A domain-oriented method for evaluating 5G core network software quality*

Viktor Gnatyuk[1,2,*,†] and Roman Hamretskyi[1,†]

[1] State University "Kyiv Aviation Institute", 1 Liubomyra Huzara ave., 03058 Kyiv, Ukraine
[2] State Scientific and Research Institute of Cybersecurity Technologies and Information Protection, 3/6 Maksym Zalizniak str., 03142 Kyiv, Ukraine

Abstract

This paper examines approaches to assessing the quality of fifth-generation (5G) core network software. A method is proposed that combines standard quality metrics in accordance with ISO/IEC 25010 with additional criteria relevant to 5G telecommunication systems, such as latency, scalability, energy efficiency, adaptability, and fault tolerance. A three-tier model is developed, enabling multidimensional assessment of core software considering service, functional, and infrastructure levels. Simulations were conducted to evaluate the method's effectiveness under different scenarios, including load variation, component version updates, and node failures. The proposed approach demonstrates higher precision and practical applicability compared to existing methods. Graphs, block diagrams, and experimental results are presented. Prospects for implementation and future research directions are outlined.

Keywords

software, 5G, ISO/IEC 25010, QoS, software quality

1. Introduction

The rapid development of fifth-generation (5G) mobile networks has led to increased demands on the quality of the software (SW) responsible for core network operation. The 5G core is characterized by a high degree of virtualization, flexibility, scalability, and support for heterogeneous services, ranging from massive IoT connectivity to ultra-low-latency applications. Ensuring high-quality software is critically important for the stability and reliability of the entire network.

Existing software quality assessment standards, such as ISO/IEC 25010, provide general approaches to classifying and analyzing quality characteristics including reliability, performance, and usability. However, they do not consider the unique aspects of 5G environments, such as strict latency requirements, dynamic resource allocation, and energy efficiency within virtualized environments.

Thus, there is a pressing need to develop a software quality assessment method tailored to the 5G core that would objectively evaluate the suitability of software components for deployment in modern telecommunication networks. Such a method must incorporate both traditional quality indicators and specific metrics characterizing 5G core operation.

2. Analysis of recent sources

Software quality assessment is a critically important stage in the development lifecycle, particularly for high-load systems such as the 5G core. One of the most widely recognized standards for defining and evaluating quality characteristics is ISO/IEC 25010 [1], which outlines attributes such as functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, and portability.

* CSDP'2025: Cyber Security and Data Protection, July 31, 2025, Lviv, Ukraine
* Corresponding author.
† These authors contributed equally.
viktor.hnatiuk@npp.kai.edu.ua (V. Gnatyuk); roman.hamretskyi@npp.kai.edu.ua (R. Hamretskyi)
0000-0002-4916-7149 (V. Gnatyuk); 0000-0001-5514-9783 (R. Hamretskyi)

While ISO/IEC 25010 offers a universal framework, several researchers [2–4] have emphasized the need to adapt or extend the quality model for specific domains such as telecommunication networks. For example, [2] proposes incorporating real-time and resource-dependency features, which are critical in 5G environments.

Considerable attention has also been paid to the quality evaluation of cloud-based software, which shares architectural features with the 5G core, including virtualization, microservice architecture, and automated orchestration. In [3], a methodology is proposed for assessing the quality of microservice applications, taking into account parameters such as scalability, fault tolerance, and update efficiency. Reference [4] introduces a multi-level evaluation model that considers both technical and business quality indicators.

In the context of 5G, some studies [5, 6] focus on evaluating software performance within virtualized infrastructure and network functions (NFV), emphasizing the importance of metrics such as latency, throughput, orchestration delay, and resource efficiency.

Thus, an analysis of existing approaches reveals that although general software quality models exist, they do not fully account for the specific characteristics of the 5G core. Recent research also highlights that emerging technologies, such as software-defined radio (SDR) receivers, introduce new cybersecurity risks in wireless environments, which should be reflected in domain-oriented quality evaluation methods [7–12]. This justifies the need for a unified assessment method that integrates general standards with domain-specific indicators.

