

DEVELOPMENT OF A MULTI-LOOP SECURITY SYSTEM OF INFORMATION INTERACTIONS IN SOCIO-CYBERPHYSICAL SYSTEMS

Serhii Yevseiev

Corresponding author

Doctor of Technical Science, Professor, Head of Department

Department of Cyber Security*

E-mail: Serhii.Yevseiev@gmail.com

Oleksandr Milov

Doctor of Technical Sciences, Professor

Department of Cyber Security*

Nataliia Dzhenuk

Associate Professor

Department of Information Systems named after V. O. Kravets*

Maksym Tolkachov

Associate Professor*

Tetiana Voitko

Researcher

Research Department

Institute of Information and Communication Technologies and Cyber Defense

National Defence University of Ukraine

Povitroflotskiy ave., 28, Kyiv, Ukraine, 03049

Mykhailo Prygara

PhD, Associate Professor

Department of Machine Industry Technology***

Natalia Voropay

PhD, Associate Professor

Department of Cyber Security*

Oleksandr Shpak

PhD, Associate Professor

Department of Software System***

Andrii Volkov

Head of Department**

Oleksandr Lezik

PhD, Associate Professor**

*National Technical University "Kharkiv Polytechnic Institute"

Kyrpychva str., 2, Kharkiv, Ukraine, 61002

**Department of Air Defense Forces Tactics of the Ground Forces

Ivan Kozhedub Kharkiv National Air Force University

Dynamivska str., 3a, Kharkiv, Ukraine, 61021

***Uzhhorod National University

Narodna sq., 3, Uzhhorod, Ukraine, 88000

The object of the study is a multi-loop security system of information interactions in socio-cyberphysical systems. The dynamic nature of physical environments inherently challenges the ability of socio-cyber-physical systems to perform adequate control actions for physical assets in many contexts. However, adaptation and evolution actions must be evaluated before implementation in the control system to ensure fault tolerance while minimizing risks. Therefore, the design of socio-cyber-physical systems must ensure not only reliable autonomy, but also operational fault tolerance and safety. The proposed approach is based on the integration of targeted (mixed) threats based on the synthesis of technical cyber threats with social engineering methods. This approach allows forming a dynamic security model based on the analysis of the interaction of various agents in socio-cyberphysical systems, which makes it possible to increase the level of counteraction to targeted (mixed) cyber threats.

The results of modeling are based on the proposed classification of threats using social engineering methods, which allows cyber-attackers to ensure the probability of implementing targeted threats up to 95–98 %. The proposed classification of threats based on social engineering methods will allow forming an additional parameter for the objectivity of target threats, taking into account their integration and synergy. At the same time, the presented model will make it possible to timely provide knowledge about the possibility of implementing a targeted attack and timely take preventive countermeasures. This approach will improve the set of protection measures, as well as promptly create an increase in the level of resistance of the company's personnel (organization, enterprise, etc.) to threats of social engineering

Keywords: socio-cyber-physical system, security model of information interactions, social engineering, targeted attacks

Received date 11.08.2023

Accepted date 20.10.2023

Published date 30.10.2023

How to Cite: Yevseiev, S., Milov, O., Dzhenuk, N., Tolkachov, M., Voitko, T., Prygara, M., Voropay, N., Shpak, O., Volkov, A.,

Lezik, O. (2023). Development of a multi-loop security system of information interactions in socio-cyberphysical systems. *Eastern-European Journal of Enterprise Technologies*, 5 (9 (125)), 53–74. doi: <https://doi.org/10.15587/1729-4061.2023.289467>

1. Introduction

Development of Cyber-Physical Systems (CPS) is a challenging task in critical applications such as avionics, automotive, etc. This task is getting increasingly challenging as CPS become more complex: indeed, CPS

currently consist of a large number of components and subsystems of heterogeneous nature and different levels of criticality. Additionally, non-functional aspects or perspectives such as time, memory, power consumption, reliability, temperature, and security are as important as functionality.

There are many analysis methods and tools for validating CPS, but their underlying model is always specific to a single viewpoint, and there is currently limited support for semantically linking viewpoints. In practice, the assumptions that analysis of a particular viewpoint makes about other viewpoints remain largely implicit, and when they are explicit, they are handled largely manually. The current design process overly restricts the set of possible system designs and there is a need for methods and tools to formally link viewpoint-specific models and corresponding analysis results [1, 2].

In particular, the works that evolve from CPS to incorporate human aspects can be divided into two main computing paradigms, namely the Cyber-Physical-Human System (CPHS) and the Cyber-Physical-Social System (CPSS). According to the state of the art, both computing paradigms define the interaction space in which people and CPS objects live together [1–3]. It can also be seen that CPHS and CPSS are used interchangeably by researchers focused on the presence of humans and their interactions with machines in the overall socio-technical system. Despite the roughly equivalent use of the two acronyms, “social” has a broader meaning. The term “social” reflects emotional and cognitive characteristics. It also conveys the socio-technical principles that govern human behavior in a context that must eventually be transferred to machines.

Based on this conceptual distinction between human and social, CPHS and CPSS are distinguished by the way they take into account social factors, especially those related to machines. Therefore, aiming at better human-machine synergy, systems should be designed with a social aspect in mind, like CPSS. Especially when developing industrial systems that involve close collaboration between humans and CPS; taking into account social aspects provides a better interaction experience and increases employee efficiency [4]. Thus, human-machine interaction can rise to a more cognitive level in which machines can adapt their behavior by identifying situations, understanding and reasoning about human needs in context [5]. However, in Industry 4.0, human-centered computing paradigms do not yet have a strong system foundation. Consequently, they often fail to fully connect social aspects in CPS [6].

Similar to traditional computing, a new computing paradigm called cyber-physical-social computing or physical-cyber-social computing [7] has emerged and has attracted worldwide attention in recent years, which focuses on research into the digital fusion between people, computers and things. In fact, this paradigm originates in the development of cyber-physical systems (CPS) and cyber-social systems (CSS) technologies. Additionally, social characteristics and interaction are being introduced into CPS due to the growing number of human-centered computing. The corresponding computing systems are called cyber-physical-social systems (CPSS) [8].

The key technologies for CPSS development are closely related to transdisciplinary technologies spanning CPS and CSS. Unfortunately, designing CPSS is challenging. First, CPSS are quite complex systems due to network heterogeneity, software and hardware complexity, and lack effective design approaches to systematically address these issues in a unified manner. Secondly, the social part in CPSS has not yet been given due attention; there are no feasible approaches to unified modeling of cyberspace, physical and

social space. Finally, the security issue has not been sufficiently researched because CPSS are still in their infancy.

Thus, an effective, safety-critical design of the elements of the security system for the interaction of continuous business processes in CPSS is an urgent task. The solution requires the development of a security model based on an analysis of the interaction of continuous business processes in CPSS, taking into account the multi-loop arrangement of infrastructure elements and the variety of technologies used to build CPSS.

2. Literature review and problem statement

Cyber-Physical Systems (CPS) are complex distributed systems driven or controlled by computer algorithms and performing computational procedures in their distributed closed-loop environment. The term “cybernetic systems” appeared around 2006 when it was proposed by Helen Gill at the National Science Foundation in the United States [9].

Access networks with sensors and actuators are controlled and managed through the computing nodes of the cyber-physical system. They are typically developed using model-based approaches for large structures. They are pre-programmed for specific situations based on a set of rules and regulated by a traditional feedback control loop [10]. CPS can be implemented at various scales, ranging from the nanoworld to large-scale systems. Their complex interaction with the environment, interaction with social systems and the possibility of external malicious control can lead to unpredictable consequences [11]. Most CPS applications have not considered human factors as an internal element. To this end, socio-cyber-physical systems (SCPS) [12] have emerged, which are usually considered as an extension of CPS and integrate cyberspace, physical space and social space.

With the development of the Internet of Things and social networks, a large amount of data is being generated. Social engineering techniques are combined with cyber, hybrid and targeted attacks. In this regard, a hybridity of targeted attacks appears. How to protect an enterprise network from the influence of heterogeneous data with a sociological component is a very difficult question.

A comprehensive, multi-level approach is required, which, in addition, solves the problem of protection against the influence of targeted attacks associated with different social groups in society. A compositional structure of protection is required. This protection must integrate the functional elements of systems to implement system-level properties that cannot be achieved by integrating the local properties of system components [10].

Further development of computing resources and capabilities of artificial intelligence (AI) has made a decisive shift in the worldview and social culture of electronic communication.

OpenAI, GoogleAI and DeepMind companies are improving modern AI technologies. Modern developments of OpenAI, Bard, DALL-E, ImageNet, etc. based on the Generative Pre-trained Transformer model actively use a set of databases [13]. They transform and generate text, images, and speech. But the most important thing to take into account in the development trend of system security is that these developments have an open commercial and non-commercial AI API used in social systems.

According to [14], as AI-based systems continue to improve, there are concerns that attackers will have more reasons to use them for malicious purposes. Artificial intelligence systems carry a number of risks that are not fully taken into account by existing structures and approaches to risk management in the content analysis of socio-cyber-physical systems. On the other hand, with proper control, artificial intelligence systems can mitigate security threats, their consequences and manage them [15].

Easy availability of information from open sources and uncontrolled intelligence make it easier to gather information. Specific targets can be carefully selected to create more robust and targeted attacks. A large group of victims can be targeted simultaneously, and some open-source tools can be used to launch semi-automated attacks. Technologies such as machine learning and artificial intelligence make attacks on sociophysical systems more effective and aggressive. Targeted, large-scale, robotic, automated attacks become possible. SCPS systems are becoming a serious, universal and persistent security threat.

Data transmitted from SCPS systems and information resulting from processing often generate different forms of data that require different levels of security. Managing data and information requires effective methods for processing, interpreting and reusing them.

For cybersecurity, new components are being added related to social influence, cognition, emotion, and decision-making. These components are present in methods of attacks on socio-cyberphysical systems [16].

NIST has released a white paper outlining standards for identifying and managing bias (prejudice) in AI [17]. This guide is not intended to offer a definitive solution for eliminating AI bias. Its goal is to “identify significant issues in the complex field of AI bias and provide the first step in a roadmap for developing detailed socio-technical guidance for identifying and addressing AI bias”.

Advantages:

- provides a clear and standardized framework for identifying and managing bias in AI systems;
- offers a comprehensive approach to identifying bias, including different stages of the AI development lifecycle;
- offers practical guidance for implementing the framework in a variety of organizational contexts;
- offers a number of tools and techniques that can be used to identify and eliminate bias in AI systems;
- can help organizations build more reliable and transparent artificial intelligence systems.

Disadvantages:

- the publication may not be accessible to non-experts in AI development or policy;
- the framework may be too prescriptive for some organizations, limiting their flexibility to adapt to their specific needs;
- the publication does not provide guidance on how to measure the effectiveness of bias reduction techniques and mitigate malicious influence;
- this framework focuses on identifying and managing bias rather than addressing the root causes of bias in society and data.

A comprehensive guide to generative models and their applications is provided by [18]. The book covers a range of topics, including autoencoders, variational autoencoders, generative adversarial networks, and deep belief networks. The author provides clear explanations of complex concepts

and offers practical examples and code snippets to help readers implement the techniques discussed. The book also includes detailed descriptions of popular open-source libraries such as TensorFlow and PyTorch, which are commonly used to build deep learning models.

One of the strengths of the publication is its focus on practical application. The author gives many examples of how generative models can be used in real-world scenarios such as creating images, text, and even music. The publication includes detailed explanations of the mathematics behind generative models that can be useful for implementing cyber defense systems.

While the publication is a good introduction to deep learning and generative models, it does not cover all aspects. Some important topics regarding system-level reasoning and its means of implementation in terms of processed knowledge, awareness raising, reasoning mechanisms have been omitted. The publication primarily focuses on specific deep learning frameworks such as Keras and TensorFlow, which may limit the applicability of the concepts presented to practitioners using other tools or languages.

The paper [19] provides an understanding of data management of the cyber-physical-social system (D-CPSS) using the 7C Framework. It describes how cyber-physical systems (CPS) integrate with social systems to form data-driven cyber-physical and social systems (D-CPSS). Integrating these systems brings many benefits, including increased efficiency, productivity and flexibility. However, challenges still exist in understanding and implementing D-CPSS. To better understand D-CPSS, the 7C framework is proposed. This framework provides a holistic view of D-CPSS, taking into account various components and their relationships. The 7C framework consists of seven dimensions:

- cyber: this dimension refers to the digital technologies used in D-CPSS, including sensors, data analysis, and machine learning algorithms. These technologies enable the collection, analysis and interpretation of data that can be used to optimize production processes;
- physical: this parameter refers to the physical components of D-CPSS, including machines, robots, and other equipment. These components are integrated with cyber systems to provide real-time monitoring and control of production processes;
- social: this aspect refers to the human aspects of D-CPSS, including social networks and relationships between stakeholders. These relationships are critical to the success of social production as they enable collaboration, knowledge sharing and innovation;
- cognitive: this dimension refers to the cognitive capabilities of D-CPSS, including decision-making and problem-solving. These capabilities are enhanced by the integration of cyber, physical and social systems;
- communication: this parameter refers to the communication networks and protocols used in D-CPSS. Effective communication is essential to coordinate production activities and ensure that all stakeholders are informed and involved;
- control: this parameter refers to the control mechanisms used in D-CPSS, including feedback loops and autonomous decision-making. These mechanisms allow the system to respond to changes in the environment and optimize performance;
- context: this dimension refers to the broader context in which D-CPSS operates, including the regulatory envi-

ronment, market demand, and societal expectations. Understanding the context is essential to design and implement a D-CPSS that meets the needs of all stakeholders.

The 7C framework provides a useful tool for analyzing and developing D-CPSS considering cyber, physical, social, cognitive, communication, control, and context dimensions.

Although the paper provides a good overview of the 7C framework for understanding data-driven cyber-physical and social systems (D-CPSS) in the context of social production, there are some shortcomings:

- lack of in-depth analysis: the paper provides a general overview of the 7C concept and its potential applications in social production. However, it does not give a detailed analysis of the structure or its limitations;

- lack of empirical data: the paper does not provide any empirical data or case studies to support its arguments. It would be useful to have some examples of successful D-CPSS implementations and their impact on social production;

- lack of discussion of ethical and social implications: although the paper briefly mentions the social aspects of D-CPSS, it does not discuss the ethical and social implications of these systems. For example, the paper does not address the potential impact of D-CPSS on job displacement or privacy issues related to data collection and analysis;

- lack of discussion of technical challenges: although the paper briefly mentions the cyber and physical aspects of D-CPSS, it does not discuss technical challenges associated with implementing these systems. For example, the paper does not address potential cybersecurity risks or problems of integrating disparate systems.

