Cyber Resilience and Fault Tolerance of Artificial Intelligence Systems: EU Standards, Guidelines, and Reports

Oleksandr Lemeshko[0000-0002-0609-6520], Maryna Yevdokymenko[0000-0002-7391-3068], Oleksandra Yeremenko[0000-0003-3721-8188], and Ievgeniia Kuzminykh[0000-0001-6917-4234]

Kharkiv National University of Radio Electronics, Ukraine
oleksandr.lemeshko.ua@ieee.org

Abstract. The problem of ensuring cyber resilience and fault tolerance of artificial intelligence systems is urgent. The paper proposes methods for ensuring cyber resilience and fault tolerance of an artificial intelligence system based on existing European standards, recommendations, and reports. Collectively, the use of these methods and recommendations will make it possible to ensure complex cyber resilience and fault tolerance of the artificial intelligence system, namely databases (knowledge bases), the functionality of the system itself as a whole. The considered methods are based on the aspects of ensuring cyber resilience and fault tolerance of data centers or clouds as platforms for the deployment and implementation of artificial intelligence systems. Using the proposed solutions will increase the trust of artificial intelligence systems and will allow them to be implemented more intensively in many industries.

Keywords: Cyber Resilience, Fault Tolerance, Cybersecurity, Artificial Intelligence, Database, Personal Data, Data Center, Cloud.

1 Introduction

Given increasingly widespread and implemented computer systems, information security occupies an important place in the modern world. Therefore, existing security technologies require constant revision and modernization. One of the most effective areas of cybersecurity development, which allows detailed detection of attacks and preventing them faster than specialists in this field, is artificial intelligence [1–3].

Today, there are several classes of solutions that successfully apply modern technologies, which are part of the field of artificial intelligence. These classes include User and Entity Behavior Analytics (UEBA), Next-Generation Firewall (NGFW). Also, modern information security services (ISSs) from the unauthorized access are ready to recognize objects through a webcam and record facts of violation of security policies in real-time, for example, to detect an illegitimate person, a smartphone photographing a screen, an IP camera [4–9], etc. These capabilities are especially important today when many organizations relocated their employees to work remotely, but do not want to lose control over them. Moreover, in almost all classes of ISSs,

machine learning is actively used, which allows us to take a serious step in the development of cybersecurity and increase the resulting security level of organizations. Also, along with the advent of new machine learning algorithms, their scope has expanded. For several years now, machine learning in the field of information security has been used not only to detect attacks but also to carry them out. However, despite many advantages of AI, the artificial intelligence system (AIS) itself is susceptible to attacks such as model theft, framework vulnerabilities, the substitution of data for training, logical vulnerabilities. It is also worth noting that data processing using artificial intelligence methods leads to the fact that the final decision depends not only on the decision-making algorithm but also on the data processed earlier and currently being processed. As a result, two completely new types of attacks on AISs arise data poisoning, which characterizes the manipulation of input data during training to change the subsequent decision-making process; and data evasion, which characterizes the selection of input data at the decision-making stage, leading to their misclassification. Also, the processing of large amounts of data in machine learning systems certainly jeopardizes, first of all, the data of the users themselves. Hence, at present, there already are attempts to combine systems of this class with such actively developing promising directions in cryptography as homomorphic encryption and confidential computing protocols. However, these mechanisms are only at the stage of development and have not yet been implemented. Based on this, it follows that despite the many advantages of using artificial intelligence, the vulnerabilities of the AIS are the data processing system and data storage, i.e. the knowledge base based on which the entire artificial intelligence system is trained. Following this, the following urgent tasks arise:

- Ensuring the security of the AIS performance.
- Ensuring the protection of the AIS data storage.
- Ensuring the cyber resilience and fault tolerance of the AIS throughout its life cycle.

This paper is devoted to the analysis of existing solutions to each of the above tasks, as well as the subsequent review of the development and application of standards and recommendations in this area in Ukraine and at the International level.

