UDC 004.056.5:621.317

# Determination of Cyberattack Parameters on the Measurements System of Critical Infrastructure Facility

Oleksii Novikov [1], Mykola Ilin [2], Mykola Ovcharuk [3], Iryna Stopochkina[4] and Andrii Voitsekhovskyi[5]

[1,2 3 4,5] *National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine*

**Abstract**

The paper solves the problem of determination and researching the parameters of stealth attacks on the linear Kalman filter data measurement system that bypasses the standard fault diagnosis detector. The relevance of the research is determined not only by the importance of solving cyber security problems, but also by the active use of the Kalman filter in large industrial power supply networks to evaluate the indicators of system nodes, in industrial automation systems, and others. A cyber attack on the measuring system of the Kalman filter is under consideration, the purpose of which is to disrupt the normal functioning of the filter by distorting the measurement signal, which is a mandatory component of the filter. The filtering system is equipped with a fault detector, which detects the presence of an attack on the measurement signals. The condition of the attack is invisibility for the fault detector, that is, the attacker implements a class of stealth attacks on the integrity of the information that circulates and is processed in the system. The task of finding a distorted measurement signal was solved using the variational optimization method and the gradient method of the fastest descent. A computational experiment was conducted, the quantitative characteristics of the algorithm were obtained and analyzed. The proposed method and the corresponding algorithm for determining the parameters of stealth attacks on the measurement system of critical infrastructure objects can be used to solve the problems of testing cyber defense systems.

*Keywords:* Cyberattacks, critical infrastructure, stealthy attacks, Kalman filters, optimization methods

## Introduction

In the past decade, the intensity of cyberattacks on critical infrastructure objects has significantly increased. This rise can be explained by the growing motivation of malicious actors, as well as vulnerabilities in the automated control systems of technological processes (Industrial Control Systems, ICS), Supervisory Control and Data Acquisition (SCADA) systems, distributed control systems (DCS), and programmable logic controllers (PLC). In some cases, SCADA systems may perform the functions of automated control systems (ACS) for managed objects and their components.

Malicious actors are often motivated by political or ideological reasons, as well as vandalism or hacktivism. In cases of political motivation, cyberattacks may be used as a tool of state influence or terrorism to destabilize the society or economy of a particular country. Political groups may target critical infrastructure to demonstrate their power or achieve political goals. Economically motivated cybercriminals may seek financial gain through extortion (e.g., ransomware) or intellectual property theft. Some hackers attack critical infrastructure to showcase their technical skills or express protests against certain organizations or governments.

The vulnerabilities of ACS and SCADA systems include outdated infrastructure, insufficient security measures, lack of network segmentation, human error, Internet connectivity, and others. Many ACS and SCADA systems were developed decades ago when cybersecurity was not a priority. These systems may contain outdated software and hardware, making them vulnerable to modern cyber threats. Often, critical systems lack necessary security measures such as data encryption, multi-factor authentication, and regular security updates, creating opportunities for malicious actors. The absence of proper network segmentation allows attackers who

breach one part of the system to easily move to other critical components. Misconfigured systems, inadequate staff training, and neglecting security policies can significantly increase the risk of successful cyberattacks. Finally, the growing connection of critical systems to the Internet for remote management and monitoring substantially increases their vulnerability to cyberattacks.

Attacks on critical infrastructure can lead to power outages, disruption of water supply, and transport systems, which can have serious social and economic consequences. Attacks on healthcare systems, emergency services, or industrial facilities can endanger public health and safety.

The most critical consequences of attacks on ACS and SCADA subsystems are for lower-level process control systems using programmable logic controllers (PLCs). This class of systems directly ensures the controllability, stability, and resilience of technological parameters in critical technologies.

Common attacks on lower-level process control systems or attacks on PLCs include Denial of Service (DoS) attacks, replay attacks, false data injection attacks, zero dynamic attacks, covert attacks, stealthy attacks, and others [1-5]. The last three types of attacks are the most dangerous. The goal of such attacks is to introduce harmful changes into control systems that cannot be immediately detected by standard monitoring and diagnostic methods. The effectiveness of these attacks depends on the presence of modern attack detection systems in industrial control systems.

Currently, a wide range of attack detection methods and systems have been developed. These methods range from traditional fault detection in software to modern anomaly detection methods based on behavioral system characteristics, artificial intelligence, and data analysis. The classification of attack detection methods and systems in industrial automation can be divided into several categories, such as by method type, integration level with automation systems, and the types of attacks they can detect [6].

In [7–10], traditional fault diagnostic methods were analyzed, including Bayesian detection with binary hypotheses, weighted least squares methods, $\chi^2$ detectors based on Kalman filters, quasi-fault detection and isolation techniques, and others.

The methods for detecting attacks in industrial automation systems are being improved, so the task of developing new and improving existing cyberattacks as a tool for testing cybersecurity systems remains relevant. In particular, stealthy attacks, zero dynamic attacks, and covert attacks are promising areas for research as they bypass standard diagnostic detectors.

## 1. Main concepts

### 1.1. Objective of the study

The aim of this paper is to develop and investigate the parameters of stealthy attacks on the data measurement system of the Kalman filter, as a tool for testing cybersecurity systems, which bypasses the standard fault detection diagnostics.