A comprehensive overview of social engineering in cybersecurity is given in [20]. The paper begins by defining social engineering and highlighting the key differences between social engineering and traditional hacking methods. Various types of social engineering attacks are discussed. The authors provide detailed descriptions of each type of attack, as well as examples of real incidents. One of the strengths of the paper is a detailed exploration of the mechanisms that make social engineering attacks successful. The authors identify several factors that contribute to the success of social engineering attacks, including the human tendency to trust others, the desire for social acceptance, and the tendency to rely on heuristics or mental shortcuts when making decisions. Various types of human vulnerabilities that cybercriminals exploit in social engineering attacks are also discussed. These include cognitive biases, emotional manipulation, and social influence. The authors give examples of how cybercriminals use these vulnerabilities to manipulate people into divulging sensitive information or performing actions they would not otherwise do. The authors also discuss some countermeasures to reduce the risks associated with social engineering attacks. In general, the content of the paper is comprehensive and informative. It provides a detailed overview of social engineering in cybersecurity.

Although the paper gives a comprehensive overview of social engineering in cybersecurity, there are some disadvantages and weaknesses to be considered. One of the drawbacks of the paper is that it relies heavily on theoretical concepts and does not contain enough practical examples or case studies. Although the authors cite some real-life examples of social engineering attacks, their number is limited and they are not analyzed in detail. This can make it difficult to fully understand the impact of social engineering attacks and how they work in practice. Additionally, the paper does

not provide a comprehensive discussion of countermeasures to prevent or mitigate social engineering attacks. Although the authors briefly mention some countermeasures, they do not go into detail about how they can be implemented or how effective they are in practice.

The paper [21] highlights the importance of detecting anomalies and attacks in supervisory control networks. The work focuses on ensuring the safety and reliability of these systems by detecting abnormal behavior or malicious attacks that could potentially compromise their functionality. The paper includes statistical analysis, machine learning algorithms, rule-based systems, and a combination of these approaches. The advantages and limitations of each method are described, but possible solutions or improvements are not suggested. The research contributes to improving safety measures for CPS.

An approach to building a cyber security system that allows for a comprehensive analysis of various attack vectors is described in [22]. The main idea is to develop an intelligent gateway system that effectively solves security problems through the use of virtual enterprise. The paper suggests that this approach can be useful in identifying vulnerabilities and weaknesses in a system. Simulation of various attack scenarios allows you to analyze potential security threats and develop appropriate countermeasures. The advantages of using virtualization technology in the development process are emphasized. It provides flexibility and adaptability during testing.

The paper [23] presents a socio-technical modeling approach as a valuable tool for understanding and mitigating cyber-physical threats. By incorporating social and organizational factors in the analysis, the proposed approach offers a more complete understanding of vulnerabilities and their potential impact.

The work [24] focuses on highlighting the concept of the cyber-physical universe, which encompasses a vast network of interconnected devices, systems and infrastructures connecting the digital and physical worlds. The authors argue that by viewing this nexus of cyber and physical entities as a cohesive system, one can gain valuable insight into emerging patterns and understand the dynamics shaping the modern world. The paper suggests that this cyber-physical universe represents a paradigm shift in the understanding of complex systems. By integrating data from sensors, devices and digital networks in real time, we can analyze and model the behavior of this interconnected system, thereby providing a deeper understanding of its emergent properties. An integrated approach to interacting data analysis, machine learning, and artificial intelligence is described to understand the vast amounts of data generated by the cyber-physical universe. It is claimed that using these technologies, we can identify hidden patterns, predict future trends and optimize system performance in various fields such as smart cities, transport, healthcare and industrial automation. The paper acknowledges the challenges associated with the cyber-physical universe, including privacy, security concerns and ethical implications. This requires a comprehensive approach that includes not only technical knowledge, but also considerations of social impact, policy frameworks and legal regulations. The study of patterns and dynamics in the context of the cyber-physical universe is discussed.

In [25], a socio-technical approach to creating sustainable connected transport systems is considered. The socio-technical approach combines technical elements such as advanced sensor technologies, data analysis and intelligent transport systems. By integrating technical elements with

social factors, the proposed approach aims to improve the adaptability and responsiveness of transport infrastructure to disruptions.

Statistics on the interaction between technological infrastructure (such as sensors, communication networks and data analysis) and social factors (such as user behavior, policy frameworks, governance models) is provided. Options for designing and implementing this approach as a multi-aspect approach for managing transport systems are described. This includes understanding commuter behavior and preferences, considering the impact on urban mobility patterns, addressing privacy and security concerns, and ensuring equitable access to transport services. Resilience is a key aspect of this approach, emphasizing the ability of transport systems to withstand and recover from failures. The application for narrow tasks is described, which can be expanded to broader aspects.

The concept of digital twins and their potential impact on society are explored in [26]. Digital twins enable real-time monitoring, analysis and optimization, bridging the gap between the physical and digital worlds. The authors emphasize that this approach is universal and can be applied in industries such as manufacturing, healthcare, transport and urban planning. One of the key aspects discussed in the paper is the integration of digital twins into SCPS. The authors suggest that integrating digital twins into SCPS can lead to more efficient decision-making, as well as improved resource allocation and use. The need for measures to ensure confidentiality and data security is emphasized. The concept of twins is presented without examples of practical implementation.

The paper [27] describes the transition from traditional approaches to a more comprehensive approach that includes the development of scenarios for reliable AI. It describes the design and development of specific scenarios and use cases to evaluate and verify the performance and reliability of AI systems. Methods of identification, interpretation, monitoring, correction and validation to create reliable, fair, and ethical AI systems are considered.

The analysis of publications allows us to conclude that when creating socio-cyber-physical systems and systems for ensuring their safe functioning, issues related to information processes of the social component have not been sufficiently studied. This is due to the fact that the processing and perception of information by a person differs significantly from those in a cyber-physical system. Therefore, these issues of perception analysis, assessment, exchange and processing of information are relevant and explain the urgency of developing methods for assessing these processes, which can be considered as a problem facing the authors.

3. The aim and objectives of the study

The aim of the study is to create a multi-loop security system in socio-cyberphysical

systems, taking into account the integration of cyber threats with threats based on social engineering, as well as a security model for information interactions. This approach makes it possible to ensure the synergy of security systems between the functional elements of security systems and ensure the implementation of system-level properties that cannot be achieved by integrating the local properties of the security system components.

To achieve the aim, the following objectives must be accomplished:

- to analyze the processes of influence in socio-cyber-physical systems;
- to create a threat classifier taking into account threats based on social engineering methods and multi-loop information security systems;
- to develop a security model of information interactions in socio-cyberphysical systems;
- to conduct a study of the security model of information interactions in socio-cyberphysical systems.

4. Materials and methods

4.1. Processes of information interaction in the socio-cyberphysical system

The object of the study is the process of forming security mechanisms for information interactions in socio-cyber-physical systems, the subject is a security model of information interactions in socio-cyberphysical systems. The study proposes an approach that is based on passive and active stages, including:

- creation of new data by artificial intelligence for automatic alerts;
- information collection templates;
- search for threat indicators;
- procedural analysis;
- automation of artificial intelligence procedures for data analysis and creation of new ones.

The general problems of timely detection of cyber threats (targeted attacks) on socio-cyber-physical systems are presented in Fig. 1.

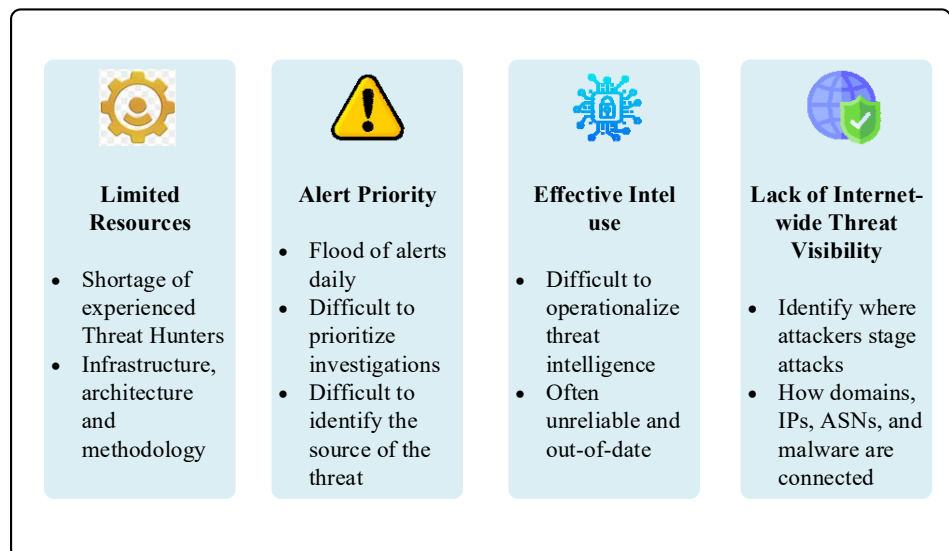


Fig. 1. General threat detection issues

From a general point of view, a cyber-physical-social system (CPSS) is a combination of two systems: cyber-physical (CPS) and social. CPS belongs to a generation of systems with integrated computing and physical capabilities, closely related to the 4th industrial revolution [28]. The social aspect refers to interacting individuals having their own consciousness, preferences, motivation and behavior. The development of CPSS is still at an early stage. Over the past decade, different researchers have used different terms to refer to the integration of the human dimension in CPS, proposing different concepts. For example, [29–31] used cyber-physical-human systems (CPHS), defined as “*systems of interconnected systems (computers, cyber-physical devices and people) “talking” to each other in space” and time, while also allowing other systems, devices and data streams to connect and disconnect*”. In [32], the concept of cyber-physical-social thinking (CPST) hyperspace was introduced for the geological information system. This work defines CPSS as “*a system deployed with a focus on people, knowledge, society and culture, in addition to cyberspace and physical space. Therefore, it can bind nature, cyberspace and society with certain rules*”, while for CPST this is established through the merging of a new dimension of thought space into the CPS space. A thought space is a high-level thinking or idea that arose during the people’s intellectual activity. The work visualizes human intelligence separately from the social aspect of CPSS as a thought space. On the other hand, the term “cyber-physical social systems” (CPSS) was also used in [33, 34] and was defined as “*complex socio-technical systems in which human and technical aspects are closely intertwined*”. According to this definition, the scope of SCPS extends to the intangible objects of social context, which include social culture and norms, personal beliefs and attitudes, and informal institutions of social interactions. The concept of CPSS has been formulated in many works; however, the usage is not uniform. Moreover, the perspective and method of definition

also vary from domain to domain. In an attempt to address this gap, [3] proposed a holistic definition and domain-independent conceptualization of CPSS based on the general framework provided by systems theory. Let us introduce the following concepts [4]:

– CPSS is a system strictly consisting of a Cyber-Physical System (CPS) and a Social System (SS) in which system components interact in a virtual and physical environment, where CPS and SS are defined respectively as follows;

– a Cyber-Physical System (CPS) is a system that encompasses all systems and subsystems of Cyber and Physical Systems, their components and interactions between them, as well as the integration of computing with physical processes;

– a Social System (SS) is a system consisting of interacting individuals, each having his own consciousness, preferences, motivation and behavior.

Elements united on the basis of information transfer processes form cybernetic space. Social space is represented by people with their knowledge, mental abilities and sociocultural elements. Cybernetic space exchanges information with physical space (endpoints) and social space (people).

All three spaces are closely interconnected and are represented by sets of their components (physical objects with software elements and people). Information interaction between physical, cybernetic and social spaces (S-CPS) is carried out through the interaction of the components that form these spaces. Because physical resources cannot interact without the support of information technology, cyber protection of information transfer processes is required. Since interaction also affects social resources, cyber defense must use artificial intelligence methods as a mandatory component. At the same time, cybernetic and social resources are active resources that can initiate interaction, as shown in Fig. 2 by a double-headed arrow.

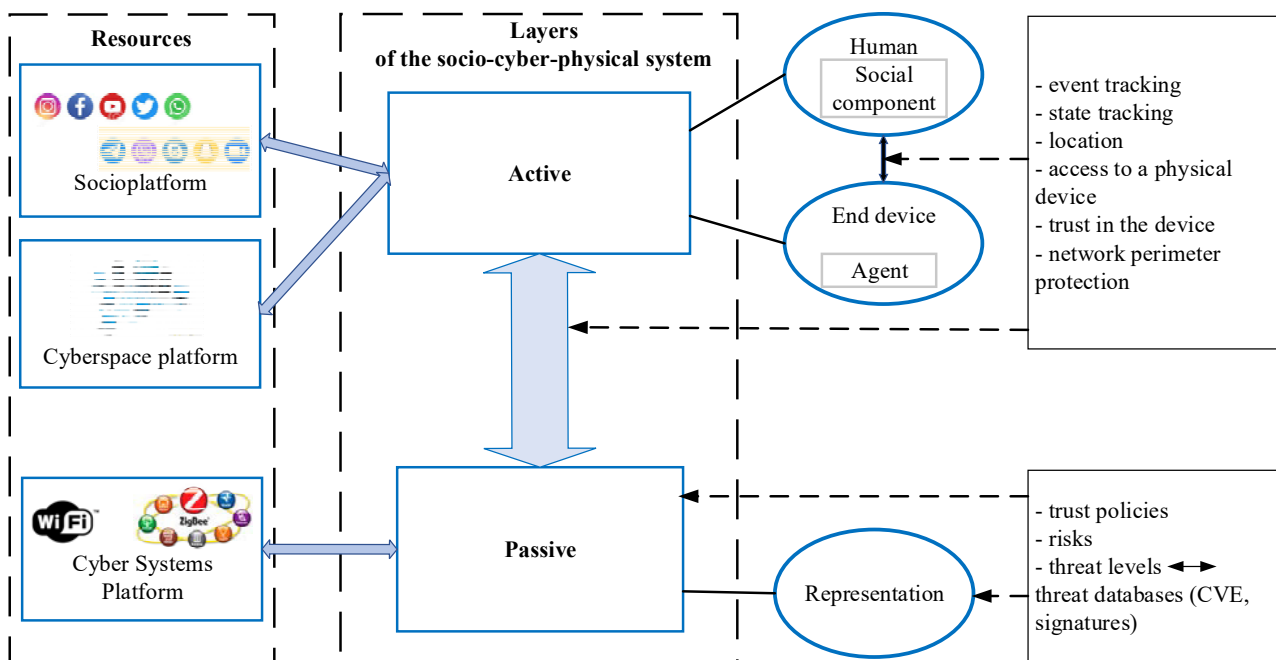


Fig. 2. Cybernetic and social resources that initiate interaction

In addition to cyber and social resources, S-CPS includes a set of representations that serve to represent knowledge about the problem area in which S-CPS operates and provides a figurative component for information interaction. As one of the possible knowledge models, a model can be used that is formed on the basis of a dictionary of signatures, trust policies, risks, threat level (CVE behavioral databases, standard policies). The listed elements participate in interaction as a passive element – through requests from active elements (cybernetic and social resources). In Fig. 2, interaction with the passive element is indicated by unidirectional arrows.

