On the construction of data protection algorithms in cyberphysical systems

Andrey Iskhakov*

Financial University under the Government of the Russian Federation, 49, Leningradskiy ave.
Moscow, 125993, Russia

Abstract. Technology is playing an increasingly important role in our lives, and at the same time, data protection is becoming a critical issue. The article considers a particularly important sector - ensuring data security in cyberphysical systems of adaptive type, which combine physical and informational components. This paper proposes an approach presented in the form of a sequence of actions to adjust (optimize) the protection mechanisms by analyzing the state of the complex of automated systems involved in the processing of restricted access information obtained during the operation of cyber-physical systems for compliance with the required level of information security. Each subsystem processing the target information is assigned a specific security class. Optional barriers can be installed between subsystems of the same class, which ensures blocking of uncontrolled access to restricted data; management of components of all protected data. The results of approbation, representing the evaluation of using the proposed algorithm to modernize the subsystem of CPS protection against malware introduction, indicate its high efficiency.

1 Introduction

Combination of cyberphysical systems (CPS) for realization of industrial tasks leads to the formation of a new class of complex systems - heterogeneous adaptive CPS, whose specificity lies in the presence of mechanisms of adaptation to dynamically changing environment. In contrast to traditional technical or information systems, the synthesis of control systems for cyber-physical systems requires the development of new fundamental approaches to their safety management. Moreover, often the results of cyber-physical systems and the data circulating in it are restricted information, accumulated in data centers, and require proper procedures to ensure their confidentiality according to organizational rules as well as legal requirements.

Information security of cyber-physical systems is a critical issue that requires constant attention and development.

This paper proposes a developed and formalized algorithm to protect the data generated as a result of adaptive CPS information processing and accumulation. The urgency of this research is determined by the gap between the growing complexity of new-type cyber-
physical systems and the insufficient efficiency of existing control mechanisms for heterogeneous CPS of adaptive type.

2 Current state of research

Cyberphysical systems are the result of high-level integration of computing, networking, and physical processes [1-3]. A central characteristic of them is the intensive interaction of hardware and software resources to solve computational and communication problems along with the management of physical components. The principle of CPS functioning is to continuously acquire data from the environment, process it and apply it to further optimize control processes.

CPS have a serious economic and social potential, since their polysensory representation along with the wide application of human-machine interfaces also causes the emergence of new classes of CPS models that provide an opportunity to provide a flexible representation for both physical and informational components of the observed objects. In [4] four classes of cyberphysical models are proposed, the set of which allows describing various systems in general terms, freeing the upper layers of the monitoring infrastructure from redundant data and information noise.

One of the main requirements to CPS besides functional efficiency is safety of interaction of its components taking into account complex influence on controlled objects. Therefore, the formation of a secure sensor-information-communication network is the most important stage of creation and implementation of any modern CPS.

A relevant direction of research in the highlighted problem area is the development of methods of adaptive authentication of individual CPS components in M2M protocols, as well as authentication and identification of CPS operators/administrators, including those that provide for intelligent analysis of data flows using machine learning methods. The protection paradigm applied today, based on static rules of single-factor and multi-factor verification algorithms, not only leads to strict limitations that hinder the operational management of CPS, but has a number of significant drawbacks related to vulnerabilities in the methods of identity adaptation. Therefore, this project is planned to develop new risk-based authentication methods for CPS that make it much more difficult to compromise service commands by increasing the level of trust in authentication procedures.

CPS security is related to a wide range of branches of scientific knowledge, each of which has its own specific features [5-9]. At the same time, safety and security are two key properties of CPS. In fact, they differ only in their underlying data, as security aims to ensure the protection of the system from accidental failures, and safety aims to protect the system from deliberate attacks. With the common goal of ensuring the secure operation of a CPS, they share much in common in their approaches to achieving this goal. The alignment of defensibility and security is the foundation for the secure operation of a CPS. The need for such alignment has been recognized by both the research community and industry, which is reflected in research papers and standards: IEC TS 62443-1-1:2009 and IEC 62859:2016+A1:2019.

At the same time, ensuring secure interaction of CPS components is a complex problem involving various aspects. The practical realization of CPS is related to the technologies of the Internet of Things and wireless sensor networks.

3 Proposed algorithm

Taking into account the above-mentioned features and using the system-conceptual approach to the organization of effective protection of data obtained during the operation of cyber-
physical systems, an important task is to preserve the principles of adaptability, functional independence and reasonable sufficiency of protection subsystems, while ensuring the minimization of granted rights, completeness of control and active response to threats of violation of integrity, confidentiality and availability.

Taking into account the previously highlighted features, nature and conditions of processing of restricted access information obtained during the operation of cyber-physical systems, the following algorithm was formed, which allows to ensure the protection of target information, adaptively based on the input conditions of the information processing system. Figure 1 shows the conceptual block diagram of the algorithm operation.

Input: Conditions of data acquisition, information about protection mechanisms.

Output: Optimized list of mechanisms and recommendations to protect the target object.

Fig. 1. Block diagram of the proposed algorithm.

The labels in Figure 1:

- DB – Databases;
- P – Data Providers;
- S – List of system subjects having access to protected data;
- F – Protection functionality \( R, P \);
- \( P_s \) – Level (probability) of information security;
- \( P_{s0} \) – Required level (probability) of information security;
- M – Complex of applied solutions for threat neutralization;
The proposed algorithm is offered as a basic algorithm for the formation of information security policy, which provides the required level of protection of restricted access information obtained during the operation of cyber-physical systems. The algorithm proposes a sequence of actions to adjust (optimize) the protection mechanisms by analyzing the state of the complex of automated systems involved in the processing of restricted access information obtained during the operation of cyber-physical systems for compliance with the required level of information security.