2 Analysis of Existing Methods for Protecting Data Centers and Clouds

As a rule, data centers (DCs) and cloud technologies are used as data storages for deploying an AI system (infrastructure) for modern solutions. The probability of an attack on the network of a cloud hosting provider or a DC is high. This is caused by the large volume of resources placed there. In the DC case, due to the nature of the information placed, its high price, and criticality, we cannot exclude the threat of a professionally prepared and performed attack aimed at obtaining or destroying information, as well as achieving control over the resource [10–12].

Mandatory protection methods for DCs are shown in Fig. 1.

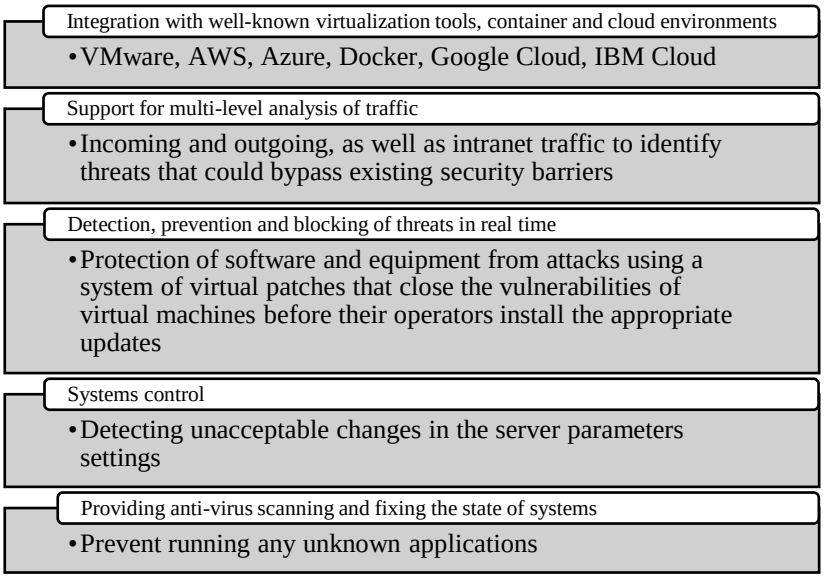

Fig. 1. Protection methods for DCs.

The following are the mandatory properties of the cloud protection system [13–16]:

- The ability to classify and manage cloud assets, which implies the classification, labeling, and processing of information.
- Security issues related to personnel; adding information security issues to job responsibilities, confidentiality agreements; educating and training of personnel in the field of information security.
- Physical protection of cloud storage, including perimeter protection and access control.
- Management of data transfer and operational activities, including operational procedures and responsibilities, isolation of development and production environments; control of information processing facilities by third parties and/or organizations; planning the performance and load of systems; protection against malicious software.
- Access control, including business requirements for control over logical access; user registration, control over user passwords; user identification and authentication; management of user privileges and access rights; protection of diagnostic ports during remote access; the principle of separation in networks; control of network connections; network routing management; security of using network services; control of access to the operating system; control of access to applications; restriction of continuous access to information.
- Monitoring of system access and use, including work with portable devices and work in the remote mode; measures to ensure information security when auditing systems.
- Development and maintenance of systems, including requirements for security and resiliency of systems, taking into account cyber resilience; information protection

measures related to the use of cryptography, encryption, digital signatures, the security of system files; software control; hidden channels of data leakage and Trojans, etc.

Particular attention should be paid to the International Recommendations, namely ANSI/TIA-942-B Telecommunications Infrastructure Standard for Data Centers [17], as well as the document "Cloud Computing Benefits, risks and recommendations for information security" by The European Network and Information Security Agency (ENISA) [18].

3 Methods of Data Storage Security of AIS

Usually, when discussing the security of databases, the risk of compromising and losing confidential information unwittingly comes to the fore. Modern conditions make us consciously approach security issues, obliging us to use more and more advanced methods of protecting the database [19, 20].

Basic database protection is setting up firewalls in front of the DBMS to block any access attempts from dubious sources, setting up and maintaining up to date password policy and role-based access model followed by auditing user actions. Today, there is a more effective approach—the use of specialized information security systems in the field of database protection—solutions of the Database Activity Monitoring (DAM) and Database Firewall (DBF) classes.

