formula (7):

$$S_{combined} = \frac{1}{(1-0.4-0.3+(-0.1\times0.4\times0.3))+\frac{0.4}{3}+\frac{0.3}{2}}.$$

The calculations yielded the following:

$$S_{combined} \approx 1.87.$$

Thus, when each optimisation is applied separately, the expected acceleration would be approximately 1.67-1.75 times, but due to the positive effect of interaction (caching enhances rendering efficiency), the overall acceleration increases to 1.87 times. This example demonstrated that partial rendering combined with API request caching provides a greater total performance gain than each method separately (1.87 × vs. ~1.7 ×). This synergistic effect has also been confirmed in the works of other authors. In particular, V. Jain (2022) noted that the simultaneous use of lazy loading and JS code splitting technologies reduced web page loading time by up to 40% – significantly more than expected from each technique separately. Thus, the results were consistent with the current case: optimisations that complement each other (partial interface updates and data caching) can provide an enhanced cumulative effect. Moreover, a study by J. John (2024) showed that effective caching alone significantly speeds up the performance of a web system, reducing latency and increasing server throughput. This is consistent with choice of caching as one of the methods: in Example 1, caching API requests reduced the time to retrieve data by approximately half, which is consistent with the researcher's conclusions about the impact of cache on system latency. Thus, the literature data confirmed that combining rendering optimisation (on the client side) with request caching (on the server side) is an effective approach to speeding up content display.

Example 2: JavaScript code optimisation + CDN usage. Minification and merging of JavaScript files optimised $P_A = 0.25$. Using a CDN reduces the loading time of static resources, optimising $P_B = 0.35$. Since both optimisations reduce the overall load on the network, the effect is partially overlapping, so $R = 0.2$ (some files are already cached, so minification does not have the full effect).

Calculation:

$$S_{combined} = \frac{1}{(1-0,25-0,35+(0,2\times0,25\times0,35))+\frac{0,25}{2,5}+\frac{0,35}{3}}.$$

Result:

$$S_{combined} \approx 1,55.$$

In other words, in this example, it was found that minification and concatenation of JavaScript files together with the use of CDN provide a smaller total gain (≈ 1.55×) than would be expected from the independent effects of each method. The reason is the partial overlap of effects: both methods are aimed at accelerating the delivery of static resources and reducing network delays, so their effects are partially duplicated. A similar phenomenon – a drop in growth when optimisations overlap – has been noted by other researchers. For example, M.N.Y. Utomo *et al.* (2025) combined the use of a distributed CDN and caching (via Varnish) in an experiment to improve the performance of a web platform. The result was a significant improvement: throughput increased by 175%, average response time was reduced by more than half (54%), and network latency was reduced by 82%. These figures confirmed the effectiveness of using CDN and caching simultaneously. However, the authors also noted that both methods improve similar aspects of the system's performance (the speed of content delivery to the user). The current results clarified this point: when optimisations affect the same segment of the loading time (in this case, the time of file transfer and loading), the performance gain from combining them may be less than the total sum of the gains. Accordingly, when developing an optimisation strategy, it is worth considering possible areas of

overlap between methods to avoid a situation where the resources invested have a decreasing effect (the phenomenon of diminishing returns). Example 2 shows just such a case of partial duplication.

Example 3: Code splitting + server caching. Code splitting involves dividing the web application code into smaller modules that are loaded only when needed (for example, when a user accesses a specific section or function). Server caching reduces the time it takes to respond to requests, as repeat requests are processed using data already prepared in the cache on the server side. In the proposed scenario, it is assumed that: Code splitting covers 30% of the code (= 0.3) with an acceleration factor of = 3. Server caching affects 25% of operations (= 0.25) and provides a fourfold gain = 4). The interdependence coefficient between optimisations is = −0.05, which indicates mutual reinforcement: reducing the total code volume reduces the number of requests that require caching, and the presence of a cache, in turn, speeds up the loading of split modules. The total acceleration is calculated using the already known formula (7). The corresponding values are substituted as follows:

$$S_{combined} = \frac{1}{(1-0.3-0.25+(-0.05\times0.3\times0.25))+\frac{0.3}{3}+\frac{0.5}{4}}.$$

Calculations give approximately:

$$S_{combined} \approx 1.64.$$

Separate use of code splitting would speed up the system by approximately 1.25 times, and server-side caching alone by about 1.23 times. Thanks to the negative interdependence coefficient ($R < 0$), these two optimisation methods complement each other and provide a noticeably better cumulative result. In cases where $R > 0$ (for example, if parts of the functionality or data are duplicated, creating additional load), the overall acceleration could decrease due to mutual overlap of effects.