3. Objective of the study

The objective of this paper is to develop a method for assessing the quality of 5G core software that is based on international standards but adapted to the requirements of modern telecommunication infrastructure. The paper provides an overview of existing approaches to software quality evaluation, formulates the requirements for a new assessment method, describes its structure, and presents an experimental validation of its effectiveness.

4. Method description

The 5G network core performs critical functions such as access control, mobility management, traffic routing, quality of service (QoS) assurance, security, and network resource orchestration. These functions are implemented through a set of software components deployed in a cloud environment using Network Function Virtualization (NFV) and Software-Defined Networking (SDN) [13, 14].

Due to the architectural complexity, high workload dynamics, and strict performance guarantees, traditional software quality assessment methods are insufficient to comprehensively evaluate the quality of the 5G core. In particular, it is essential to consider parameters such as:

- Data transmission latency (crucial for uRLLC-type services).
- Scalability (the system's ability to adapt resources to varying loads).
- Reliability (the system's ability to function without failure under critical conditions).
- Energy efficiency (especially important for distributed edge infrastructures).
- Fault and attack tolerance (including DDoS protection and self-healing).
- Integration with cloud-native and containerized infrastructures [15, 16].

Therefore, it is necessary to develop a software quality assessment method for the 5G core that:

1. Accounts for both traditional quality characteristics (based on ISO/IEC 25010) and domain-specific indicators.
2. Is adapted to the 5G Core architecture: service-oriented, containerized, and orchestrated.
3. Supports formal representation and quantitative computation of quality indicators.

4. Allows for comparison of alternative implementations of core components (e.g., AMF, SMF, UPF, etc.).
5. Is suitable for integration into CI/CD pipelines and monitoring systems [17, 18].

The proposed method for assessing the quality of 5G core software is based on adapting the ISO/IEC 25010 model to the specifics of the 5G Core architecture. The main idea is to supplement traditional quality characteristics (functionality, performance efficiency, reliability, and others) with domain-oriented metrics inherent to cloud-native and next-generation telecommunication systems.

4.1. Method architecture

The method comprises three hierarchical levels:

1. Basic Level (Standard): Traditional ISO/IEC 25010 characteristics—functional suitability, performance efficiency, reliability, security, maintainability, and portability.
2. Extension Level (5G-Specific): Latency, scalability, fault tolerance, NFV/SDN integration, and energy efficiency.
3. Evaluation Level: Module for collecting, weighting, and interpreting quality indicators. Metrics are normalized to a [0;1] scale and aggregated to form a comprehensive quality score.

$$Q_i = \sum_{j=1}^{n} w_{ij}\, m_{ij}, \quad (1)$$

where Q_i is the aggregated value of the i-th quality characteristic, m_{ij} is the value of the j-th metric for the i-th characteristic (Table 1), and w_{ij} is the weight of the j-th metric (determined either through expert judgment or based on Analytic Hierarchy Process (AHP) analysis).

The overall quality indicator of the 5G core software is calculated as a weighted sum of the quality characteristics:

$$Q_{total} = \sum_{i=1}^{k} W_i\, Q_i, \quad (2)$$

Table 1
Formation of the Metric List.

Characteristic	Metric	Designation	Description
Latency	Average signaling delay	m_1	Delay between AMF ↔ SMF ↔ UPF
Scalability	Maximum number of users	m_2	Number of supported sessions without degradation
Reliability	MTBF (Mean Time Between Failures)	m_3	Average time between failures
Energy Efficiency	Energy consumption per session	m_4	Power consumption at the user level
Performance	Throughput (Gbps)	m_5	Core network traffic processing speed

The method can be implemented as a monitoring module integrated into a CI/CD pipeline or a testing environment (e.g., OpenAirInterface, free5GC). The weighting coefficients can be determined through expert evaluation or the Analytic Hierarchy Process (AHP) method. The method can also be adapted to specific service types such as eMBB, mMTC, or uRLLC.