At the same time, DAM is a solution for independent monitoring of user actions in a DBMS. Moreover, independence denotes the absence of the need to reconfigure and tune the DBMS themselves. Systems of this class can be deployed passively, working with a copy of the traffic and not having any effect on business processes, the part of which the databases are.

DBF is a related solution, which also can "proactively" protect information. This is achieved by blocking unwanted requests. To solve this problem, it is no longer enough to work with a copy of the traffic, and it is necessary to install the protection system components "in the gap." In other words, database security mechanisms can be implemented in various ways: from designing a database with built-in security mechanisms to integrating the database with third-party products. The main direction in the development of methods for ensuring database security is the analysis of existing threats and risks. Thus, the existing international standards NIST, ISO/IEC, and COBIT constantly carry out such an analysis and put forward ever higher requirements for methods of ensuring security [21–29].

4 Methods for Ensuring Cyber Resilience and Fault Tolerance of AIS and Its Databases

Based on the above requirements for the protection of data centers and clouds, as well as the basic protection of databases and knowledge bases of the AIS, we can conclude that despite the use of various architectures, systems, virtualization tools, operating systems and software, the functionality of the AIS, the given means for data storing and

processing need to ensure their continued functioning and provision of services, which is determined by their fault tolerance and cyber resilience [21–24].

Thus, there arises a task of ensuring integrated security, including the AIS, its databases, and knowledge bases by ensuring the cyber resilience of the AIS and the fault tolerance of the data storage. Thus, we can conclude the direct dependence of the AIS functioning on its security and fault tolerance, which is shown in Fig. 2.

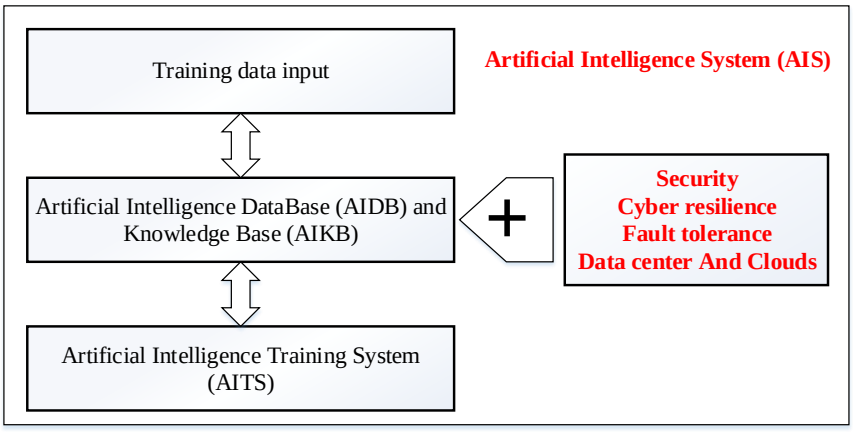

Fig. 2. Ensuring AIS cyber resilience and fault tolerance.

In other words, the DB and DBMS must be a cyber-resilient and fault-tolerant system that must maintain its operability when at least one node fails.

In this regard, the following requirements are put forward for a professional data storage system, shown in Fig. 3.

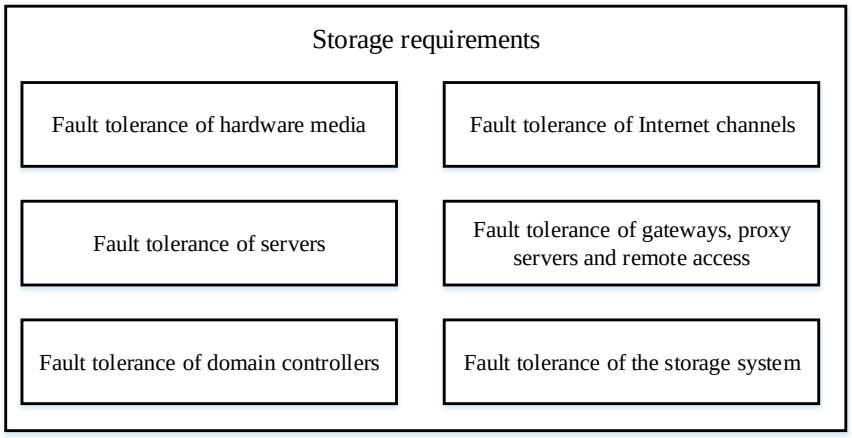

Fig. 3. The storage requirements.