Possible processes of information interaction among S-CPS elements relate to the cyber defense of the current area of interaction of the network or networks. Cyber defense is based on the Trust Security principle [35] and includes:

- event tracking;
- state tracking;
- location of physical devices;
- access to physical devices;
- trust level;
- security level of the network segment.

The action of security system mechanisms (cyber defense) extends not only to terminal devices, but, most importantly, to the information channel that S-CPS resources use to interact with each other [36]. An online community can act as an information channel, i.e. a virtual community whose participants interact via the Internet. Unlike social networks, an online community unites people based on common interests or goals. Since S-CPS involves the integration of physical, cyber and social spaces to solve specific problems, the commonality of goals was a key factor that determined the choice of the Internet community concept for organizing information interaction among S-CPS resources.

To understand the processes occurring in online communities, in particular, to understand and analyze the information influence on members of such a community from interested parties, agent-based models can be useful. The agent-based approach provides simulation modeling of the behavior of agents playing various roles in communication processes, agent characteristics, decision-making, adaptive behavior and mobility, as well as agent interaction with the environment.

4.2. Multidimensionality as a characteristic of the structure of a security system

The widespread idea that we are at the beginning of the “fourth industrial revolution” has attracted significant attention from both business and academia. This movement, often referred to as Industry 4.0 or Smart Manufacturing, has become more prominent following a program launched by the German government and the development of similar initiatives in the United States and other countries [37]. This was also facilitated by programs such as the “Factory of the Future” of the European Union [38]. The original idea primarily focused on the convergence of the physical and virtual worlds, represented by the term “cyber-physical system (CPS)”, thereby promoting a “CPS-based industry”. As a result, the term CPPS (cyber-physical production system) even appeared [39, 40]. Soon, this idea gradually grew into a combination of CPS, Internet of Things (IoT), and Internet of Services (IoS), demonstrating the evolution towards digitalization or digital transformation. The original point of view has been complemented by the aspect of “smartness” or “intelligence”, as evidenced by the terms “smart machines”,

“smart sensors”, “smart factory”, “smart environment”, “smart products”, etc. [41]. Thus, the next industrial revolution is the result of a close combination of contributions from various fields of technology, computer science, manufacturing and, in particular, artificial intelligence.

To adequately understand the holistic vision brought by Industry 4.0 and the associated digital transformation, it is necessary to view it through the lens of collaborative networks (CNs). Of course, we can say that this is “another partial view”. However, the prospect of collaboration is explicitly or implicitly present in most Industry 4.0 requirements. In addition, CNs are interdisciplinary and multidisciplinary in nature, which can be useful in gaining a holistic understanding of challenges associated with this transformation. Therefore, “collaboration” is the decisive challenge of the fourth industrial revolution. As a consequence, the area of “Collaboration Networks”, among others, must be seen as a major factor in this transformation. To support this assertion, some keywords associated with Industry 4.0 should be noted, including “networks”, “vertical and horizontal integration”, “value chains”, and “co-design/end-to-end design” [42]. The review [43] also indicates that “interconnection” and “collaboration” are among the main “terminology clusters” found.

Thus, it can be argued that the classical methods of developing network perimeter protection are no longer valid. The amount of data, the amount of self-describing data that systems generate increase by an order of magnitude every 5 years. At the same time, the problem is that, in parallel with data growth, the subset of redundant data (i.e. noise) grows even faster than the set of useful data as a whole. To assess the dynamics of these processes, it is necessary to analyze the interaction in complex dynamic environments. Classical security methods are difficult to consider as classical paradigms of designing security systems. Solutions are needed that enable the design of security systems that remain operational in complex, constantly changing environments. Such solutions exist. They are cognitive processes that the brain actually uses to work in a complex, dynamic environment. It should be noted that in order to cope with the complexity and dynamism of the environment, the brain uses methods such as reverse engineering, cognitive thinking and modeling, which belong to the field of artificial intelligence (AI).

This is about taking existing security algorithms and applying them to the full range of IT operations, taking into account artificial intelligence methods and models. Considering the speed at which structures and situations change in IT environments, it is simply impossible to implement effective protection of these processes without using AI, without using similar approaches.

An analysis of artificial intelligence methods shows that the algorithms it contains are divided into different types (Fig. 3) [21]. Particular attention should be paid to algorithms associated with the selection of data subject to subsequent analysis. These are algorithms related to the detection of patterns. There are algorithms associated with the results of processing information contained in these patterns. Much of the work in deep learning is largely related to pattern detection. Thus, algorithms are used that determine the choice of which data to analyze first. With regard to means of ensuring the security of computer and communication systems, it can be argued that the emphasis in protection is shifting towards detecting patterns of network traffic.

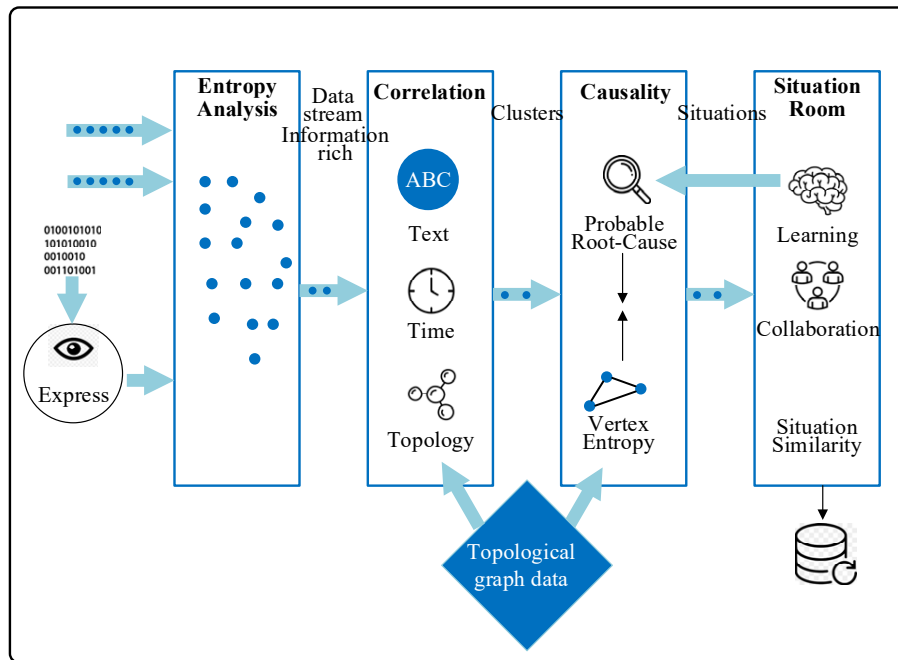


Fig. 3. Types of artificial intelligence algorithms

Thus, the analysis of artificial intelligence algorithms shown in Fig. 3 allows us to formulate a vector for further improvement of the security system for the interaction of SCPS continuous business processes. This approach ensures the emergence of the formation of multi-loop security systems, taking into account the integration of both CPHS infrastructure elements and CPSS formation technologies.

To understand the processes occurring in online communities, in particular, to understand and analyze the information influence on members of such a community from stakeholders, agent-based models can be useful. The agent-based approach provides simulation modeling of the behavior of agents playing various roles in communication processes, agent characteristics, decision-making, adaptive behavior and mobility, as well as agent interaction with the environment.

Not only analytical methods, but also software tools have been developed to use this modeling approach using the NetLogo agent-based modeling framework. This environment is one of the most popular for such purposes today, and the built-in graphical tools and constant support and updating of the versions used allow you to make a clear choice in favor of this programming and modeling environment. Ultimately, the use of the selected tools ensures the creation of more realistic and management-relevant forecasts, and also opens up new opportunities for the exchange of models and connection with other methods.

The use of any modeling environment must involve a preliminary analysis of the assumptions and limitations of the model being developed. First of all, this relates to the time of process modeling. Although it is possible to build continuous-time models of information influence, discrete-time models were chosen as more common. Discrete-time models of influence dynamics reflect changes in the set of opinions of agents at a given time t to the set of opinions of these agents at time $t+1$. In the proposed model of the dynamics of information influence, opinions of individual agents are formed on the basis of a discrete set of values.

It should be kept in mind that agent-based modeling, despite all its strengths, can also have weaknesses. One of them is the amount of calculations required. Interactions and

associated computations tend to increase exponentially as the number of agents increases. Another potential weakness is the tendency to ignore elements of system-level behavior when the focus is on agent-level behavior.

5. Results of developing a security model for information interactions in socio-cyberphysical systems

5.1. Analysis of influence processes in socio-cyber-physical systems

To understand what patterns of influence processes should be used for certain agents to analyze and identify a threat, you need to understand how it is formed. This provision can be considered as the basic principle for constructing a multidimensional structure model in SCPS.

Against the backdrop of repeated online influence operations (hidden or overt), attempts to mislead or influence the opinion of the target audience [44], individual behavioral data is collected and diagnosed in the cyber environment [45]. Based on this, linguistic (language) models are formed, using which content can be generated in real time, which can be added to the content of the current interaction at the place where this inter-individual interaction occurs. As a result, a situation arises where distorted social norms are generated in distorted content [46]. The result of this process is the formation of the necessary behavior of the target audience.

It is advisable to describe the processes of influencing the target audience using agent-based modeling methods. This approach is natural for representing the intentional manipulation of information to influence beliefs or opinions.

Suppose we have a set of N agents, each representing an individual in the target audience. Each agent is characterized by a binary belief variable Y_i , where $Y_i=1$ indicates agreement with the expressed point of view or direction of behavior, and $Y_i=0$ indicates disagreement. Agents also have a set of characteristics or attributes, including age, gender, education level, political affiliation, etc.

We model the flow of information as a set M of information sources, each of which is represented by a binary variable I_j , where $I_j=1$ indicates that the source expresses and supports a desired point of view or desired behavior, and $I_j=0$ indicates that it expresses the exact opposite. Each information source also has a set of characteristics or attributes, which may include things such as media, authorship, design.

Agents and information sources interact with each other at successive discrete moments in time, allowing their behavior to be modeled using the following set of rules or algorithms. The following rules are proposed:

1. Dynamics of the agent's beliefs. At each time step, each agent updates its belief based on the information received. This can be modeled using a simple rule, for example:

$$Y_i(t+1) = \max\{I_j(t) * w_{ij} + (1 - I_j(t)) * w'_{ij}\} \text{ for all } j, \quad (1)$$

where Y_i – belief variable,

I_j – binary variable of the desired topic,

w_{ij} – the weight or importance of information source j for agent i ,

w'_{ij} – the weight of the opposite point of view for this source.

These weights can be based on factors such as the perceived credibility, relevance, or resonance of the information source with the agent's characteristics or attributes.

2. Selecting a source of information. At each time step, each agent chooses for himself those sources of information that he should pay attention to. This rule can be formally written as follows:

$$I_j(t+1) = \begin{cases} 1 & \text{if } \sum_i w_{ij} \cdot Y_i(t) \geq \text{threshold}_j, \\ 0 & \text{else,} \end{cases} \quad (2)$$

where threshold_j is a threshold value representing the minimum level of collective influence required for a source of information to be considered relevant or influential to agents.

This threshold can be based on factors such as the size or diversity of the audience, the prominence or influence of the source, or the contextual relevance of the topic.

3. Manipulations with information sources. At each time step, each information source can manipulate the information it provides to influence agents' beliefs. This approach allows us to model this using the rule:

$$I_j(t+1) = \begin{cases} 1 & \text{if } \sum_i w_{ij} \cdot Y_i(t) \geq \text{threshold}_i - M_{ij}, \\ 0 & \text{else,} \end{cases} \quad (3)$$

where M_{ij} is the level of manipulation, representing the degree of intentional bias or distortion introduced by the source of information. This level can be based on factors such as the incentives or goals of the source, the level of confidence or accountability in the informa-

tion ecosystem, or the influence of external factors such as rumors or disinformation.

Modeling the behavior of these agents and information sources over a selected time interval allows analyzing the emergent dynamics of the system and identifying patterns or trends that can be used to inform interventions or countermeasures against deliberate manipulation of information.

5. 2. Formation of a threat classifier based on combining with threats of social engineering methods

To form a threat classifier, we will use the approaches presented in [47–51].

At the same time, we will introduce a classification of threats based on social engineering, which will increase the level of objectivity of possible social threats and ensure synergy in the formation of a dynamic model for the formation of security systems. Fig. 4 presents the proposed classification of threats.

The presented classification of social engineering threats (Fig. 4) allows us to formulate the main phases of their integration with targeted attacks and form a unified threat classifier (Fig. 5). In addition, this approach will make it possible to form a mathematical apparatus for the formation of multi-loop security systems.

This takes into account the multi-loop aspects of socio-cyberphysical systems, their multi-platform nature and the synthesis of various technologies.

To formally describe the mathematical apparatus of the cybernetic threat model, we will use the approach in [52], which will allow us to form an objective assessment of threats with signs of synergy and hybridity in multi-loop systems – socio-cyberphysical systems.