Action description:

Step 1. Input data analysis is performed, including operating databases with protected information, arrays of system subjects and data providers, and technical parameters of the infrastructure.

Step 2. The construction of a security graph with an arbitrary level of nesting, the vertices of which are subsystems, is performed and the classification of assets is performed on it. As a result, each subsystem processing target information is assigned a certain security class. Optional barriers can be set between subsystems of the same class; mandatory barriers must be set between subsystems of different classes, and on the part of the subsystem with a higher security level.

After classification, segmentation (isolation) between unrelated groups of assets is performed. Particular attention is paid to building a multi-layered echeloned defense and isolation from publicly accessible networks.

Step 3. Regulatory requirements are assessed and the threat model is updated.

Step 4. When building/modernizing protection mechanisms, protection subsystems in each of the structural components and some controlling link (core) should be provided. The functions of the core should be: to organize and ensure blocking of uncontrolled access to restricted data; to manage the components of all protected data; to organize and ensure checks of the correctness of the protection system operation; to organize and maintain the reference data arrays of the protection system; to ensure response to signals of unauthorized actions; to maintain the protocols of the protection system.

Step 5. Security assessment is carried out. It is recommended to carry out a comprehensive audit with the involvement of third-party independent expert organizations, licensees of regulators.

Step 6. In case of non-compliance of the realized level of protection in comparison with the required one, the transition to Step 2 is carried out.

Step 7. Audit procedures should be carried out on a regular basis in accordance with the adopted safety policy of the operating organization. According to the results of the operation completion, the data should be cleaned by means of guaranteed destruction facilities.

Thus, the proposed algorithm allows to provide an informal description of the system components (architecture) and informal setting of security policy rules for an automated system that processes restricted access information obtained during the operation of cyber-physical systems. The algorithm provides for the interpretation of conditions for the implementation of the security policy and the distribution of responsibilities for the implementation of these rules between the subjects of the system on the basis of complex categorization of data and multidimensional analysis of existing means and mechanisms for information protection that perform the required functions.

4 Results of approbation

For testing purposes, it was decided to evaluate the use of the proposed algorithm for modernizing the CPS protection subsystem against malware introduction [10].

---

$T$ – Life cycle term;  
$D$ – Procedure of guaranteed data destruction.
The initial defense subsystem at the cyberpolygon was implemented based on an audit conducted by CPS security experts and using machine learning techniques [11, 12].

The first dataset contains over 10,000 rows of data with dynamic API calls, which are per hash function eigenvectors and markups on identifying legitimate and malicious events. This dataset contains sequences of malware and legitimate API calls. Each API call sequence consists of the first 100 non-repeating API calls associated with the parent process. The second dataset contains over 1,000 API calls to the VPO with their various families, which include additional information about the type, launch, and execution process of the program in an isolated environment.

Figure 2 shows the performance metrics of the resulting system.

![ROC curves on classification approaches in CPS malware data protection system.](image)

**Fig. 2.** ROC curves on classification approaches in CPS malware data protection system.

After classifying in accordance with the proposed algorithm, segmentation between unrelated groups of assets was performed. In accordance with the principles of building multi-layered echeloned protection and isolation from public networks, three CPS technology segments were identified, allowing them to be placed in separate isolated VLANs, reducing the risk of malware propagation through publicly available management interfaces.
Also, according to the algorithm steps, the need for an additional sandbox solution \( M \) for the communication gateway to download the CPS firmware update file was determined. As a result, the following characteristics of the protection subsystem were obtained, which are presented in Table 1.

Table 1. Estimates of the effectiveness of the modernized protection subsystem.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of isolated segments (out of 12)</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>(</td>
<td>M</td>
<td>) - Number of neutralization tools</td>
</tr>
<tr>
<td>Accuracy metric, % (malware detection)</td>
<td>92.5</td>
<td>95.2</td>
</tr>
<tr>
<td>Completeness metric, % (malware detection)</td>
<td>89.2</td>
<td>92.5</td>
</tr>
<tr>
<td>F1 metric (malware detection)</td>
<td>90.8</td>
<td>93.2</td>
</tr>
</tbody>
</table>

The conducted experiment and the given results show the improvement of the quality of information protection in CPS on a private example of development of subsystems of protection against malicious software. The verification of the conditions of security policy implementation based on complex data categorization and analysis of the existing means and mechanisms of information protection allowed not only to identify unnecessary interconnections, but also to supplement the system with the necessary security tools and improve the overall accuracy of malware detection.

5 Conclusion

Adaptive cyberphysical systems are used in a wide variety of applications such as automotive, industrial manufacturing, healthcare and others. They can reduce maintenance costs, improve operational efficiency, and enhance overall system performance. However, along with the benefits, cyber-physical systems also carry certain data security threats.

An important challenge is to maintain the principles of adaptability, functional autonomy and reasonable sufficiency of protection subsystems, while minimizing the rights granted, full control and proactive response to threats to integrity, confidentiality and availability. For example, malware detection in CPS can be difficult due to the complexity of training a large dataset of all malicious events and the large number of indications of malicious activity that can significantly degrade quality metrics during model training. Current malware data is collected in special environments, while malicious activity data can only be collected in a secure environment.

The proposed approach to building defense algorithms has shown its effectiveness on the example of the anti-malware subsystem. Meanwhile, it is necessary to apply complex measures to protect CPS, including technical, organizational and training components.

The article is based on the results of research carried out at the expense of budgetary funds on the state assignment of the Financial University.

References

2. R. Clauberg, World civilizations 3-4(5), 1-7 (2020)
5. Y. Klots, L. Koretska, Herald of Khmelnytskyi National University 303, 106-110 (2021)
12. Wu x-W, Y. Wang, Y. Fang, P. Jia, Neural Computing and Applications 34(8), 8643-8656 (2022)