Thus, the combination of code splitting and server caching can yield a significant synergistic gain, but achieving optimal results depends on a thorough analysis of each method's impact on different parts of the code, as well as considering their degree of interdependence. The results obtained indicate that the overall gain from combining several optimisations depends not only on the proportion of code that can be accelerated but also on their degree of interdependence. A negative interdependence coefficient ($R < 0$) can ensure a better cumulative effect if optimisations complement each other and do not duplicate functionality. At the same time, a positive coefficient ($R > 0$) can reduce the final acceleration when optimisation methods partially overlap. Taking these conclusions into account, more effective optimisation strategies can be developed, aligning the nature of interaction between the methods used. This, in turn, allows for predicting potential performance gains and rationalising the costs of implementing web application performance improvement solutions.

Thus, the overall acceleration is ~ 1.64×, while the separate application of code splitting would accelerate rendering by ~ 1.25×, and server caching would reduce response generation time by ~ 1.23×. In other words, these methods complemented each other better than initially expected. A similar conclusion can be found in the work of J. van Riet & T.A. Ghaleb (2023), which described the sequential optimisation of a large web application by implementing 13 different improvements over a period of four months. As a result of this comprehensive approach, the first contentful paint metric improved by almost 98% (for the desktop version of the site), and other performance metrics also improved significantly. This case clearly demonstrated that combining several optimisation techniques (including those similar to those presented in this study: caching, resource loading optimisation, etc.) leads to a significant acceleration of the web application. In addition, the review by H.O. Ekpobimi *et al.* (2024) emphasised the need for an integrated approach: the authors noted that the simultaneous implementation of front-end optimisation strategies (such as code splitting, lazy loading, script minimisation) together with back-end improvements (different levels of data caching) allows for the greatest overall efficiency. The data obtained confirmed this thesis using the example of combining code splitting and server caching.

In the study by L. Kvurt & L. Tsyhylyk (2009), Amdahl's Law was considered a fundamental principle for evaluating the potential acceleration of systems through the optimisation of their components. Its application allows for theoretically determining the ultimate effectiveness of any optimisation, taking into account the limitations imposed by the unoptimised portion of the code. Importantly, to quantitatively predict such an effect, a model based on Amdahl's Law was applied, which considers the proportion of optimised code P and the acceleration N of each method, as well as their interdependence coefficient R. A similar modelling approach was proposed by C. Poolla & R. Saxena (2022), who extended the classical Amdahl's Law for the case of several simultaneous improvements and confirmed its validity with experimental data. The obtained results aligned with the conclusion that mathematical prediction of multifactor acceleration is possible and quite accurate. The distinction of the current study lies in the introduction of the R parameter to assess the nature of optimisation interaction. This allowed for explaining why in some cases (Examples 1 and 3) a greater than additive gain (synergy, $R < 0$) was observed, while in others (Example 2) – less (overlap effect, $R > 0$). This approach has not yet received widespread coverage in works, where most attention has been paid either to single optimisations or their empirical combination.

M. Hevery (2025) examined the application of Amdahl's Law in JavaScript code optimisation. The author

emphasised the importance of focusing on optimising those parts of the code that significantly affect the overall application performance, instead of expending resources on less significant fragments. The interaction between modern processors and virtual machines in the context of JavaScript execution was also discussed. Although the author provided valuable insights into JavaScript optimisation, the work did not propose specific mathematical models for evaluating the effectiveness of optimisations. The development of such models will allow for quantitatively assessing the impact of optimisations on web interface performance, which is critical in multi-project environments. Therefore, the contribution of this work lies in confirming the presence of both positive interaction and conflicts between different acceleration methods, as reported in previous studies, and in proposing a formalised method for evaluating this phenomenon. This has enabled developers to evaluate in advance the potential effect of combining several optimisations and to make informed decisions regarding web application performance improvement based on both theoretical calculations and best practices known from the literature.

CONCLUSIONS

Amdahl's classic law does not take into account the interaction between several optimisations, which makes it impossible to apply it directly to assess the cumulative acceleration during the simultaneous implementation of different methods. The proposed generalised model made it possible to assess the cumulative increase in performance, taking into account the interdependence between optimisations. The interdependence coefficient R, added to the model, made it possible to determine whether the optimisation methods interact synergistically and reinforce each other, or whether they overlap and reduce the overall acceleration. The model can be used for automated analysis of web application performance in the early stages of design, which contributes to a more rational allocation of resources and a well-founded choice of interface acceleration strategies.