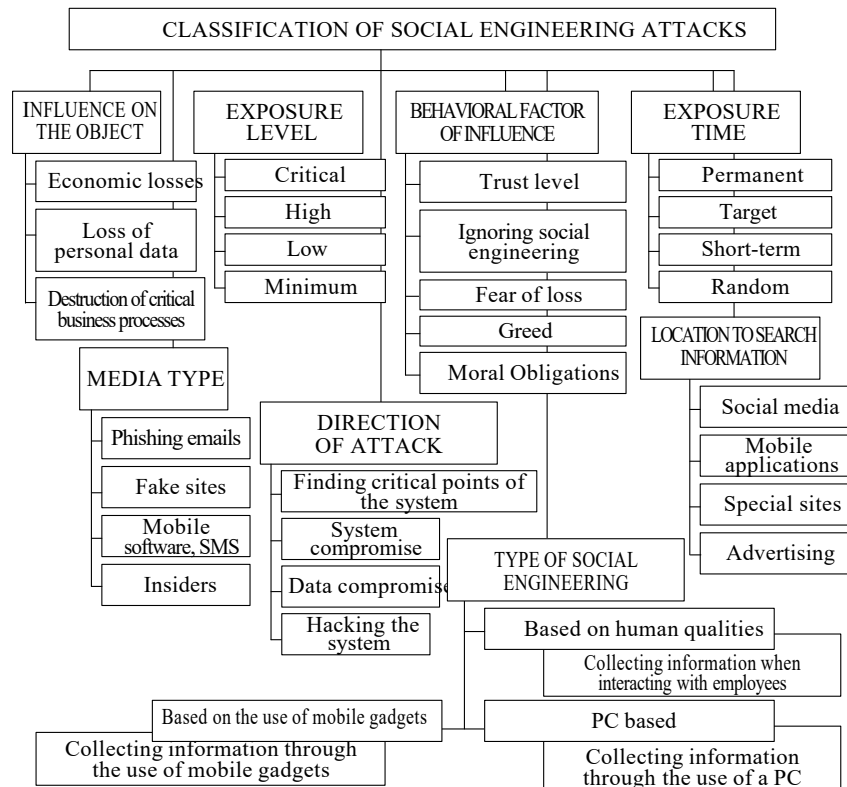


Fig. 4. Classification of threats based on social engineering methods

To ensure the security of the entire protection system, it is necessary to take into account the threats of the internal and external loops for each of the platforms, taking into account their integration with threats based on social engineering methods (Fig. 4):

– internal loop threats, taking into account the hybridity and synergy of threats for platform 1 – social networks:

$$\begin{aligned}
 &W_{\text{hybrid } C, J, A, Au, Af \text{ synerg}_{1\text{platform}}}^{SS \text{ ISL}} = \\
 &= \left(W_{\text{synerg}_{1\text{platform}}}^{CS \text{ ISL}} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL}} \cap \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL}} \cap \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL}} \cap \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL}} \cap \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right), \tag{4}
 \end{aligned}$$

where $W_{\text{synerg}_{1\text{platform}}}^{CS \text{ ISL}}$ is the synergy of threats to the confidentiality service; $\text{soc. eng}_{.1}$ – finding critical points of the system, $\text{soc. eng}_{.2}$ – system compromise, $\text{soc. eng}_{.3}$ – data compromise, $\text{soc. eng}_{.4}$ – system hacking, $\text{soc. eng}_{.5}$ – collection of information, α_i – weighting factor of the possibility of implementing a threat based on social engineering methods, $i \in \{0.25; 0.5; 0.75; 1.0\}$, where 0.25 is the probability of using a threat based on social engineering methods once a year (low level), 0.5 is the probability of using a threat based on social engineering methods once a month (medium level), 0.75 – probability of using a threat based on social engineering methods once a week (high level), 1.0 – probability of using a threat based on social engineering methods once a day (critical level); $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL } I}$ – synergy of threats to the integrity service; $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL } A}$ – synergy of threats to the availability service; $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL } Au}$ – synergy of threats to the authenticity service; $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL } Af}$ – synergy of threats to the affiliation service; $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ISL } Inv - Af}$ – synergy of threats to the affiliation service; – internal loop threats, taking into account the hybridity and synergy of threats for platform 2 – cyberspace:

$$\begin{aligned}
 &W_{\text{hybrid } C, J, A, Au, Af \text{ synerg}_{2\text{platform}}}^{CS \text{ ISL}} = \\
 &= \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } C} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } I} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } A} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } Au} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } Af} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right), \tag{5}
 \end{aligned}$$

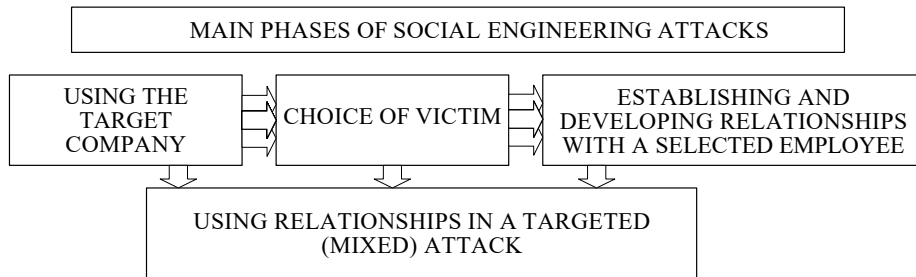


Fig. 5. Main stages of attacks based on social engineering methods

where $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } C}$ is the synergy of threats to the confidentiality service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } I}$ is the synergy of threats to the integrity service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } A}$ is the synergy of threats to the availability service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } Au}$ is the synergy of threats to the authenticity service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ISL } Af}$ is the synergy of threats to the affiliation service; – internal loop threats, taking into account the hybridity and synergy of threats for platform 3 – cyber-physical systems:

$$\begin{aligned}
 &W_{\text{hybrid } C, J, A, Au, Af \text{ synerg}_{3\text{platform}}}^{CPS \text{ ISL}} = \\
 &= \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL}} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } I} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } A} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } Au} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 &\cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } Af} \cup \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right), \tag{6}
 \end{aligned}$$

where $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } C}$ is the synergy of threats to the confidentiality service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } I}$ is the synergy of threats to the integrity service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } A}$ is the synergy of threats to the availability service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } Au}$ is the synergy of threats to the authenticity service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ISL } Af}$ is the synergy of threats to the affiliation service.

General assessment of threats to the internal loop, taking into account the technologies of the socio-cyber-physical system and threats based on social engineering methods:

$$\begin{aligned}
 &W_{\text{ISL}}^{CPSS} = W_{\text{hybrid } C, J, A, Au, Af \text{ synerg}_{1\text{platform}}}^{SS \text{ ISL}} \cup \\
 &\cup W_{\text{hybrid } C, J, A, Au, Af \text{ synerg}_{2\text{platform}}}^{CS \text{ ISL}} \cup \\
 &\cup W_{\text{hybrid } C, J, A, Au, Af \text{ synerg}_{3\text{platform}}}^{CS \text{ ISL}}. \tag{7}
 \end{aligned}$$

General assessment of threats to the internal loop, taking into account the form of ownership of elements and technologies of the socio-cyberphysical system and threats based on social engineering methods (Fig. 4):

$$W_{\text{ISL}}^{CPSS} = W_{\text{ISL}_{\text{private}}}^{CPSS} \cup W_{\text{ISL}_{\text{state}}}^{CPSS} \cup W_{\text{ISL}_{\text{corporativ}}}^{CPSS}, \tag{8}$$

$W_{\text{ISL}_{\text{private}}}^{CPSS}$ – general assessment of internal threats to the personal property system;

$W_{ISL_{state}}^{CPSS}$ – general assessment of internal threats to the state property system;
 $W_{ISL_{corporativ}}^{CPSS}$ – general assessment of internal threats to the corporate property system;
 – external loop threats, taking into account hybridity and synergy for platform 1 – social networks:

$$\begin{aligned}
 & W_{\text{hybrid } C, I, A, Au, Af \text{ synerg}_{1\text{platform}}}^{SS \text{ ESL}} = \\
 & = \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL}} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL}} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right)^I \cap \\
 & \cap \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL}} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right)^A \cap \\
 & \cap \left(W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL}} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right)^{Au} \cap \\
 & \cap \left(W_{\text{synerg}_{1\text{platform}}^3}^{SCPS \text{ ESL}} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right), \quad (9)
 \end{aligned}$$

where $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL } C}$ is the synergy of threats to the confidentiality service, $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL } I}$ is the synergy of threats to the integrity service, $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL } A}$ is the synergy of threats to the availability service, $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL } Au}$ is the synergy of threats to the authenticity service, $W_{\text{synerg}_{1\text{platform}}}^{SS \text{ ESL } Af}$ is the synergy of threats to the affiliation service; *soc. eng.*₁ – finding critical points of the system, *soc. eng.*₂ – system compromise, *soc. eng.*₃ – data compromise, *soc. eng.*₄ – system hacking, *soc. eng.*₅ – collection of information, α_i – weighting factor of the possibility of implementing a threat based on social engineering methods, $i \in \{0.25; 0.5; 0.75; 1.0\}$, where 0.25 is the probability of using a threat based on social engineering methods once a year (low level), 0.5 is the probability of using a threat based on social engineering methods once a month (medium level), 0.75 – probability of using a threat based on social engineering methods once a week (high level), 1.0 – probability of using a threat based on social engineering methods once a day (critical level);

– external loop threats, taking into account the hybridity and synergy of threats for platform 2 – cyberspace:

$$\begin{aligned}
 & W_{\text{hybrid } C, I, A, Au, Af \text{ synerg}_{2\text{platform}}}^{CS \text{ ESL}} = \\
 & = W \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } C} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } I} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } A} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } Au} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } Af} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right), \quad (10)
 \end{aligned}$$

where $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } C}$ is the synergy of threats to the confidentiality service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } I}$ is the synergy of threats to the

integrity service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } A}$ is the synergy of threats to the availability service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } Au}$ is the synergy of threats to the authenticity service, $W_{\text{synerg}_{2\text{platform}}}^{CS \text{ ESL } Af}$ is the synergy of threats to the affiliation service;
 – external loop threats, taking into account the hybridity and synergy of threats for platform 3 – cyber-physical systems:

$$\begin{aligned}
 & W_{\text{hybrid } C, I, A, Au, Af \text{ synerg}_{3\text{platform}}}^{CPS \text{ ESL}} = \\
 & = \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL}} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } I} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } A} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } Au} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right) \cap \\
 & \cap \left(W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } Af} \bigcup_{i=1}^5 \text{soc. eng}_i \times \alpha_i \right), \quad (11)
 \end{aligned}$$

where $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } C}$ is the synergy of threats to the confidentiality service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } I}$ is the synergy of threats to the integrity service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } A}$ is the synergy of threats to the availability service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } Au}$ is the synergy of threats to the authenticity service, $W_{\text{synerg}_{3\text{platform}}}^{CPS \text{ ESL } Af}$ is the synergy of threats to the affiliation service.

General assessment of external threats taking into account the technologies of the socio-cyberphysical system:

$$\begin{aligned}
 & W_{ESL}^{CPSS} = W_{\text{hybrid } C, I, A, Au, Af \text{ synerg}_{1\text{platform}}}^{SS \text{ ESL}} \cup \\
 & \cup W_{\text{hybrid } C, I, A, Au, Af \text{ synerg}_{2\text{platform}}}^{CS \text{ ESL}} \cup W_{\text{hybrid } C, I, A, Au, Af \text{ synerg}_{3\text{platform}}}^{CPS \text{ ESL}}. \quad (12)
 \end{aligned}$$

General assessment of external threats, taking into account the form of ownership of the elements and technologies of the socio-cyberphysical system (Fig. 4):

$$W_{ESL_{\text{general}}}^{CPSS} = W_{ESL_{\text{private}}}^{CPSS} \cup W_{ESL_{\text{state}}}^{CPSS} \cup W_{ESL_{\text{corporativ}}}^{CPSS}, \quad (13)$$

where $W_{ESL_{\text{private}}}^{CPSS}$ – general assessment of internal threats to the personal property system;

$W_{ESL_{\text{state}}}^{CPSS}$ – general assessment of internal threats to the state property system;

$W_{ESL_{\text{corporativ}}}^{CPSS}$ – general assessment of internal threats to the corporate property system.

Based on expressions (4), (9), an assessment of threats in socio-cyberphysical systems in the internal and external security loops of CPSS is formed, and based on expressions (5), (10) – taking into account forms of ownership (separately).

To provide a generalized assessment of a multi-loop security system, we use the formula:

$$W_{\text{final}}^{CPSS} = W_{ISL_{\text{general}}}^{CPSS} \cup W_{ESL_{\text{general}}}^{CPSS}. \quad (14)$$

Each element of information resources $I_{A_i} \in \{I_A\}$ can be described by a vector:

$$I_{A_i} = (\text{Type}_i, A_i^C, A_i^I, A_i^A, A_i^{Au}, A_i^{Af}, \beta_i),$$

where $Type_i$ is the type of information asset, described by a set of basic values: $Type_i = \{CI_i, PD_i, CD_i, TS_i, StR_i, PubI_i, ContI_i, PI_i\}$, where CI_i is confidential information, PD_i is payment documents, CD_i is credit documents, TS_i is trade secret, StR_i – statistical reports, $PubI_i$ – public information, $ContI_i$ – control information, PI_i – personal data. $A_i^C, A_i^I, A_i^A, A_i^{Au}, A_i^{Af}$ – security services (A_i^C – confidentiality, A_i^I – integrity, A_i^A – availability, A_i^{Au} – authenticity, A_i^{Af} – affiliation); β_i is a metric for the relationship between time and the degree of information secrecy for an asset (critical – 1.0; high – 0.75; medium – 0.5; low – 0.25; very low – 0.01).

Then the general (current) level of security of socio-cyberphysical systems based on wireless mobile technologies is described by the expression:

– for additive convolution:

$$L_{W_{security}^{CPS}} = L_{ISL} \sum_{j=1}^3 \sum_{i=1}^{12} (I_{A_{ij}} \times \beta_{ij}) + L_{ESL} \sum_{j=1}^3 \sum_{i=1}^{12} (I_{A_{ij}} \times \beta_{ij}); \quad (15)$$

– for multiplicative convolution:

$$L_{W_{security}^{CPS}} = 1 - \left[1 - L_{ISL} \sum_{j=1}^3 \sum_{i=1}^{12} (I_{A_{ij}} \times \beta_{ij}) \right] \times \left[1 - L_{ESL} \sum_{j=1}^3 \sum_{i=1}^{12} (I_{A_{ij}} \times \beta_{ij}) \right]. \quad (16)$$

Fig. 6 presents a tuple of the proposed classifier, which takes into account the synergetic threat model, taking into account the hybridity of cyber threats and their integration with threats based on social engineering methods.

In addition, the multi-loop nature of socio-cyberphysical systems, their multi-platform nature and the integration of its constituent elements are taken into account.

This approach makes it possible to unify not only the construction of a threat classifier, but also to provide a comprehensive assessment of targeted attacks based on the emergent properties of the complex formation of attacks.

5.3. Development of a security model for information interactions in socio-cyberphysical systems

With the unification of the cyber-physical and social space, cyber defense faces the following challenges, which are common to various groups of methods (Fig. 7):

- the need to track information exchange in real time;
- the need to recognize multimodal information, i.e. information not only in the form of text messages, but also voice or gesture information;
- the need to generate and apply patterns to control information interaction between resources based on S-CPS representations.

If analysis is carried out on the basis of damaged, noisy data contained in the general flow of different types of information, the results can be very problematic. Therefore, before using a particular pattern (say, when using neural network training) as part of detection technology, it is necessary to ensure that the data is not noisy, or the noise level is reduced to a minimum, and that the information redundancy of the message is minimized.