At the same time, the fault tolerance of Internet channels means a recommendation to connect two stable Internet channels from two independent providers. To ensure fault

tolerance of gateways, proxy servers, and remote access, it is recommended to create a DBMS on professional hardware solutions with the ability to combine devices with synchronization of their configuration files into a failover cluster. Such an architecture will be able to ensure the stability of the Internet connection and uninterrupted access to internal and external key services. Fault tolerance of hardware media is provided by a single cluster of virtual machines with "live migration" technology through VMware, Hyper-V, etc.

Server fault tolerance refers to ensuring the continuous operation of all types of servers, such as terminal servers, DBMS servers, application servers, file, mail, and document servers, as well as Web servers. This task is solved through reservation and duplication.

Fault tolerance of domain controllers is achieved through the standard mechanism of the primary and backup domain controller roles.

Thus, the implementation of the recommendations shown in Fig. 3 will provide online reservations for critical services. In other words, if any problems with the main resources will appear, the entire system instantly switches to backup resources, almost imperceptibly for the user. At the same time, clusters are self-sufficient systems. In the event of emergencies, the cluster is designed for automatic actions to eliminate them and maintain the efficiency of services. The structure of a fault-tolerant DBMS is built in such a way that all important data, for example, personal data of customers, is centrally located in a reliable data storage system. The rest of the servers are required only for operational information processing and interactive work with the system. This means that if one of the servers fails, it can be easily replaced without the risk of losing or damaging company information. Also, along with fault tolerance, the distribution of the operational load among all cluster members is provided.

5 International Recommendations for Cyber Resilience and Fault Tolerance of Artificial Intelligence Systems

Analyzing the requirements for the AIS, it is important to note that to ensure the security of these systems, should be guided by the complex recommendations, aimed at protecting the functionality of the AIS, data storage, their overall cyber resilience, and fault tolerance. Based on this, there is a classification of recommendations for the development, implementation, and effective functioning of AIS. This classification of the main International standards is presented in Table 1.

International information security standards constitute an extensive system, which includes both mandatory provisions and provisions-recommendations for ensuring information security. At the same time, the development of new standards is an ongoing process that responds to all new challenges and incidents of information security and is aimed at designing a universal and reliable model for protecting personal data and information, including in cyberspace.

Table 1. Classification of the international standards for security AIS [26–33].

Application area	Standards, guidelines, and reports
Data center security and protection	• ISO / IEC 27001:2005 • ISO / IEC 27001:2017 • Telecommunications Infrastructure Standard for Data Centers (TIA-942B) • NIST 800-53 • SSAE 18 Audit Standard & Certification
Cloud protection	• ISO-27001 / ISO-27002 • ISO/IEC 27017:2015 • ISO/IEC 27018:2017 • CSA. Cloud Controls Matrix • CSA. Security Guidance for Critical Areas of Focus in Cloud Computing • NIST. SP 800-146. Cloud Computing Synopsis and Recommendations • ENISA. Cloud Computing: Benefits, Risks, and Recommendations for Information Security • ISACA. IT Control Objectives for Cloud Computing: Controls and Assurance in the Cloud • LCCA. Legal Cloud Computing Association: (SECTION II) HIPAA Cloud storage security
Personal data protection	• General Data Protection Regulation (GDPR) • Payment Card Industry Data Security Standard (PCI DSS) PCI compliance checklist 3.2 • Health Insurance Portability and Accountability Act (HIPAA)
System security and protection	• System and Organisation Controls (SOC) Reporting • CIS AWS Foundations v1.2 • CIS Controls Top 20 • ACSC Essential Eight • ETSI GS NFV-SEC 001 “Security Problem Statement”
Artificial Intelligence Security	• The ETSI Industry Specification Group on Securing Artificial Intelligence (ISG SAI) • ENISA’s WP2020 Output O.1.1.3 on Building knowledge on Artificial Intelligence Security • ENISA EC White Paper on Artificial Intelligence • Policy recommendations for safe and secure use of artificial intelligence, automated decision-making, robotics, and connected devices in a modern consumer world; • The Information Technology Industry (ITI) AI Policy Principles • The Malicious Use of Artificial Intelligence