The study made it possible to adapt Amdahl's law to take into account asynchronous requests, multi-level caching, and dynamic resource loading. This integration makes the performance assessment of web interfaces more accurate, as it covers specific aspects of modern technologies. The generalised model of combined optimisations considered several acceleration methods at once and made it possible to calculate their combined effect, taking into account possible interactions, which is reflected in the R coefficient. If the methods do not complement each other, the overall gain may be lower than the sum of the individual effects, while their synergy, on the contrary, can increase the final performance indicator.

Empirical examples have shown that combining caching with partial rendering can enhance the result by reducing the number of network calls and page load times. In this case, code splitting covers 30% of the code ($P_A = 0.3$) and provides acceleration with a coefficient of $N_A = 3$. Server caching affects 25% of operations ($P_B = 0.25$), resulting in a fourfold increase in performance ($N_B = 4$). The interdependence coefficient between these optimisations is $R = -0.05$. On the other hand, combining script minification with CDN may result in a more modest increase, as both methods partially overlap in their impact on network operations. The proposed formalised approach made it possible to predict such scenarios in advance and select the most appropriate optimisation strategies for each specific type of application (high-load service, content site, SPA, etc.).

The practical value of the results obtained lies in the ability to predict the effectiveness of web application optimisations before their implementation, as well as in a better understanding of the interaction of different approaches to acceleration. Web developers and DevOps engineers can make informed choices about the best strategies for improving the interface, avoiding unnecessary expenditure of resources on methods that do not provide significant improvement or conflict with each other. Further research should focus on verifying the models in real-world web system operating conditions, particularly with large amounts of data and involving large categories of users. This will provide additional evidence of the correctness and universality of the developed approach, as well as contribute to its further development and improvement.

ACKNOWLEDGEMENTS

None.

FUNDING

None.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Abounacer, R., Afdel, K., & Bouaouda, A. (2023). Resource utilization and cost implications of container live migration in clouds: An approach performed on Amazon Web Services (AWS). *Research Square*. doi: 10.21203/rs.3.rs-3286731/v1.

[2] Ahmed, F., Erman, J., Ge, Z., Liu, A. X., Wang, J., & Yan, H. (2017). Detecting and localizing end-to-end performance degradation for cellular data services based on TCP loss ratio and round trip time. *IEEE/ACM Transactions on Networking*, 25(6), 3709-3722. doi: 10.1109/TNET.2017.2761758.

[3] Araújo, G.R., Gomes, R., Ferrão, P., & Gomes, M.G. (2024). Optimizing building retrofit through data analytics: A study of multi-objective optimization and surrogate models derived from energy performance certificates. *Energy and Built Environment*, 5(6), 889-899. doi: 10.1016/j.enbenv.2023.07.002.