Data sets from different domains typically contain data defined by a wide range of attributes, among which there are varying degrees of correlation. Identifying data objects that do not correspond to these hidden correlations is challenging. Moreover, attributes can often play different roles in applications. In particular, some features can be perceived as independent variables that are responsible for determining the context in which the dependent variable exhibits abnormal behavior.

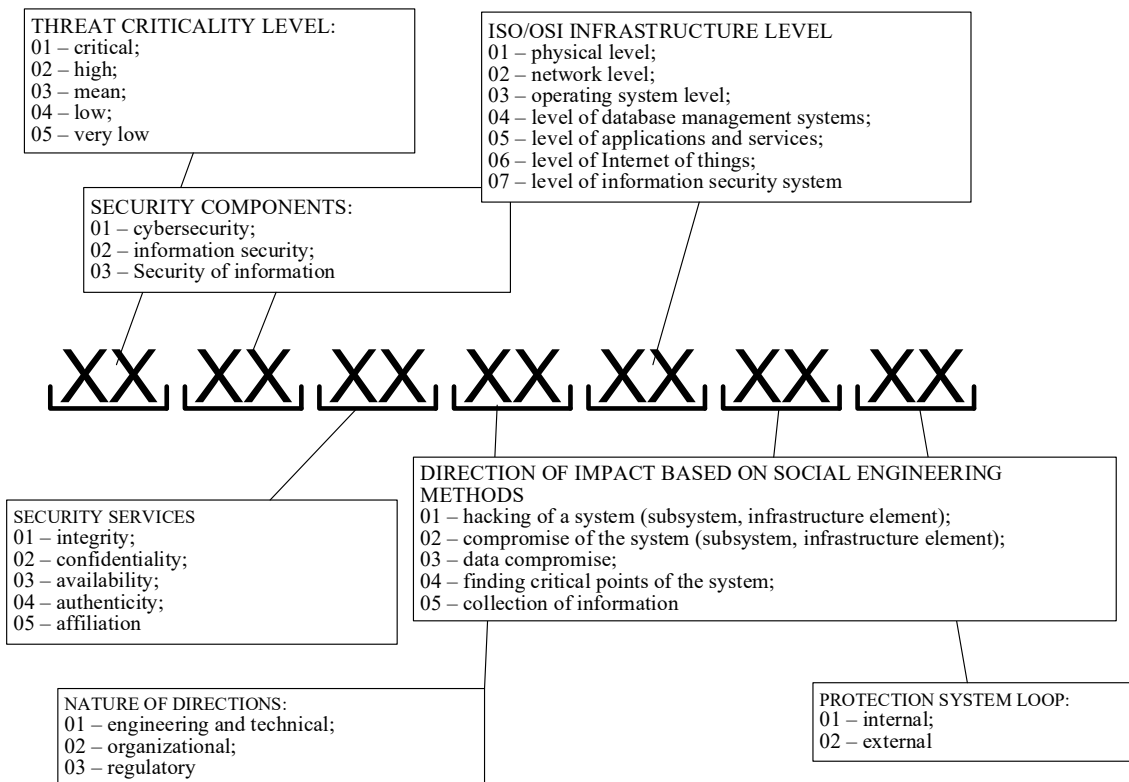


Fig. 6. Classifier of synergistic threats with aggregation of threats based on social engineering methods

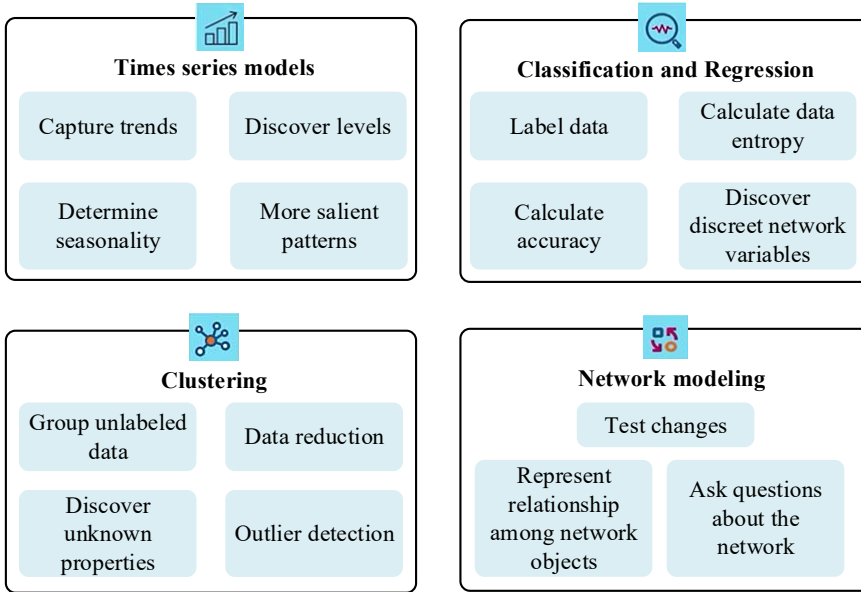


Fig. 7. Detection and correlation of artifacts in cybersecurity systems

Consequently, the work focuses on detecting data objects that exhibit abnormal behavior in a subset of attributes called behavioral, in relation to them some others called contextual. As a major contribution, a model is proposed to describe the correlation laws hidden in data distributions across pairs of behavioral and contextual attributes. A probabilistic measure is introduced aimed at assessing subsequently observed objects based on how much their behavior deviates from the detected correlation laws [53].

Correlation between these attributes exists both at the topological level and at the level of semantic content of data sets, as well as in the time domain. Thus, the result of this stage will be a set of clusters of correlated data. These clusters indicate events that correlate with each other in terms of temporal, semantic, and topological dimensions.

A mathematical model of this process can be considered as a model, using which it is necessary to determine manipulative factors based on observed patterns of beliefs or opinions of the target audience:

Suppose, as in the previous model, we have a set of N agents, each of which has a binary belief variable Y_i . There is also a set M of potential manipulative factors that may reflect phenomena such as media framing, linguistic features, visual cues, or social influence.

Manipulative factors can be represented using a set of binary variables X_j , where $X_j=1$ indicates the presence of a potential manipulative factor, and $X_j=0$ indicates its absence. For example, the presence of a particular framing style can be represented using a binary variable, where $X_j=1$ indicates that framing is present and $X_j=0$ indicates its absence.

Thus, it is necessary to determine which manipulative factors are most strongly associated with the observed patterns of beliefs or opinions of the target audience. This relationship can be represented as a logistic regression model, where the probability of an agent having a certain belief is presented as a function of manipulative factors:

$$P(Y_i = 1) = \text{logistic}(b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n), \quad (17)$$

where X_m is a set of binary variables corresponding to manipulative factors, $b_0, b_1, b_2, \dots, b_n$ are coefficients of the logistic

regression model, logistic is a logistic sigmoid function that maps a linear combination of manipulative factors to probabilities in the range from 0 to 1:

$$\text{logistic}(z) = \frac{1}{1 + \exp(-z)}. \quad (18)$$

The parameters $b_0, b_1, b_2, \dots, b_m$ are the coefficients of the logistic regression model and can be estimated using maximum likelihood or other methods.

Once the coefficients of the logistic regression model have been estimated, we can proceed to identifying the manipulative factors that are most strongly associated with the observed patterns of beliefs or opinions in the target audience. This can be done by examining the magnitude

and sign of the coefficients, which indicate the strength and direction of the relationship between each manipulative factor and the likelihood of the agent having a particular belief.

By identifying the manipulative factors that are most closely related to observed patterns of beliefs or opinions, interventions or countermeasures can be developed to reduce the influence of the manipulative factors and promote balanced decision-making.

To perform these actions, both entropy analysis of information flows and extraction of relevant data using a library of patterns can be used. Let's assume that we have a flow of information consisting of a set of N messages. Each message is characterized by a set of attributes that describe its content and context. Each message can be represented as a feature vector, where each feature is a binary variable indicating the presence or absence of a certain characteristic.

We also define a set of patterns M or templates that capture the key characteristics of messages associated with certain topics. For example, we may have a pattern that captures the key features of politics-related posts, another for sports-related posts, and so on.

To analyze the entropy of an information flow and extract relevant data, a technique such as pattern matching or clustering can be used. This allows you to identify messages that match library patterns. The pattern library can be represented as a matrix P , where each row corresponds to a pattern and each column corresponds to a function:

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & \dots & p_{1,F} \\ p_{2,1} & p_{2,2} & \dots & p_{2,F} \\ \vdots & \vdots & & \vdots \\ p_{M,1} & p_{M,2} & \dots & p_{M,F} \end{bmatrix}, \quad (19)$$

where $p_{i,j}$ is a binary variable indicating the presence or absence of feature j in sample i .

To identify messages that are similar to patterns in the proposed library, we represent messages as feature vectors and calculate their similarity to each pattern using a measure such as cosine similarity or Jaccard similarity. For example, when representing a message as a vector M , its

cosine similarity to each pattern in the proposed library is calculated as follows:

$$similarity(M, P_i) = \frac{dot(M, P_i)}{(norm(M) \cdot norm(P_i))}, \quad (20)$$

where $dot(M, P_i)$ is the scalar product of the vectors M and P_i , $norm(M)$ and $norm(P_i)$ are the Euclidean norms of the vectors M and P_i .

This similarity measure can be used to identify the top k patterns that best match each message and use them to extract relevant data or insights. For example, if a message most closely matches a policy pattern, you can infer that it contains political content and add it to the political message database.

To analyze the entropy of the information flow, you can also calculate the entropy of pattern distribution in the library. For example, the entropy H can be calculated as:

$$H = -\sum_i p_i \cdot \log_2(p_i), \quad (21)$$

where p_i is the proportion of messages similar to pattern i , \log_2 is the base 2 logarithm.

By analyzing the entropy of pattern distribution, one can get an idea of the diversity and balance of topics in the information flow. For example, low entropy may indicate that the information flow is dominated by a few specific topics, while if entropy is high, the information flow is diverse and covers a wide range of topics.

Thus, the constructed mathematical model allows us to analyze the entropy of information flows and extract the corresponding data using a library of patterns. This approach can be useful for understanding the dynamics of information dissemination and developing effective interventions or countermeasures.

To identify the correlation of the received data with the time of receipt, text content and computer network topology according to previously obtained clusters, it is proposed to use a mathematical model based on clustering algorithms. The construction of such a model is a sequence of the following steps.

Let there be a data set of N messages. Each message is characterized by three main features: time of receipt t , message content C and computer network topology T . Each message can be represented as a three-dimensional vector:

$$M_i = (t_i, C_i, T_i), \quad (22)$$

where t is the time of receipt, C is the text content, T is the computer network topology.

To group messages based on their similarity, it is suggested to use a distance metric such as Euclidean distance or cosine distance to calculate the distance between each pair of messages:

$$d(M_i, M_j) = \sqrt{(t_i - t_j)^2 + d_C(C_i, C_j)^2 + d_T(T_i, T_j)^2}, \quad (23)$$

where $d_C(C_i, C_j)$ and $d_T(T_i, T_j)$ are the distances between content and topology elements, respectively.

Distances for categorical features are calculated using the Jaccard index or Hamming distance, and for continuous features, it is cosine similarity or Euclidean distance.

After calculating the distance between all pairs of messages, it is possible to use a clustering algorithm such as k -means or hierarchical clustering to group similar messages into clusters. The algorithm works by iteratively assigning

each message to the nearest cluster center and updating the cluster centers based on the new assignments.

The number of clusters k can be determined using methods such as the elbow method or the silhouette method. These methods aim to find the value of k that maximizes the similarity within a cluster and minimizes the similarity between clusters.

Once messages are grouped, it becomes possible to analyze the characteristics of each cluster to draw conclusions about correlations between the time of receipt, text content, and computer network topology. Next, to identify a pattern or trend, it is necessary to calculate the average time of receipt, content similarity, and network topology similarity for each cluster and compare them with each other.

To analyze clusters using cause-and-effect relationships, it is proposed to use a Bayesian network, which is a probabilistic graphical model representing cause-and-effect relationships between variables.

Building a mathematical model using Bayesian networks can be done as follows.

Suppose messages are grouped into k clusters and a set of variables is defined that can affect the time of receipt, text content, and computer network topology (such as sender, recipient, message type, and network location). Each variable can be represented as a node in a Bayesian network, where the edges between the nodes represent the cause-and-effect relationships between the variables.

For example, we can assume that the sender has a causal effect on the text content and network topology, and the time of message has a causal effect on the time of receipt and text content. This hypothesis can be represented as a Bayesian network with the following nodes: sender, message type, time of receipt, content, network topology.

Since the edges between nodes indicate cause-and-effect relationships between variables, it can be assumed that the sender has a causal effect on the nodes of content formation and network topology, which is reflected by the edges of the graph from the sender node to the content and topology nodes.

To determine cause-and-effect relationships between variables, conditional probabilities represented by a Bayesian network are used. For example, you can calculate the probability of receiving the contents of a message given the sender and topology:

$$P(Content | Sender, Topology).$$

The probability is estimated from the data by counting the frequency of each combination of variables in the dataset and normalizing them by the total number of messages.

Once conditional probabilities are determined, they can be used to predict cause-and-effect relationships between variables. For example, it is possible to predict the probability of receiving a message from a particular sender, given the content and topology:

$$P(Sender | Content, Topology).$$

This probability can be used to analyze the causal effect of the sender on the content and topology of messages.

Options for displaying this information may include Bayesian network and conditional probability visualizations, as well as pre-actions and explanations of cause-and-effect relationships between variables. This allows you to obtain a list of the most likely senders given a specific content and topology, or a list of the most likely content given a specific sender

and time of receipt. In this case, it becomes possible to obtain an explanation of cause-and-effect relationships, for example, “messages sent by sender *A*, as a rule, have a topology similar to messages sent by sender *B*, but different content”.

Thus, the proposed version of the security model of information interactions in security-cyberphysical systems provides cluster analysis using cause-and-effect relationships. This approach can be represented as a Bayesian network, where

nodes represent variables and edges represent cause-and-effect relationships between them. Conditional probabilities represented by the network can be used to make predictions and explanations of cause-and-effect relationships. Options for displaying this information may include visualizations, predictions, and explanations of cause-and-effect relationships between variables. Fig. 8 shows a security model of information interactions in socio-cyberphysical systems.