At the moment, according to the Cybersecurity Strategy of Ukraine (2016–2020), the main task of developing the cybersecurity system is to ensure the cyber resilience and cybersecurity of the national information infrastructure, in particular in the context of digital transformation. Technical and applied recommendations are provided by “achieving compatibility with the relevant standards of the European Union and NATO,” as well as taking into account “the best world practices and international standards on cybersecurity and cyber defense.” Existing such documents in Ukraine in the field of cybersecurity does not meet the requirements of today's cyber defense, and in the field of artificial intelligence are absent. Therefore, the main direction of

development in the field of cybersecurity is artificial intelligence is the study of recommendations for their further implementation, operation, monitoring, analysis both at the state level and in individual sectors and industries.

6 Conclusions

This paper examines the main methods of ensuring the security and protection of an artificial intelligence system, including data storage, system functionality, as well as the cyber resilience and fault tolerance of artificial intelligence systems in general. Analysis of the existing International recommendations in this area showed the need for the development of relevant regulatory documents in Ukraine, due to their discrepancy or absence. In this regard, the introduction of AI systems in Ukraine is slow, which entails a lag in many industries.

References

1. White, R., Banks, E.: Computer networking problems and solutions: An innovative approach to building resilient, modern networks (2017)
2. Palhares, R. M., Yuan, Y., Wang, Q.: Artificial Intelligence in Industrial Systems. IEEE Trans. Ind. Electron. **66**(12): 9636–9640 (2019). https://doi.org/10.1109/TIE.2019.2916709
3. Bresniker, K., et al.: Grand Challenge: Applying Artificial Intelligence and Machine Learning to Cybersecurity. Comput. **52**(12): 45–52 (2019). https://doi.org/10.1109/MC.2019.2942584
4. Smelyakov, K., et al.: Efficiency of Image Convolution. IEEE 8th International Conference on Advanced Optoelectronics and Lasers: 578–583 (2019). https://doi.org/10.1109/CAOL46282.2019.9019450
5. Smelyakov, K., et al.: Comparative efficiency analysis of gradational correction models of highly lighted image. IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology: 703–708 (2019). https://doi.org/10.1109/PICST47496.2019.9061356
6. Smelyakov, K., Chupryna, A., Hvozdiev, M., Sandrkin, D.: Gradational correction models efficiency analysis of low-light digital image. Open Conference of Electrical, Electronic and Information Sciences: 34–39 (2019). https://doi.org/10.1109/eStream.2019.8732174
7. Hu, Z., Buriachok, V., Bogachuk, I., Sokolov, V., Ageyev, D.: Development and operation analysis of spectrum monitoring subsystem 2.4–2.5 GHz range. Lect. Notes Data Eng. Commun. Technol. **48**: 675–709 (2020). https://doi.org/10.1007/978-3-030-43070-2_29
8. Lemeshko, O., Yevsieieva, O., Yevdokymenko, M.: Tensor flow-based model of quality of experience routing. 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering: 1005–1008 (2018). https://doi.org/1010.1109/TCSET.2018.8336364
9. Kuzminykh, I., Carlsson, A., Yevdokymenko, M., Sokolov, V.: Investigation of the IoT device lifetime with secure data transmission. Lect. Notes Comput. Sci. **11660**: 16–27 (2019). https://doi.org/10.1007/978-3-030-30859-9_2
10. Kant, K., Le, M., Jajodia, S.: Security considerations in data center configuration management. 4th Symposium on Configuration Analytics and Automation: 1–9 (2011). https://doi.org/10.1109/SafeConfig.2011.6111676