[4] Dyvak, M., & Kindzerskyi, O. (2024). Investigation of the efficiency of parallel computational scheme for identification of interval discrete models based on swarm intelligence. *Herald of Khmelnytskyi National University. Technical Sciences*, 331(1), 29-37. doi: 10.31891/2307-5732-2024-331-3.
[5] Ekpobimi, H.O., Kandekere, R.C., & Fasanmade, A.A. (2024). Conceptual framework for enhancing front-end web performance: Strategies and best practices. *Global Journal of Advanced Research and Reviews*, 2(1), 99-107. doi: 10.58175/gjarr.2024.2.1.0032.
[6] Felani, R., Al-Azam, M.N., Adi, D.P., & Widodo, A. (2020). Optimizing virtual resources management using Docker on cloud applications. *Indonesian Journal of Computing and Cybernetics Systems*, 14(3), article number 319. doi: 10.22146/ijccs.57565.
[7] GeeksForGeeks. (2023). *Computer organization. Amdahl's law and its proof*. Retrieved from https://www.geeksforgeeks.org/computer-organization-amdahls-law-and-its-proof.
[8] Hevery, M. (2025). *Amdahl's law. Bare metal JavaScript: The JavaScript virtual machine*. Retrieved from https://frontendmasters.com/courses/javascript-cpu-vm/amdahl-s-law.
[9] Jain, V. (2022). Optimizing web performance with lazy loading and code splitting. *International Journal of Core Engineering & Management*, 7(3), 193-199. doi: 10.5281/zenodo.14956631.
[10] John, J. (2024). *Optimizing application performance: A study on the impact of caching strategies on latency reduction*. Retrieved from https://www.researchgate.net/publication/385916660_OPTIMIZING_APPLICATION_PERFORMANCE_A_STUDY_ON_THE_IMPACT_OF_CACHING_STRATEGIES_ON_LATENCY_REDUCTION.
[11] Junghans, C., Agarwal, A., & Delle Site, L. (2017). Computational efficiency and Amdahl's law for the adaptive resolution simulation technique. *Computer Physics Communications*, 215, 20-25. doi: 10.1016/j.cpc.2017.01.030.
[12] Kothapalli, M. (2022). Performance analysis of single page applications. *International Journal of Science and Research (IJSR)*, 11(1), 1631-1635. doi: 10.21275/SR24529184457.
[13] Kundos, M.H., Solovey, L.Ya., Hrysyuk, A.V., & Bahniuk, O.M. (2024). Efficiency and multi-threading of parallel calculations in systems programming. *Taurida Scientific Herald. Series: Technical Sciences*, 5, 60-64. doi: 10.32782/tnv-tech.2024.5.6.
[14] Kvurt, L., & Tsyhylyk, L. (2009). The use of Amdahl's and Gustafson's laws in evaluating the acceleration factor in multiprocessor systems. *Measurement Techniques and Metrology*, 70, 55-56.
[15] Ma, J. (2024). A high performance computing web search engine based on big data and parallel distributed models. *Informatica*, 48(20), 27-38. doi: 10.31449/inf.v48i20.6776.
[16] Majchrzak, T.A., Biørn-Hansen, A., & Grønli, T.-M. (2018). Progressive web apps: The definite approach to cross-platform development? In *51st Hawaii international conference on system sciences (HICSS 2018)* (pp. 1-10). Hawaii: Hilton Waikoloa Village. doi: 10.24251/HICSS.2018.718.
[17] Malavolta, I., Chinnappan, K., Jasmontas, L., Gupta, S., & Soltany, K.A.K. (2020). Evaluating the impact of caching on the energy consumption and performance of progressive web apps. In *Proceedings of the IEEE/ACM 7th international conference on mobile software engineering and systems* (pp. 109-119). New York: Association for Computing Machinery. doi: 10.1145/3387905.3388593.
[18] Mitton, L. (2023). *Amdahl's law: Understanding the basics*. Retrieved from https://www.splunk.com/en_us/blog/learn/amdahls-law.html.
[19] Nair, A.M., Sivaiswarya, C.K., Sidharth, S., Visakh, K.K., & Joy, J. (2024). Dockerized application with web interface. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 10(2), 412-419. doi: 10.32628/CSEIT243646.
[20] Oh, S., Kwon, Y., & Lee, J. (2025). Optimizing real-time object detection in a multi-neural processing unit system. *Sensors*, 25(5), article number 1376. doi: 10.3390/s25051376.
[21] Poolla, C., & Saxena, R. (2022). On extending Amdahl's law to learn computer performance. *ArXiv*, 2110, article number 07822. doi: 10.48550/arXiv.2110.07822.
[22] Potdar, A.M., Narayan, D.G., Kengond, S., & Mulla, M.M. (2020). Performance evaluation of Docker container and virtual machine. *Procedia Computer Science*, 171, 1419-1428. doi: 10.1016/j.procs.2020.04.152.
[23] Prus, O.V., Maidaniuk, V.P., & Arseniuk, I.R. (2024) March analysis of tools for multiproject environments management: Optimization of the software development. *Scientific Works of VNTU*, 1, 29-36. doi: 10.31649/2307-5376-2024-1-29-36.
[24] Ren, J., Gao, L., & Wang, Z. (2024). JavaScript performance tuning as a crowdsourced service. *IEEE Transactions on Mobile Computing*, 23(5), 6116-6132. doi: 10.1109/TMC.2023.3316167.
[25] Utomo, M.N.Y., Tungadi, E., & Khartika, W. (2025). Enhancing web performance for e-learning platform using content delivery network (CDN) and varnish cache. *Journal of Information Systems and Informatics*, 7(1), 831-847. doi: 10.51519/journalisi.v7i1.993.
[26] van Riet, J., Ghaleb, T.A. (2023). Optimise along the way: An industrial case study on web performance. *Journal of Systems and Software*, 198, article number 111593. doi: 10.1016/j.jss.2022.111593.
[27] Vepsäläinen, J., Hellas, A., & Vuorimaa, P. (2024). Overview of web application performance optimization techniques. *ArXiv*, 2412, article number 07892. doi: 10.48550/arXiv.2412.07892.