4.2. Mathematical model

Stage 1: Define a comprehensive set of metrics:

- Functional Metrics (F): correctness, reliability, security;
- Non-Functional Metrics (NF): performance, scalability, energy efficiency;
- Contextual Metrics (C): adaptability, automation, fault tolerance.

$$M=\{m_1, m_2, \dots m_n\},\ m_i \in [0;1]. \tag{3}$$

Stage 2: Normalize metrics.

Each metric is normalized to the interval [0;1] (if not already normalized), for example:

$$\hat{m}_i = \frac{m_i - m_i^{\min}}{m_i^{\max} - m_i^{\min}}, \tag{4}$$

where $\hat{m}_i$ is the normalized value, and m_i^{min}, m_i^{max} are the minimum and maximum bounds of the metric, respectively.

Stage 3: Apply weighting coefficients to metrics.

To account for the significance of different indicators, a system of weighting coefficients is applied:

$$w_i \in [0;1], \sum_{i=1}^{n} w_i = 1. \tag{5}$$

Stage 4: Calculate aggregated indices.

Service level (Q_1):

$$Q_1 = \sum_{i=1}^{k} w_i^{(1)} \cdot \hat{m}_i. \tag{6}$$

Functional level (Q_2):

$$Q_2 = \sum_{j=1}^{l} w_j^{(2)} \cdot \hat{m}_j. \tag{7}$$

Infrastructure level (Q_3):

$$Q_3 = \sum_{z=1}^{p} w_z^{(3)} \cdot \hat{m}_z. \tag{8}$$

where $w^{(r)}$ are the weights at the corresponding level, and $\hat{m}_i$ are the metrics grouped by level.

Stage 5: Calculate the integral quality index.

$$Q_{\text{total}} = \alpha \cdot Q_1 + \beta \cdot Q_2 + \gamma \cdot Q_3, \tag{9}$$

where $\alpha+\beta+\gamma=1$, the coefficients α, β, and γ represent the importance of each level (e.g., α = 0.4, β = 0.35, γ = 0.25).

Stage 6: Perform dynamic assessment over time.

When the version, configuration, or load changes, the metrics are tracked over time t, allowing the calculation of:

$$Q_{\text{total}} = f\left(m_1(t), m_2(t), \ldots m_n(t)\right), \tag{10}$$

and to construct graphs showing the degradation or improvement of quality over time.

All method parameters (the set of metrics, weights, levels) can be adapted to the specifics of a particular use case, such as a telecom operator's core, a private 5G network, or a test environment.

5. Experimental research

To validate the effectiveness of the proposed method, an experimental testbed was implemented based on the open-source free5GC stack, simulating the 5G core network.

Simulation setup:

- 5G Core (free5GC v3.3.0): AMF, SMF, UPF, PCF, AUSF components.
- User Emulator: UERANSIM v3.2.7.
- Virtualization Platform: Oracle VirtualBox Version 7.1.8 r16846.
- Monitoring and Metrics Collection: Prometheus + Grafana + custom scripts.
- Host Environment: Ubuntu Server 22.04 LTS, Intel Xeon, 64 GB RAM.

The entire testbed was deployed inside VirtualBox virtual machines running on the host system (Figures 1–3).