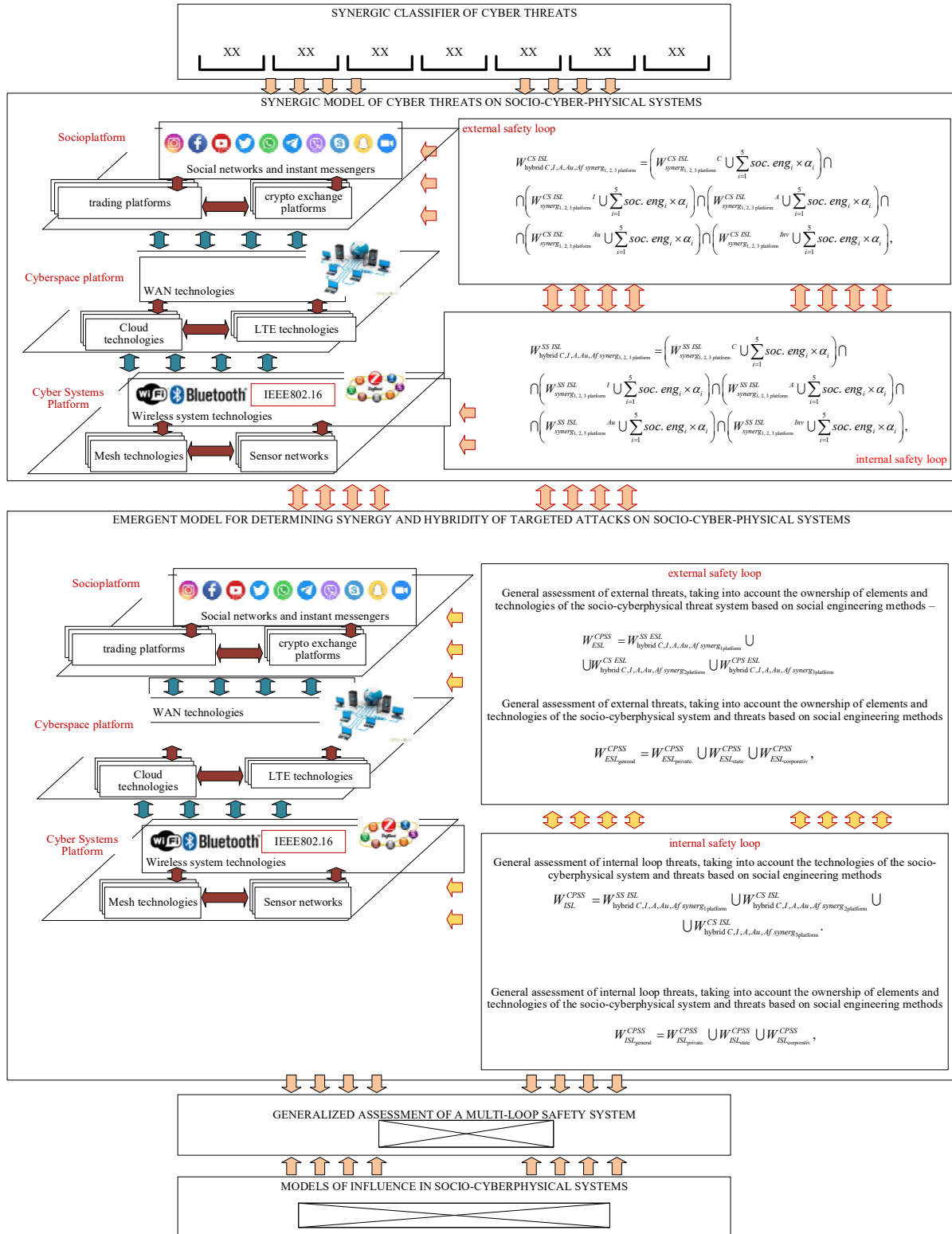


Fig. 8. Security model of information interactions in socio-cyberphysical systems

The proposed security model of information resources in socio-cyberphysical systems makes it possible to form a multi-loop security system for information interaction in socio-cyberphysical systems. This takes into account the synthesis of technologies and interaction channels; for each platform of socio-cyberphysical systems it is proposed to form internal and external security loops, which will allow taking into account the interaction of information flows, both within the platform and external. This approach will create the necessary level of objectivity in the analysis of possible cyber-physical (targeted or mixed) threats, taking into account possible integration with threats based on social engineering methods. In addition, timely identify critical points in the infrastructure of each platform and formulate preventive protection measures.

5. 4. Study of the proposed security model of information interactions in socio-cyberphysical systems

To analyze the proposed model of influence in socio-cyberphysical systems, it is necessary to perform simulation modeling of the described processes. When developing a simulation model, we will take into account some characteristics of agents and features of their behavior when interacting in socio-cyberphysical systems.

Modeling the dynamics of opinions is based on the use of so-called agent-based thinking [54]. Using this approach, it is possible to define the rules of interaction between individual agents within the model and allow social influence to spread throughout the system. This allows complex models to be built based on relatively simple rules. If the simulation results are realistic, it can be argued that the proposed rules are sufficient to create realistic emergent behavior. In other words, this behavior can be explained by the proposed mechanism.

Agents are any objects that fill a model that implements agent-based thinking. This is consistent with individual agent-based models (ABM) in the taxonomy proposed in [55].

One behavior of the system-level opinion dynamics model that can have a significant influence on information impacts is agent scheduling. Namely: what agent(s) and in what order influence (or are influenced by) what other agents at each discrete point in time. Let us include two types of agent schedules in the model: a schedule with repeated averaging (all agents are simultaneously influenced by all others they are associated with) and a schedule with limited reliability (there is one pair of agents. Thus, the influence of agents on each other in the model is simultaneously formed or all agents are simultaneously influenced by all others, subject to confidence restrictions).

The next characteristic is the taxonomy “Synchrony, subject type, scale”. This is a short method for conveying the schedule of the opinion dynamics model. Synchrony refers to whether states are constantly updated as agents act, and has two options: synchronous and asynchronous. The type of subject refers to the direction of influence and has four options: target, source, group and mixed. Scale refers to the number of actors selected for each role per time step.

Synchrony determines whether each agent’s state updates occur in parallel or sequentially. A model in which all agent updates occur in parallel will be called synchronous. A model in which some or all agent updates occur sequentially is called asynchronous.

When defining a model built on the basis of agent-based thinking, we will consider three main types of actors. Namely: source agents who influence others when they act, target agents who are influenced by others when they act, and groups of agents

who mutually influence each other. The choice of agent type affects the schedule and the order in which the influence occurs.

If the primary actors are source agents, some set of secondary actors (i.e. targets) is selected for each primary actor. In the asynchronous model, this source influences all of its targets before another source acts. If the primary actors are target agents, a certain set of secondary actors (i.e. sources) is selected for each primary actor. In the asynchronous model, this target is influenced by all of its sources before another target takes effect.

If primary actors are groups of agents, there are no secondary actors; each group influences its member agents as defined by the model. Each group action can be represented as an opinion dynamics model running on a subset of agents, so these graphs can be further refined using taxonomy. Collision rules may be required for synchronous models where one agent can be selected as a member of multiple groups.

If the primary actors are of mixed types, some or all of the above types of actors exist and act within the model. It is necessary to indicate whether actors of a given type act before actors of another type, representing the sequential application of several opinion dynamics models within a time step, or whether they act in a mixed order, representing a truly mixed opinion dynamics model.

To demonstrate the use of the taxonomy and potential differences that may arise in model results due to different schedules, we examine two models.

The first of these is the repeated averaging model. The repeated averaging model uses a linear transformation of the agents’ opinion vector to perform discrete-time updates. This model tends to converge under reasonable conditions. The modeling process is considered to have converged when the spread of opinions falls below 0.01. Graphs of opinion dynamics are shown in Tables 1–3. The agent’s volatility indicator μ was introduced as a model parameter, which reflects the agent’s willingness to change his opinion. A zero value of the parameter μ reflects the agent’s inability to change his opinion, and a unit value reflects the agent’s complete readiness to change his opinion under outside influence. In the graphs presented in Table 1, the axes correspond to the following variables. The horizontal axis corresponds to the number of information exchanges as a result of interaction between agents, the vertical axis reflects the value of two variables: the average opinion for a group of agents (mean opinion) and the dispersion (scatter) of opinions in a group of agents (range opinion).

Thus, the analysis of the results of Tables 1–3 showed that in the case of agents who are sources of influence, the convergence of the information influence process occurs at higher values of the volatility level (susceptibility to influence and readiness to change one’s beliefs). At lower values of the volatility coefficient, the convergence of the influence process occurs at later points in time. With a fixed level of volatility within a certain type of agent, changes when changing the mode of influence (synchronous or asynchronous) are practically insignificant.

The second model of influence dynamics in socio-cyberphysical systems is the generalized model of limited confidence. The model differs in one key point – after initializing a set of messages and network topologies, they are filtered, allowing the inclusion of only secondary participants whose opinions are within the confidence threshold of the main actor. An individual agent, if influenced at time t , influences other agents while maintaining a certain level of self-confidence, which can be changed before the start of the simulation. This ensures the convergence of clusters of agents having the same formed beliefs (points of view).

Table 1

Analysis of synchrony and bias at $\mu=0.5$

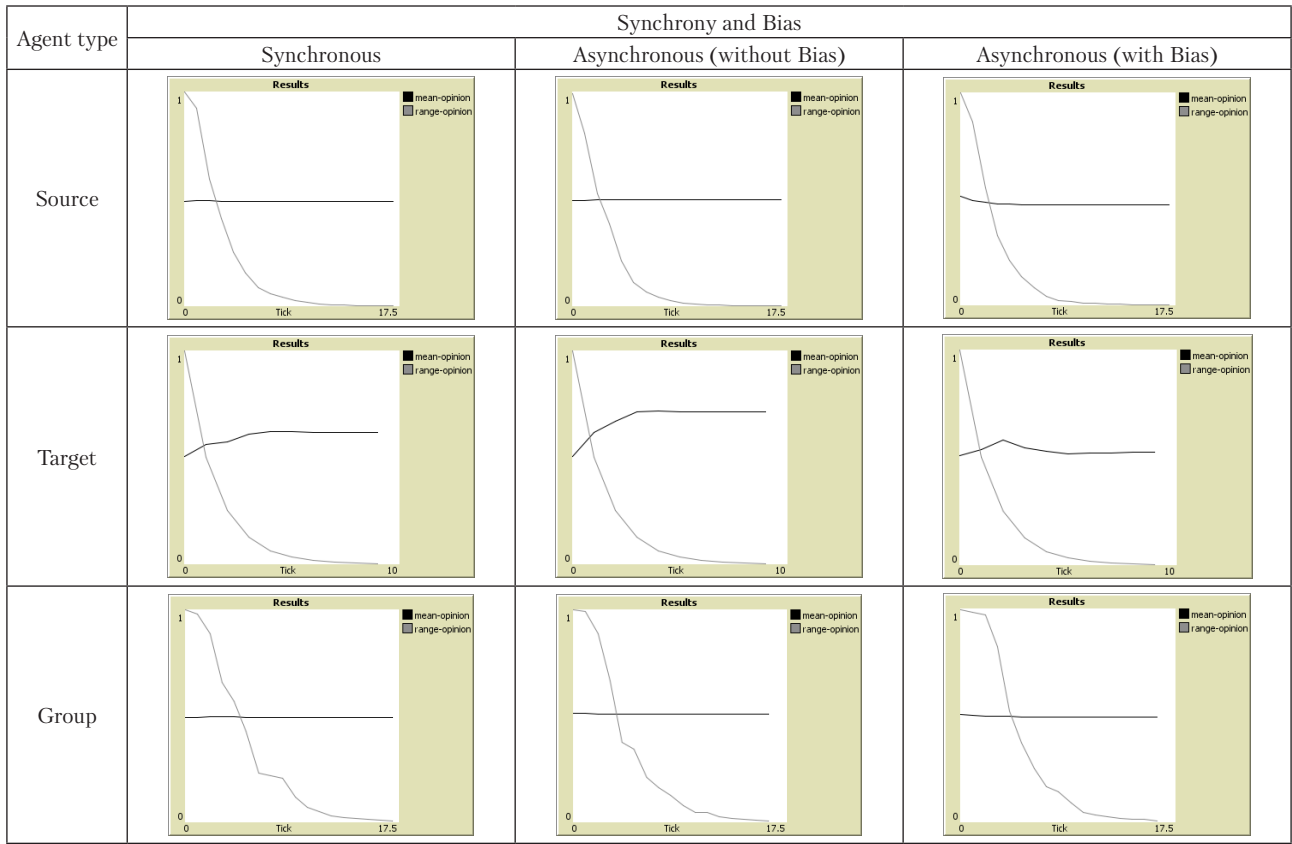
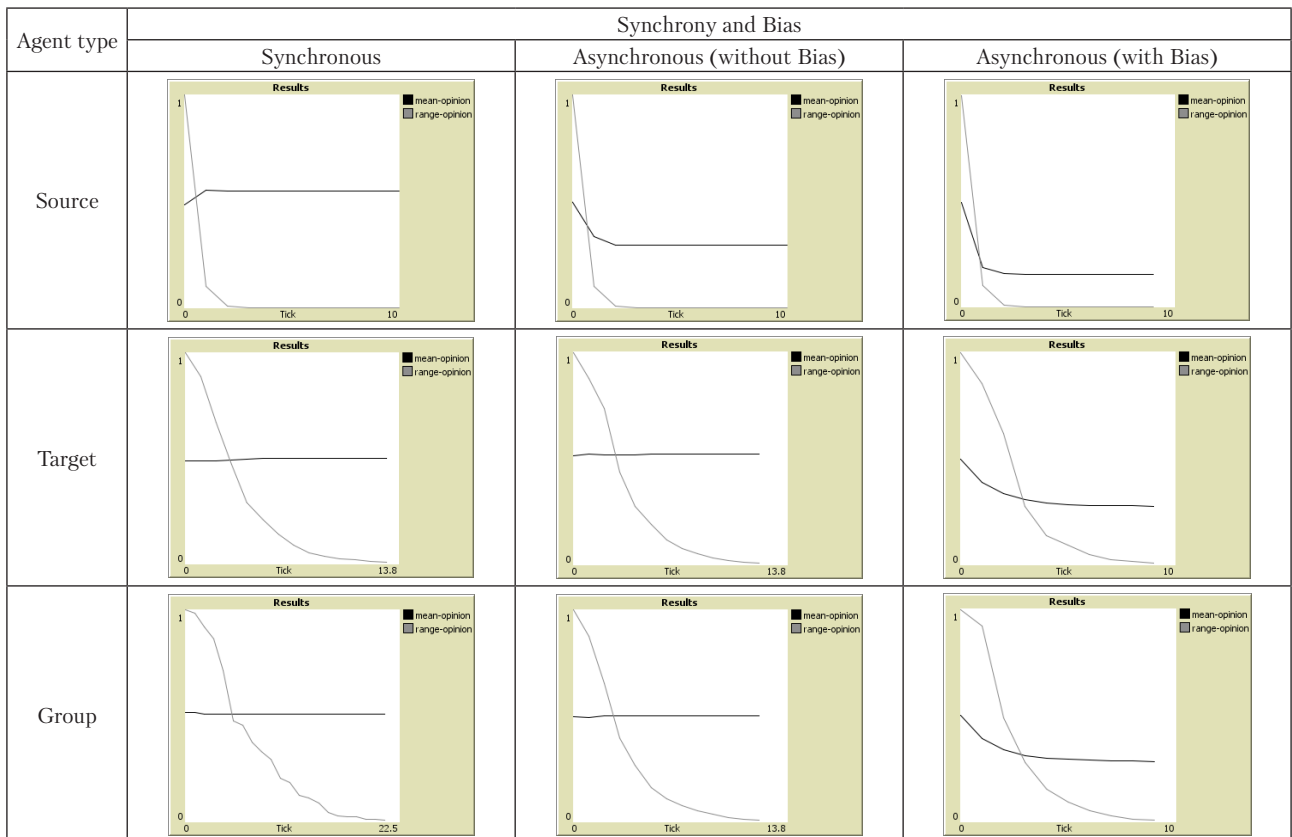


Table 2

Analysis of synchrony and bias at $\mu=0.9$



In Table 4 presents the simulation results regarding the formation of agent clusters. For all simulated situations, the level of volatility (i.e. willingness to change one's opinion) was chosen to be 0.5. This means that with a probability of 0.5, the

agent is ready to change his opinion under the influence of the agents around him. The horizontal axis still corresponds to the number of information exchanges, and the vertical axis reflects the conditional index of the opinions of agents in the cluster.