Оптимізація продуктивності веб-інтерфейсів із використанням закону Амдала

Олег Прус
Аспірант
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0009-0007-2514-9871
Володимир Майданюк
Кандидат технічних наук, доцент
Вінницький національний технічний університет
21021, вул. Хмельницьке шосе, 95, м. Вінниця, Україна
https://orcid.org/0000-0002-3345-2578

Анотація. У статті розглянуто побудову та застосування узагальнених математичних моделей на основі закону Амдала для визначення максимально можливого прискорення роботи веб-інтерфейсів з урахуванням їхніх ключових особливостей. Запропоновано розширення класичного підходу шляхом включення до моделі асинхронних процесів, механізмів різнорівневого кешування та методів динамічного завантаження ресурсів, що дає змогу точніше оцінювати сумарний вплив різних оптимізацій на швидкодію. Зокрема, обґрунтовано доцільність урахування асинхронного обміну даними, який дає змогу обробляти запити паралельно й уникати блокувань у процесі оновлення контенту. Розроблено формулу, що враховує ефективність клієнтського і серверного кешу та надає кількісну оцінку скорочення часу відгуку при повторному використанні вже завантажених даних. Особливу увагу зосереджено на методиках покрокового отримання вмісту, коли початкове завантаження сторінки мінімізується за рахунок відкладеного додавання окремих скриптів, зображень або стилів, що дає змогу пришвидшити початкове відображення важливого контенту і зробити інтерфейс більш чутливим до дій користувача. Окрім того, розглянуто вплив комплексного поєднання оптимізаційних стратегій на продуктивність веб-інтерфейсу та запропоновано відповідну узагальнену модель, яка за допомогою коефіцієнта взаємозалежності дозволяє визначити, наскільки одна оптимізація підсилює або, навпаки, нівелює дію іншої. Це надає можливість прогнозувати сумарний приріст швидкодії та зіставляти витрати на впровадження кількох рішень із потенційною економією часу. Запропонований формалізований підхід може стати основою для створення автоматизованих засобів оцінки продуктивності веб-інтерфейсів, інтегрованих у процес розробки. Перевірка моделі в трьох практичних сценаріях – частковий рендеринг з кешуванням API, мініфікація JavaScript з content delivery network (CDN) та розділення коду з кешуванням на стороні сервера – дала приріст продуктивності відповідно в 1,87 ×, 1,55 × та 1,64 ×, що повністю відповідає теоретичним прогнозам. Отримані дані підтверджують здатність коефіцієнта взаємозалежності R точно відображати синергію або накладання ефектів оптимізацій і роблять модель придатною для попереднього вибору найефективніших стратегій прискорення на етапі CI/CD-аудиту

Ключові слова: математичні моделі; асинхронне виконання; кешування даних; динамічне завантаження; комбіновані стратегії; прогнозування швидкодії

ВІСНИК
Черкаського державного технологічного університету

Науковий збірник

Том 30, № 2. 2025

Заснований у 1996 р. Виходить чотири рази на рік

Оригінал-макет видання виготовлено у редакційно-видавничому відділі
Черкаського державного технологічного університету

Відповідальний редактор:
А. Лавданський

Підписано до друку 16 липня 2025 р.
Формат 60*84/8
Умов. друк. арк. 11,4
Наклад 50 прим.

Адреса видавництва:
Черкаський державний технологічний університет
18006, бульв. Шевченка, 460, м. Черкаси, Україна
Тел.: +380638013283
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/uk

BULLETIN
of Cherkasy State Technological University

Scientific Collection

Volume 30, No. 2. 2025

Founded in 1996. Published four times per year

The original layout of the publication is made in the editorial and publishing department
of Cherkasy State Technological University

Managing editor:
A. Lavdanskyi

Signed for print of July 16, 2025.
Format 60*84/8
Conventional printed pages 11.4
Circulation 50 copies

Editors Office Address:

Cherkasy State Technological University
18006, 460 Shevchenko Blvd., Cherkasy, Ukraine
Tel.: +380638013283
E-mail: info@bulletin-chstu.com.ua
https://bulletin-chstu.com.ua/en