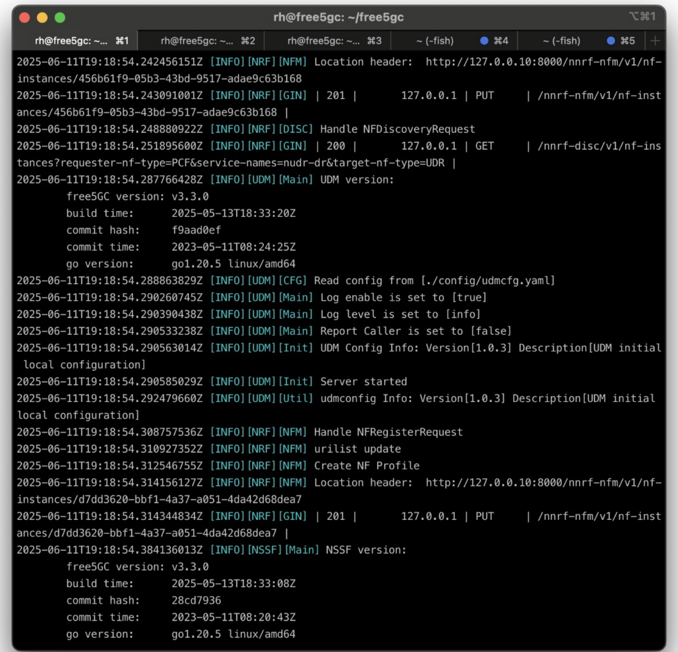

Figure 1: Free5GC Initialization Log

Load scenarios:

- Baseline load (up to 100 concurrent users).
- Peak load (over 1000 eMBB sessions).
- Failure condition (UPF process termination or SMF overload).

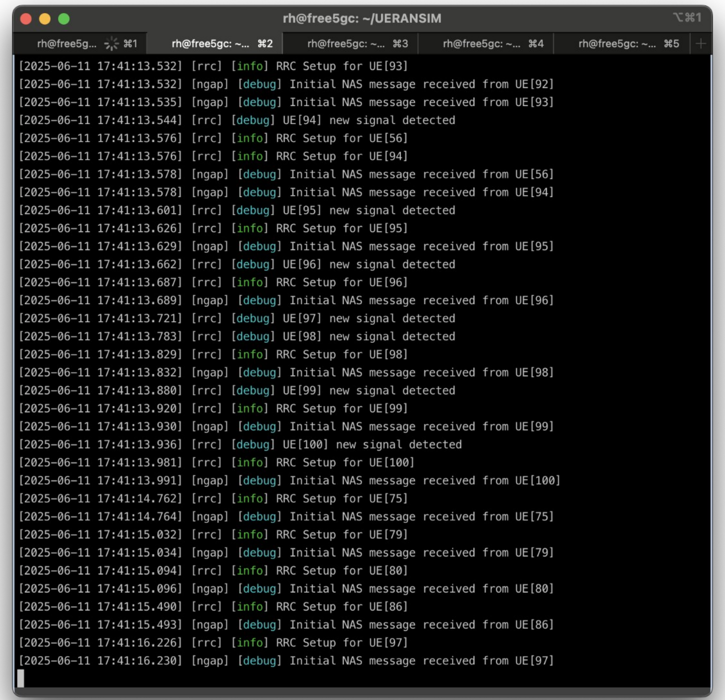

Figure 2: UERANSIM Connection Log

Figure 3: Example Grafana Dashboards

The method was compared with two other approaches:

- Method A—standard quality assessment based on ISO/IEC 25010 (without 5G-specific considerations).
- Method B—the model by [4], focused on microservices (Table 2, Figure 4).

The experimental results emphasize the importance of evaluating accuracy, performance, and applicability. Assessment accuracy refers to the deviation observed during repeated runs, which remained within ±3%, indicating stable metric interpretation.

Method performance is defined by the evaluation time—approximately 3 seconds per scenario with over 500 sessions—enabled by automated data collection. Applicability refers to ease of integration into CI/CD pipelines (e.g., testing new core component versions), suitability for comparing alternatives (such as UPF implementations from different vendors), and potential use as a QoS criterion in SLA agreements.

The proposed method demonstrated higher flexibility and sensitivity to system parameter changes, which is critical in the dynamic environment of the 5G Core. Its main advantage lies in the ability to quantitatively account for 5G-specific factors, which are not captured by general-purpose standards. Due to its modular structure, the method can also be readily adapted to future requirements, such as integration with AI-based core optimization modules.

Table 2
Comparison with Existing Methods.