Table 3

Analysis of synchrony and bias at $\mu=0.1$

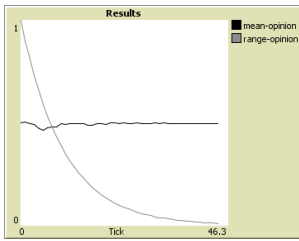
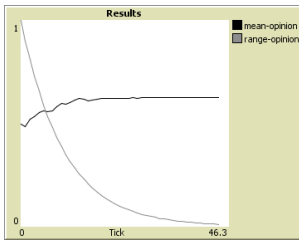
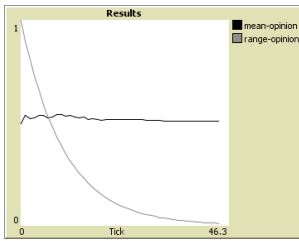
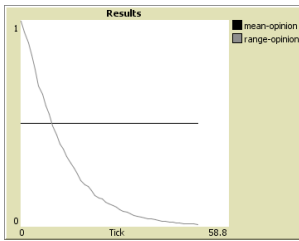
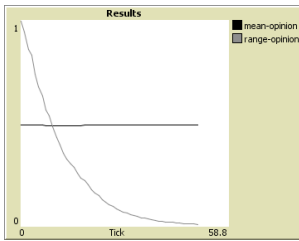
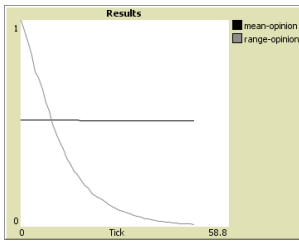
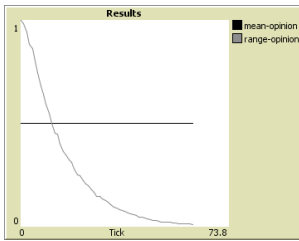
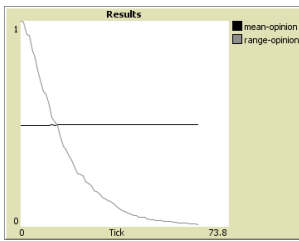
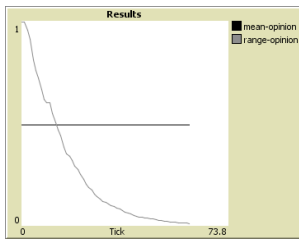
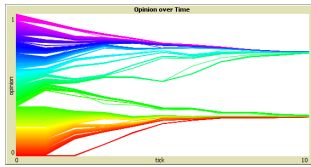
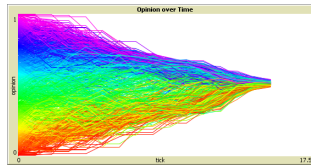
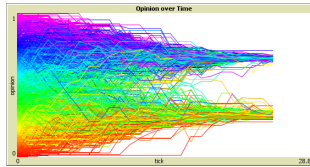
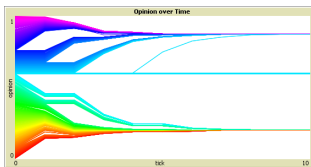
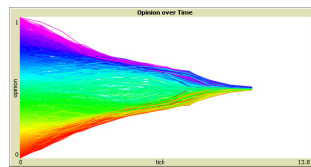
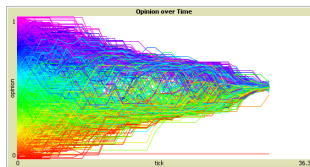
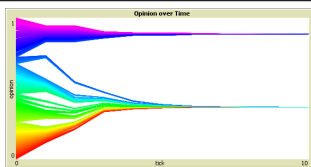
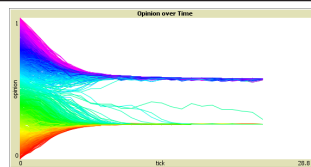
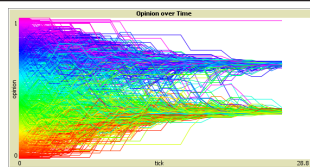
Agent type	Synchrony and Bias		
	Synchronous	Asynchronous (without Bias)	Asynchronous (with Bias)
Source			
Target			
Group			

Table 4

Results of modeling the processes of forming clusters of information interaction agents

Schedule Type/Agent Type		
Synchronous		
Influence Source	Influence Target	Group
		
Asynchronous without Bias		
Influence Source	Influence Target	Group
		
Asynchronous with Bias		
Influence Source	Influence Target	Group
		

Thus, the analysis of Table 4 showed that for agents who are sources of influence, the formation of their stable clusters occurs more quickly in the case when these agents have certain biases and interaction occurs in an asynchronous mode. In addition, for agents who are the target of information influence. The difference is that this requires a longer period of influence. For a group of agents, the formation of a stable cluster occurs most quickly in the case of asynchronous interaction, when the agents in the group have biases.

6. Discussion of the results of creating a multi-loop security system for information interactions in socio-cyberphysical systems

Conducted research on the creation of a multi-level security system for information interactions in socio-cyberphysical systems (Tables 3–4) allows us to form an objective assessment of the possible influence of source agents based on social engineering methods.

A significant difference of the proposed approach is the presence of a classifier of threats to socio-cyberphysical systems, which makes it possible to increase the level of objectivity in assessing the integration of targeted (mixed) attacks with threats based on social engineering methods. The proposed approach to generating classifier tuples is unified and intuitive, which allows its use in various areas of IT technologies and systems. The difference from known threat assessment methods [49–51] is the formation of a model based on a comprehensive threat assessment (with signs of synergy and hybridity) with threats based on social engineering methods. This will allow us to obtain a more objective assessment of the impact not only on the “victim” when implementing a targeted attack, but also to timely implement preventive measures taking into account measures to counteract social engineering methods.

The developed mathematical apparatus makes it possible to take into account the formation of multi-loop systems, taking into account the specifics of software (hardware and software) tools and mechanisms of both socio-cyberphysical systems and critical infrastructure objects. In this case, it is proposed to use the division of socio-cyberphysical systems into platforms: platform 1 – cyber systems, platform 2 – cyberspace, platform 3 – social networks. The cyber systems platform, as a rule, contains sensor, video surveillance systems, sensors, etc., as well as elements of mechanisms for performing tasks from the control system. The cyberspace platform is usually formed on the basis of cloud technologies and allows you to manage elements of the cyber systems platform, as well as interact with the social networking platform. The social platform circulates both management and general information in instant messengers and social networks. Thus, the proposed classifier takes into account the signs of synergy and hybridity of targeted attacks, as well as the possibility of combining with threats based on social engineering methods. This approach differs from well-known approaches [52, 56–64] by the ability to take into account the entire variety of targeted attacks combined with threats based on social engineering methods, considering the signs of synergy and hybridity.

Comparison of the proposed approach with known studies leads to the following conclusions.

The general conclusion is that well-known studies have focused exclusively on modeling and analyzing the dynamics

of information influence in terms of social research, without taking into account the influence of processes on the overall security level of the socio-cyberphysical system. In particular, in [65], two agents simultaneously influence their immediate neighbors in the same dimension. In repeated averaging models, all agents are simultaneously influenced by all others they are associated with, and in limited confidence models, there is one pair of agents simultaneously influencing each other [66] or all agents are simultaneously influenced by all others in the model [67] subject to confidence restrictions. A number of publications do not consider the issue of scheduling the work of agents [55, 68], although the geographic location of the agent is taken into account. This approach in the context of a socio-cyberphysical system is a significant drawback.

The main disadvantage of the proposed approach is the difficulty of assessing the statistical data of social engineering methods and the impact of threats based on social engineering on the implementation of targeted attacks.

As part of the simulation, it is necessary to take into account that agent-based modeling requires significant computing resources, which can increase exponentially depending on the increase in the number of agents. In addition, this approach to modeling “forces” restrictions on the behavior of the system as a whole when studying the agents themselves. At the same time, the presented classifier will allow us to systematize and form hybrid and/or synergetic threats, complex (general) threats based on the complete probability theorem. This approach can be used in almost any field of information and communication systems, as well as in smart technologies to create multi-loop (multi-level) information security systems, taking into account the integration of targeted threats with social engineering methods (mechanisms).

The proposed approach will further allow us to formulate the concept of building multi-loop information security systems based on the formation of security loops (internal and external), as well as taking into account the physical location and functioning of infrastructure elements and the logical structure of information and communication networks of both socio-cyber-physical systems and critical infrastructure facilities.

7. Conclusions

1. The processes of information influence in socio-cyberphysical systems are analyzed. The analysis showed insufficient research into the information processes of the social component of socio-cyber-physical systems and systems for ensuring their safe functioning. Based on the results of the analysis, a conclusion was made about the relevance of developing methods for assessing these processes.

2. A classification of threats based on social engineering methods is proposed, which makes it possible to form a unified, objective classifier of threats, taking into account the signs of their hybridity and synergy. This approach allows us to ensure the formation of a comprehensive assessment of targeted (mixed) attacks on socio-cyberphysical systems, taking into account the construction of multi-loop information security systems.

3. A security system for information interactions in socio-cyberphysical systems has been developed. Based on

the proposed model, an understanding of the cause-and-effect relationships that exist in the social component of socio-cyberphysical systems is provided, and also an analysis of clusters of system agents in the form of a Bayesian network is given. In general, this approach to developing a security system will allow us to take into account the synthesis of technologies and channels of interaction of circulating information in socio-cyberphysical systems. For each platform of socio-cyberphysical systems, it is proposed to form an internal and external security loop, which will allow taking into account the interaction of information flows, both within the platform and externally. The formed multi-loop security model provides timely counteraction to targeted attacks, taking into account their integration with social engineering threats.

4. A study of the proposed security system of information interactions in socio-cyberphysical systems was carried out in the form of simulation modeling. Qualitative characteristics of the influence mode and quantitative indicators have been determined that ensure more rapid formation of agent

clusters and convergence of the process of information influence on them.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this research, whether financial, personal, authorship or otherwise, that could affect the research and its results presented in this paper.

Financing

The study was conducted without financial support.

Data availability

The manuscript has no associated data.

References

- Graf, S., Quinton, S., Girault, A., Gessler, G. (2018). Building Correct Cyber-Physical Systems: Why We Need a Multiview Contract Theory. *Lecture Notes in Computer Science*, 19–31. doi: https://doi.org/10.1007/978-3-030-00244-2_2
- Bereket Abera, Y., Naudet, Y., Panetto, H. (2020). A new Paradigm and Meta-Model for Cyber-Physical-Social Systems. *IFAC-PapersOnLine*, 53 (2), 10949–10954. doi: <https://doi.org/10.1016/j.ifacol.2020.12.2841>
- Yilma, B. A., Naudet, Y., Panetto, H. (2019). Introduction to Personalisation in Cyber-Physical-Social Systems. *Lecture Notes in Computer Science*, 25–35. doi: https://doi.org/10.1007/978-3-030-11683-5_3
- Yilma, B. A., Panetto, H., Naudet, Y. (2019). A Meta-Model of Cyber-Physical-Social System: The CPSS Paradigm to Support Human-Machine Collaboration in Industry 4.0. *IFIP Advances in Information and Communication Technology*, 11–20. doi: https://doi.org/10.1007/978-3-030-28464-0_2
- Naudet, Y., Yilma, B. A., Panetto, H. (2018). Personalisation in Cyber Physical and Social Systems: the Case of Recommendations in Cultural Heritage Spaces. 2018 13th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP). doi: <https://doi.org/10.1109/smap.2018.8501890>
- Zeng, J., Yang, L. T., Lin, M., Ning, H., Ma, J. (2020). A survey: Cyber-physical-social systems and their system-level design methodology. *Future Generation Computer Systems*, 105, 1028–1042. doi: <https://doi.org/10.1016/j.future.2016.06.034>
- Sheth, A., Anantharam, P., Henson, C. (2013). Physical-Cyber-Social Computing: An Early 21st Century Approach. *IEEE Intelligent Systems*, 28 (1), 78–82. doi: <https://doi.org/10.1109/mis.2013.20>
- Wang, F.-Y. (2010). The Emergence of Intelligent Enterprises: From CPS to CPSS. *IEEE Intelligent Systems*, 25 (4), 85–88. doi: <https://doi.org/10.1109/mis.2010.104>
- Dzheniuk, N., Yevseiev, S., Opirskyy, I., Voropay, N., Korolev, R., Sydorenko, Z. (2023). Sociocyberphysical System Wireless Air Network Topology Synthesis Model. *Mizhnarodnyi naukovo-praktychnyi forum «Tsyfrova realnist»*. Kiberbezpeka ta informatsiyi tekhnolohiyi v umovakh hibrydnykh viyn. Kharkiv-Odesa, 4–10.
- Horváth, I., Rusák, Z., Li, Y. (2017). Order Beyond Chaos: Introducing the Notion of Generation to Characterize the Continuously Evolving Implementations of Cyber-Physical Systems. Volume 1: 37th Computers and Information in Engineering Conference. doi: <https://doi.org/10.1115/detc2017-67082>
- Tanik, U. J., Begley, A. (2013). An Adaptive Cyber-Physical System Framework for Cyber-Physical Systems Design Automation. *Applied Cyber-Physical Systems*, 125–140. doi: https://doi.org/10.1007/978-1-4614-7336-7_11
- Yin, D., Ming, X., Zhang, X. (2020). Understanding Data-Driven Cyber-Physical-Social System (D-CPSS) Using a 7C Framework in Social Manufacturing Context. *Sensors*, 20 (18), 5319. doi: <https://doi.org/10.3390/s20185319>
- Hao, K. (2020). OpenAI is giving Microsoft exclusive access to its GPT-3 language model. *MIT Technology Review*. Available at: <https://www.technologyreview.com/2020/09/23/1008729/openai-is-giving-microsoft-exclusive-access-to-its-gpt-3-language-model/>
- Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., Sedova, K. (2023). Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations. Available at: <https://cdn.openai.com/papers/forecasting-misuse.pdf>
- Tabassi, E. (2023). Artificial Intelligence Risk Management Framework. NIST AI 100-1. NIST. doi: <https://doi.org/10.6028/nist.ai.100-1>
- Wang, Z., Sun, L., Zhu, H. (2020). Defining Social Engineering in Cybersecurity. *IEEE Access*, 8, 85094–85115. doi: <https://doi.org/10.1109/access.2020.2992807>
- NIST Special Publication 1270. Proposes a framework for identifying and managing bias in artificial intelligence.
- Foster, D. (2023). *Generative Deep Learning*. O'Reilly Media, Inc.