11. Alhenaki, L., Alwatban, A., Alamri, B., Alarifi, N.: A survey on the security of cloud computing. 2nd International Conference on Computer Applications and Information Security: 1–7 (2019). https://doi.org/10.1109/CAIS.2019.8769497
12. Tipper, D.: Resilient network design: challenges and future directions. Telecommun. Syst. **56**(1): 5–16 (2014). https://doi.org/10.1007/s11235-013-9815-x
13. Ganesh, A., Sandhya, M., Shankar, S.: A study on fault tolerance methods in cloud computing. IEEE International Advance Computing Conference: 844–849 (2014). https://doi.org/10.1109/IAdCC.2014.6779432
14. Devi, K., Paulraj, D.: Multi level fault tolerance in cloud environment. International Conference on Intelligent Computing and Control Systems: 824–828 (2017). https://doi.org/10.1109/ICCONS.2017.8250578
15. Amoon, M.: A framework for providing a hybrid fault tolerance in cloud computing. Science and Information Conference: 844–849 (2015). https://doi.org/10.1109/SAI.2015.7237242
16. Cloud Security Alliance: Security Guidance for Critical Areas of Focus in Cloud Computing (2009)
17. European Network and Information Security Agency: Benefits, risks and recommendations for information security (2012)
18. NIST SP 800-145: A NIST definition of cloud computing (2012)
19. Kumar, B., Hamed Said Al Hasani, M.: Database security—risks and control methods. First IEEE International Conference on Computer Communication and the Internet: 334–340 (2016). https://doi.org/10.1109/CCI.2016.7778937
20. Firdhous, M. F. M., Hussien, N. A.: Data security implementations in cloud computing: A critical review. 3rd International Conference on Information Technology Research: 1–5 (2018). https://doi.org/10.1109/ICITR.2018.8736153.
21. Lemeshko, O., Yeremenko, O., Yevdokymenko, M.: Tensor model of fault-tolerant QoS routing with support of bandwidth and delay protection. 13th International Scientific and Technical Conference Computer Sciences and Information Technologies: 135–138 (2018). https://doi.org/10.1109/stc-csit.2018.8526707
22. Lemeshko, O., et al.: Design of the fast reroute QoS protection scheme for bandwidth and probability of packet loss in software-defined WAN. 15th International Conference the Experience of Designing and Application of CAD Systems in Microelectronics: 72–76 (2019). https://doi.org/10.1109/CADSM.2019.8779321.
23. Lemeshko, O., et al.: Cyber resilience approach based on traffic engineering fast reroute with policing. 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications: 117–122 (2019). https://doi.org/10.1109/IDAACS.2019.8924294
24. Yevdokymenko, M., Shapovalova, A., Voloshchuk, O., Carlsson, A.: Proactive approach for security of the infocommunication network based on vulnerability assessment. International Scientific-Practical Conference on Problems of Infocommunications Science and Technology: 609–612 (2019). https://doi.org/10.1109/INFOCOMMST.2018.8632079
25. ISO/IEC 27005: Information technology—security techniques—information security risk management (2011)
26. ISO/IEC 27017: Information technology—security techniques—code of practice for information security controls based on ISO/IEC 27002 for cloud services (2017)
27. Newman, J. C.: Toward AI Security: Global Aspirations for a More Resilient Future. Berkeley Center for Long-Term Cybersecurity (2019)
28. National Security Commission on Artificial Intelligence: First Quarter Recommendations (2020)
29. ANSI/TIA-942-A: Telecommunications Infrastructure Standard for Data Centers (2015)

30. ISACA: IT Control Objectives for Cloud Computing: Controls and Assurance in the Cloud (2011)
31. LCCA: Legal Cloud Computing Association: (SECTION II) HIPAA Cloud storage security (2011)
32. CSA: Security Guidance for Critical Areas of Focus in Cloud Computing (2017)
33. NIST: SP 800-146. Cloud Computing Synopsis and Recommendations (2012)