Method	Latency Consideration	Scalability	Automation	Overall Score (Q_total)
Method A	+	+	+	0.64
Method B	Partially	+	+	0.71
Proposed	+	+	+	0.83

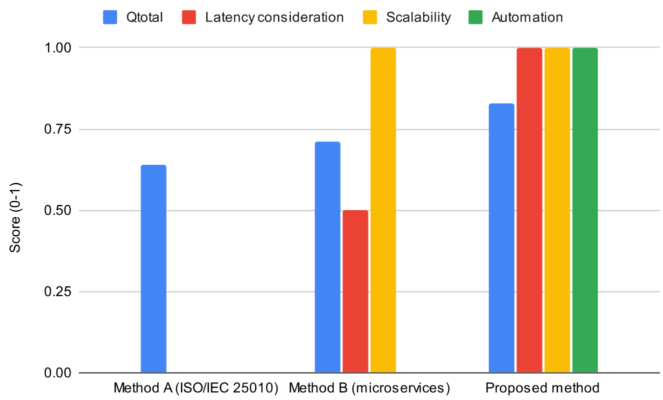

Figure 4: Comparison of 5G Core Software Quality Assessment Methods

The histogram (Figure 4) illustrates a comparison of three 5G core software quality assessment methods across four criteria: Q_{total} is the overall quality score; and consideration of latency, scalability, and automation—key characteristics essential for 5G systems.

The graph (Figure 5) shows the variation of the integral quality indicator Q_{total} depending on the number of concurrent sessions (load). The proposed method demonstrates resilience to increasing load, with only a slight decrease in Q_{total}, while Method A (ISO/IEC 25010) exhibits a significant drop in quality as the number of users increases.

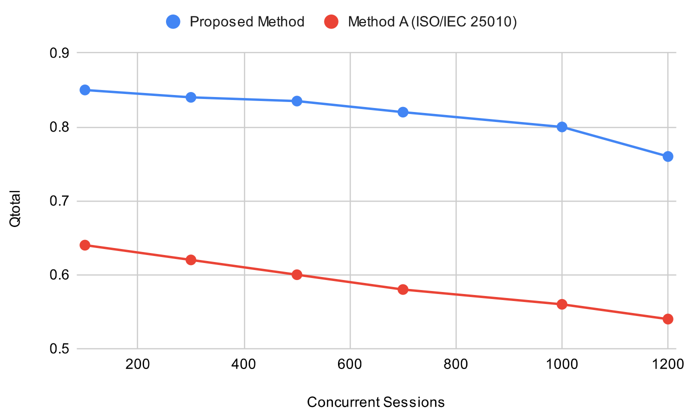

Figure 5: Variation of the Integral Q_{total} Score with Respect to Load

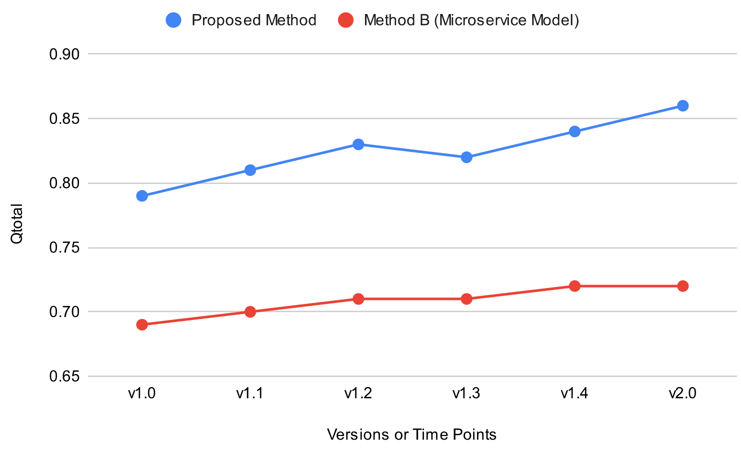

Figure 6: Q_{total} Dynamics During 5G Core Component Version Updates

The graph (Figure 6) illustrates how the Q_{total} indicator changed during updates of 5G core components (e.g., transition from version 1.0 to 2.0). The proposed method clearly captures improvements in system-level quality following updates. Method B demonstrates stability but shows lower sensitivity to changes.