19. Yevseiev, S., Milevskiy, S., Bortnik, L., Alexey, V., Bondarenko, K., Pohasii, S. (2022). Socio-Cyber-Physical Systems Security Concept. 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). doi: <https://doi.org/10.1109/hora55278.2022.9799957>
20. Wang, Z., Zhu, H., Sun, L. (2021). Social Engineering in Cybersecurity: Effect Mechanisms, Human Vulnerabilities and Attack Methods. *IEEE Access*, 9, 11895–11910. doi: <https://doi.org/10.1109/access.2021.3051633>
21. Prete, E. D., Pera, F., Faramondi, L., Fioravanti, C., Guarino, S., Oliva, G., Setola, R. (2020). Anomaly and Attack Detection in Supervisory Control Networks for Cyber-Physical Systems. Proceedings of the 30th European Safety and Reliability Conference and 15th Probabilistic Safety Assessment and Management Conference. doi: https://doi.org/10.3850/978-981-14-8593-0_4315-cd
22. Toub Blanc, T., Guillet, S., de Lamotte, F., Berruet, P., Lapotre, V. (2017). Using a Virtual Plant to Support the Development of Intelligent Gateway for Sensors/Actuators Security. *IFAC-PapersOnLine*, 50 (1), 5837–5842. doi: <https://doi.org/10.1016/j.ifacol.2017.08.541>
23. Calefato, F., Lanubile, F., Novielli, N. (2017). A Preliminary Analysis on the Effects of Propensity to Trust in Distributed Software Development. 2017 IEEE 12th International Conference on Global Software Engineering (ICGSE). doi: <https://doi.org/10.1109/icgse.2017.1>
24. Lombardi, M., Vannuccini, S. (2022). Understanding emerging patterns and dynamics through the lenses of the cyber-physical universe. *Patterns*, 3 (11), 100601. doi: <https://doi.org/10.1016/j.patter.2022.100601>
25. Roy, T., Tariq, A., Dey, S. (2022). A Socio-Technical Approach for Resilient Connected Transportation Systems in Smart Cities. *IEEE Transactions on Intelligent Transportation Systems*, 23 (6), 5019–5028. doi: <https://doi.org/10.1109/tits.2020.3045854>
26. Hamzaoui, M. A., Julien, N. (2022). Social Cyber-Physical Systems and Digital Twins Networks: A perspective about the future digital twin ecosystems. *IFAC-PapersOnLine*, 55 (8), 31–36. doi: <https://doi.org/10.1016/j.ifacol.2022.08.006>
27. Li, X., Ye, P., Li, J., Liu, Z., Cao, L., Wang, F.-Y. (2022). From Features Engineering to Scenarios Engineering for Trustworthy AI: I&I, C&C, and V&V. *IEEE Intelligent Systems*, 37 (4), 18–26. doi: <https://doi.org/10.1109/mis.2022.3197950>
28. Lezoche, M., Panetto, H. (2018). Cyber-Physical Systems, a new formal paradigm to model redundancy and resiliency. *Enterprise Information Systems*, 14 (8), 1150–1171. doi: <https://doi.org/10.1080/17517575.2018.1536807>
29. Sowe, S. K., Simmon, E., Zettsu, K., de Vaulx, F., Bojanova, I. (2016). Cyber-Physical-Human Systems: Putting People in the Loop. *IT Professional*, 18 (1), 10–13. doi: <https://doi.org/10.1109/mitp.2016.14>
30. Smirnov, A., Shilov, N., Gusikhin, O. (2017). Cyber-physical-human system for connected car-based e-tourism: Approach and case study scenario. 2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA). doi: <https://doi.org/10.1109/cogsima.2017.7929591>
31. Kumar, S. A. P., Bhargava, B., Macedo, R., Mani, G. (2017). Securing IoT-Based Cyber-Physical Human Systems against Collaborative Attacks. 2017 IEEE International Congress on Internet of Things (ICIOT). doi: <https://doi.org/10.1109/ieeee.iciot.2017.11>
32. Zhu, Y., Tan, Y., Li, R., Luo, X. (2015). Cyber-Physical-Social-Thinking Modeling and Computing for Geological Information Service System. 2015 International Conference on Identification, Information, and Knowledge in the Internet of Things (IIKI). doi: <https://doi.org/10.1109/iiki.2015.48>
33. Kannisto, J., Makitalo, N., Aaltonen, T., Mikkonen, T. (2016). Programming Model Perspective on Security and Privacy of Social Cyber-physical Systems. 2016 IEEE International Conference on Mobile Services (MS). doi: <https://doi.org/10.1109/mobserv.2016.23>
34. Xu, Q., Su, Z., Yu, S. (2018). Green Social CPS Based E-Healthcare Systems to Control the Spread of Infectious Diseases. 2018 IEEE International Conference on Communications (ICC). doi: <https://doi.org/10.1109/icc.2018.8422421>
35. Rose, S., Borchert, O., Mitchell, S., Connelly, S. (2020). Zero Trust Architecture. NIST. doi: <https://doi.org/10.6028/nist.sp.800-207>
36. Yevseiev, S., Tolkachov, M., Shetty, D., Khvostenko, V., Strelnikova, A., Milevskiy, S., Golovashych, S. (2023). The concept of building security of the network with elements of the semiotic approach. *ScienceRise*, 1, 24–34. doi: <https://doi.org/10.21303/2313-8416.2023.002828>
37. Gilchrist, A. (2016). *Industry 4.0*. Apress Berkeley, 250. doi: <https://doi.org/10.1007/978-1-4842-2047-4>
38. EFFRA. *Factories of the Future: Multi-annual Roadmap for the Contractual PPP under Horizon 2020*. European Commission. Available at: https://www.effra.eu/sites/default/files/factories_of_the_future_2020_roadmap.pdf
39. Monostori, L. (2014). Cyber-physical Production Systems: Roots, Expectations and R&D Challenges. *Procedia CIRP*, 17, 9–13. doi: <https://doi.org/10.1016/j.procir.2014.03.115>
40. Uhlemann, T. H.-J., Lehmann, C., Steinhilper, R. (2017). The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. *Procedia CIRP*, 61, 335–340. doi: <https://doi.org/10.1016/j.procir.2016.11.152>
41. Kang, H. S., Lee, J. Y., Choi, S., Kim, H., Park, J. H., Son, J. Y. et al. (2016). Smart manufacturing: Past research, present findings, and future directions. *International Journal of Precision Engineering and Manufacturing-Green Technology*, 3 (1), 111–128. doi: <https://doi.org/10.1007/s40684-016-0015-5>
42. Sniderman, B., Mahto, M., Cotteleer, M. J. (2016). *Industry 4.0 and Manufacturing Ecosystems – Exploring the World of Connected Enterprises*. Deloitte University Press. Available at: https://www2.deloitte.com/content/dam/insights/us/articles/manufacturing-ecosystems-exploring-world-connected-enterprises/DUP_2898_Industry4.0ManufacturingEcosystems.pdf
43. Hermann, M., Pentek, T., Otto, B. (2016). Design Principles for Industrie 4.0 Scenarios. 2016 49th Hawaii International Conference on System Sciences (HICSS). doi: <https://doi.org/10.1109/hicss.2016.488>

44. Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., Sedova, K. (2023). Forecasting potential misuses of language models for disinformation campaigns—and how to reduce risk. Stanford. Available at: <https://cyber.fsi.stanford.edu/io/news/forecasting-potential-misuses-language-models-disinformation-campaigns-and-how-reduce-risk>
45. Kato, T., Kudo, Y., Miyakoshi, J., Otsuka, J., Saigo, H., Karasawa, K. et al. (2020). Rational Choice Hypothesis as X-point of Utility Function and Norm Function. *Applied Economics and Finance*, 7 (4), 63. doi: <https://doi.org/10.11114/aef.v7i4.4890>
46. Heath, J. (2008). *Following the Rules: Practical Reasoning and Deontic Constraint*. Oxford University Press. doi: <https://doi.org/10.1093/acprof:oso/9780195370294.001.0001>
47. Yevseiev, S., Melenti, Y., Voitko, O., Hrebenuik, V., Korchenko, A., Mykus, S., Milov, O. et al. (2021). Development of a concept for building a critical infrastructure facilities security system. *Eastern-European Journal of Enterprise Technologies*, 3 (9 (111)), 63–83. doi: <https://doi.org/10.15587/1729-4061.2021.233533>
48. Androshchuk, A., Yevseiev, S., Melenchuk, V., Lemeshko, O., Lemeshko, V. (2020). Improvement of project risk assessment methods of implementation of automated information components of non-commercial organizational and technical systems. *EUREKA: Physics and Engineering*, 1, 48–55. doi: <https://doi.org/10.21303/2461-4262.2020.001131>
49. Yevseiev, S., Ponomarenko, V., Laptiev, O., Milov, O., Korol, O., Milevskiy, S. et. al.; Yevseiev, S., Ponomarenko, V., Laptiev, O., Milov, O. (Eds.) (2021). *Synergy of building cybersecurity systems*. Kharkiv: PC TECHNOLOGY CENTER, 188. doi: <https://doi.org/10.15587/978-617-7319-31-2>
50. Yevseiev, S., Hryshchuk, R., Molodetska, K., Nazarkevych, M., Hrytsky, V., Milov, O. et. al.; Yevseiev, S., Hryshchuk, R., Molodetska, K., Nazarkevych, M. (Eds.) (2022). *Modeling of security systems for critical infrastructure facilities*. Kharkiv: PC TECHNOLOGY CENTER, 196. doi: <https://doi.org/10.15587/978-617-7319-57-2>
51. Yevseiev, S., Khokhlov, Yu., Ostapov, S., Laptiev, O., Korol, O., Milevskiy, S. et. al.; Yevseiev, S., Khokhlov, Yu., Ostapov, S., Laptiev, O. (Eds.) (2023). *Models of socio-cyber-physical systems security*. Kharkiv: PC TECHNOLOGY CENTER, 184. doi: <https://doi.org/10.15587/978-617-7319-72-5>
52. Shmatko, O., Balakireva, S., Vlasov, A., Zagorodna, N., Korol, O., Milov, O. et al. (2020). Development of methodological foundations for designing a classifier of threats to cyberphysical systems. *Eastern-European Journal of Enterprise Technologies*, 3 (9 (105)), 6–19. doi: <https://doi.org/10.15587/1729-4061.2020.205702>
53. Angiulli, F., Fassetti, F., Serrao, C. (2023). Anomaly detection with correlation laws. *Data & Knowledge Engineering*, 145, 102181. doi: <https://doi.org/10.1016/j.datak.2023.102181>
54. Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences*, 99, 7280–7287. doi: <https://doi.org/10.1073/pnas.082080899>
55. Macal, C. M. (2016). Everything you need to know about agent-based modelling and simulation. *Journal of Simulation*, 10 (2), 144–156. doi: <https://doi.org/10.1057/jos.2016.7>
56. Walker, J. J. (2012). Cyber security concerns for emergency management. *Emergency Management*. doi: <https://doi.org/10.5772/34104>
57. Ali, N. S. (2016). A four-phase methodology for protecting web applications using an effective real-time technique. *International Journal of Internet Technology and Secured Transactions*, 6 (4), 303. doi: <https://doi.org/10.1504/ijitst.2016.10003854>
58. Park, K.-J., Zheng, R., Liu, X. (2012). Cyber-physical systems: Milestones and research challenges. *Computer Communications*, 36 (1), 1–7. doi: <https://doi.org/10.1016/j.comcom.2012.09.006>
59. Hansman, S., Hunt, R. (2005). A taxonomy of network and computer attacks. *Computers & Security*, 24 (1), 31–43. doi: <https://doi.org/10.1016/j.cose.2004.06.011>
60. Goel, S., Chen, V. (2005). Information security risk analysis – a matrix-based approach. *Proceedings of the Information Resource Management Association (IRMA) International Conference*. San Diego. Available at: <https://www.albany.edu/~GOEL/publications/goelchen2005.pdf>
61. Kjaerland, M. (2006). A taxonomy and comparison of computer security incidents from the commercial and government sectors. *Computers & Security*, 25 (7), 522–538. doi: <https://doi.org/10.1016/j.cose.2006.08.004>
62. Blackwell, C. (2010). A security ontology for incident analysis. *Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research*. doi: <https://doi.org/10.1145/1852666.1852717>
63. Ahmad, R., Yunus, Z. (2012). A dynamic cyber terrorism framework. *International Journal of Computer Science and Information Security*, 10 (2), 149–158.
64. Judin, O. K. (2015). *Derzhavni informaciyi resursi [State information resources]*. Metodologiya pobudovi klasifikatora zagroz. Kyiv: NAU, 214.
65. Sznajd-Weron, K., Sznajd, J. (2000). Opinion evolution in closed community. *International Journal of Modern Physics C*, 11 (06), 1157–1165. doi: <https://doi.org/10.1142/s0129183100000936>
66. Deffuant, G., Neau, D., Amblard, F., Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, 03, 87–98. doi: <https://doi.org/10.1142/s0219525900000078>
67. Hegselmann, R., Krause, U. (2002). Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5 (3). Available at: <https://www.jasss.org/5/3/2.html>
68. Weimer, C. W., Miller, J. O., Hill, R. R. (2016). Agent-based modeling: An introduction and primer. 2016 Winter Simulation Conference (WSC). doi: <https://doi.org/10.1109/wsc.2016.7822080>