The proposed method for assessing the quality of 5G core software offers several significant advantages that distinguish it from traditional approaches, such as ISO/IEC 25010 or microservice-oriented models. Among these advantages are:

- Comprehensiveness: the method accounts for both classical quality metrics (e.g., functionality, reliability, usability) and 5G-specific aspects such as latency, scalability, and energy efficiency.
- Architectural adaptability: the three-level model structure enables adaptation to various 5G core implementations—ranging from monolithic to cloud-distributed systems.
- Resilience under failure and load: experimental studies showed stable quality scores even under high-load conditions or failures in specific components (UPF, SMF, AMF).
- Automation: the method can be integrated into CI/CD pipelines for continuous quality monitoring.

A key advantage of this method compared to others is its orientation toward metrics relevant to 5G and information and communication systems (ICS) in general. This allows for the effective detection of anomalies during software operation in ICS environments and facilitates quality improvements through error elimination.

However, the method also has certain limitations, including:

- Configuration complexity: full implementation requires deep integration with the operator's telecom infrastructure, including access to telemetry data and logs.
- Dependency on data completeness: the accuracy of the assessment depends on the availability and reliability of input data—particularly from KPI monitoring, tracing, and load testing results.
- Need for customization: different vendor-specific 5G core implementations may require adaptations of metric weights and custom indicators.

The method can be extended to 6G systems or industrial private networks (Private 5G), which demand even higher levels of QoS/QoE. There is potential for integration with ML modules capable of automatically adjusting model weights based on changes in system load, SLA policies, or service type.

Pilot deployment is recommended in a test environment, for example using platforms such as Open5GS or ONF SD-Core. Integration with monitoring systems (e.g., Prometheus, Grafana, Zabbix) should be established to collect input metrics. Gradual integration into the operator's CI/CD pipeline is suggested for analyzing software releases within the DevOps cycle.

Conclusions

This paper proposes a method for assessing the quality of 5G core network software adapted to the specifics of modern telecommunication systems. The method combines ISO/IEC 25010 standards with additional metrics pertinent to 5G networks, including latency, scalability, energy efficiency, automation, and fault tolerance.

Key outcomes include:

- Development of a three-level quality assessment model.
- Formalization of the calculation for an integral quality index.
- Experimental validation demonstrating enhanced accuracy, stability, and relevance compared to existing methods.

The method provides significant value for mobile network operators, software developers, and system integrators by enabling:

- Objective quality assessment of core releases before deployment.
- Identification of scalability bottlenecks.
- Compliance with QoE and SLA requirements.
- Integration into continuous monitoring within DevOps pipelines.

Future work will focus on expanding the model to 6G and NTN networks, integrating machine learning for dynamic metric weighting, and deploying the method in operational environments.

Declaration on Generative AI

While preparing this work, the authors used the AI programs Grammarly Pro to correct text grammar and Strike Plagiarism to search for possible plagiarism. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

[1] ISO/IEC 25010:2011, Systems and Software Engineering—Systems and Software Quality Requirements and Evaluation (SQuaRE)—System and Software Quality Models, International Organization for Standardization, Geneva, 2011.

[2] M. Azgomi, M. Ebrahimi, Quality Model for Real-Time and Embedded Software Systems, J. Syst. Softw. 117 (2016) 436–449. doi:10.1016/j.jss.2016.02.016

[3] S. Newman, Building Microservices: Designing Fine-Grained Systems, 2nd ed., O'Reilly Media, Sebastopol, CA, 2021.

[4] A. Bucchiarone, N. Dragoni, S. Dustdar, S. Larsen, M. Mazzara, A conceptual framework for quality evaluation of microservice architectures, Service Oriented Comput. Appl. 14 (2020) 1–20. doi:10.1007/s11761-019-00273-w

[5] H. Hawilo, A. Shami, M. Mirahmadi, R. Asal, NFV: State of the Art, Challenges, and Implementation in Next Generation Mobile Networks (vEPC), IEEE Network 28 (6) (2014) 18–26. doi:10.1109/MNET.2014.6963800

[6] N. Nikaein, R. Knopp, T. Henderson, OpenAirInterface: A Flexible Platform for 5G Research, ACM SIGCOMM Comput. Commun. Rev. 44 (5) (2014) 33–38. doi:10.1145/2677046.2677052

[7] Y. Kostiuk, et al., Models and Algorithms for Analyzing Information Risks during the Security Audit of Personal Data Information System, in: 3rd Int. Conf. on Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN), Kyiv, Ukraine, vol. 3925, 2025, 155–171.
[8] Y. Kostiuk, et al., A System for Assessing the Interdependencies of Information System Agents in Information Security Risk Management using Cognitive Maps, in: 3rd Int. Conf. on Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN), Kyiv, Ukraine, vol. 3925, 2025, 249–264.
[9] S. Shevchenko, Y. Zhdanova, O. Kryvytska, H. Shevchenko, Fuzzy Cognitive Mapping as a Scenario Approach for Information Security Risk Analysis, in: Cybersecurity Providing in Information and Telecommunication Systems II, vol. 3826, 2024, 356–362.
[10] M. TajDini, V. Sokolov, P. Skladannyi, Performing Sniffing and Spoofing Attack Against ADS-B and Mode S using Software Define Radio, in: IEEE Int. Conf. on Information and Telecommunication Technologies and Radio Electronics (2021) 7–11. doi: 10.1109/UkrMiCo52950.2021.9716665.
[11] M. TajDini, V. Sokolov, V. Buriachok, Men-in-the-Middle Attack Simulation on Low Energy Wireless Devices using Software Define Radio, in: 8th Int. Conf. on "Mathematics. Information Technologies. Education:" Modern Machine Learning Technologies and Data Science, vol. 2386 (2019) 287–296
[12] R. Bybyk, I. Opirskyy, M. McIntosh, SDR Receivers as a New Challenge to Cybersecurity Wireless Technology, in: Cybersecurity Providing in Information and Telecommunication Systems, 3188, 2021, 108–119.
[13] Z. Hu, Y. Khokhlachova, V. Sydorenko, I. Opirskyy, Method for Optimization of Information Security Systems Behavior under Conditions of Influences, Int. J. Intell. Syst. Appl. 9 (12) (2017) 46–58. doi:10.5815/ijisa.2017.12.05
[14] M. Zaliskyi, R. Odarchenko, S. Gnatyuk, Y. Petrova, A. Chaplits, Method of Traffic Monitoring for DDoS Attacks Detection in e-Health Systems and Networks, in: CEUR Workshop Proceedings, 2255, 2018, 193–204.
[15] V. Kharchenko, I. Chyrka, Detection of Airplanes on the Ground using YOLO Neural Network, in: Int. Conf. on Mathematical Methods in Electromagnetic Theory, 2018, 294–297. doi:10.1109/MMET.2018.8460392
[16] J. S. Al-Azzeh, M. Al Hadidi, R. S. Odarchenko, S. Gnatyuk, Z. Shevchuk, Z. Hu, Analysis of Self-Similar Traffic Models in Computer Networks, Int. Rev. Modelling Simulations 10 (5) (2017) 328–336. doi:10.15866/iremos.v10i5.12009
[17] Y. Averyanova, et al., UAS Cyber Security Hazards Analysis and Approach to Qualitative Assessment, in: Data Sci. Secur., 290, 2021, 258–265. doi:10.1007/978-981-16-4486-3_28
[18] V. Tkachuk, Y. Yechkalo, S. Semerikov, M. Kislova, Y. Hladyr, Using Mobile ICT for Online Learning during COVID-19 lockdown, Communications in Computer and Information Science, 1308 (2021) 46–67. doi:10.1007/978-3-030-77592-6_3