### 1.2. Attack Model for Cyber-Physical Systems

Let us consider a model of a physical dynamic system:

$$\frac{dx(t)}{dt} = Ax(t) + w(t), \quad (1)$$

$$y(t) = Cx(t) + v(t) \quad (2)$$

where $x(t)$ is n-dimensional vector of the physical system state; $y(t)$ is l-dimensional vector of measuring sensors of the measuring system; $A$ and $C$ are known matrices of the coefficients of corresponding dimension; $w(t)$ is n-dimensional vector of random disturbance if the process; $v(t)$ is l-dimensional vector of random measurement errors. Perturbations $w(t)$ and $v(t)$ are Gaussian random processes (white noises), that are not correlated with each other and with $w_0$. The quantities of $x(t)$ and $y(t)$ with perturbations entered are random and are characterized by some probability distributions. The ratio (1) is a system model and (2) is a model of measuring system.

Consider a linear filter that allows you to calculate the optimal assessment of the state $\hat{x}(t)$ of the system (1) by measurements $y(t)$ of the measuring system sensors [11]:

$$\frac{d\hat{x}(t)}{dt} = A\hat{x}(t) - K(t)[C\hat{x}(t) - y(t)]; \quad (3)$$
$$K(t) = P(t)C^T R^{-1},\ \hat{x}(0) = \hat{x}_0;.$$

$$\frac{dP(t)}{dt} = P(t)A^T + AP(t) - P(t)C^T R^{-1}(t)CP(t) + Q, \quad P(0) = P_0, \quad (4)$$

where the matrix $P(t)$ is a decision of the nonlinear Riccati equation (4).

Determine the variable $x(t)$ as «physical» state of the system, and the assessment of $\hat{x}(t)$, as a «cybernetic» state of the cyber physical system [9], which consists of a physical system model and a linear filter (1) – (4). This linear filter can be used in industrial power supply networks to determine the basic indicators of systems of systems that are not equipped with appropriate measuring devices [12]. Filters are widely used as subsystems of industrial automation systems, in particular to solve common filtration and control problems [13].

Consider the cyberattack on the measuring system of a linear filter (3), (4). The purpose of the cyberattack is to violate the standard functioning of the linear filter by distorting the measurement signal used in the filter. We will analyze the case where the attacker has the ability to attack only the measurement channel from the sensor to the filter (Fig. 1). In the case of considered, the system is equipped with a malfunction detector, which detects the presence of attack in the system. The condition of the attack is the imperceptibility for the detector of the malfunction, that is, the attacking classes the class of invisible attacks (Stealthy Attacks). In this way, the attack on the integrity of information that circulates and processed in the system is realized.

For the attack, the attacker examines the information security system, learns the knowledge and architecture of the automatic control system, receive elements of software codes, knowledge of algorithms for malfunctioning, etc. The cyber-physical system (3) -(4) is equipped with a malfunction detector (Figure 1), which was built using the weighted least square method [7 -10]:

$$J = \int_{t_0}^{t_k} [C\hat{x}(t) - y(t)]^T Q^{-1} [C\hat{x}(t) - y(t)]dt, \quad (5)$$

where $Q$ is known weight coefficient.

Formulated above model attack on the system can be written as

$$a(t) = [\tilde{y}(t)] = h(S, y(t)), \quad (6)$$

where $a(t)$ is the vector of attack, $\tilde{y}(t)$ – indicators of the meter, which is provided by distorted data, $S$ is an indicator of knowledge of the attacking object of the attack.

The model of the meter that is attacked can be written in the next ratio

$$\tilde{y}(t) = y(t) + y_a(t), \quad (7)$$

where $y_a(t) \in Y_{\text{acc}}$ is the parameter of imperceptible attack (stealth attacks), harmful data added to the original measurements, $Y_{\text{acc}}$ is acceptable values set.

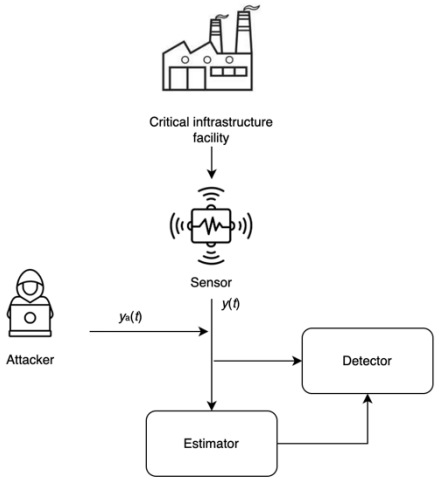


**Figure. 1:** Attack on a measuring system of a cyberphysical system

## 2. Main Results

### 2.1. Determining of harmful measurements $y_a(t)$ of an imperceptible attack (Stealthy attack), minimizing the malfunction detector $J(y_a)$.

Rewrite the ratio of the malfunction detector (5) in the form of

$$J(y_a) = \int_{t_0}^{t_k} [C\tilde{x}(t) - y(t)]^T Q^{-1} [C\tilde{x}(t) - y(t)]dt \to \min_{y_a \in Y_{\text{acc}}}, \quad (8)$$

where $\tilde{x}(t)$ is distorted assessment of the state under the influence of harmful data $y_a$, added to the original measurements $y(t)$, and the ratio of the filter (3), (4), which is under influence of cyberattacks (6), (7), as follows:

$$\frac{d\tilde{x}(t)}{dt} = A\tilde{x}(t) - K(t)[C\tilde{x}(t) - y(t) - y_a(t)], \quad (9)$$

$$\tilde{x}(0) = \tilde{x}_0. \quad (10)$$

Then formulate the problem of finding the minimum value of harmful data $y_a(t)$ of an imperceptible attack (Stealthy attack) on the meter (7) of a linear filter (9), provided (10), which bypass the malfunction detector (10).

Let us solve the problem of finding an extremum of functionality (8) with restrictions (9), (10) by Lagrange method. To take into account the restriction (9), we use Lagrange multiplier λ(t), which is vector of same dimension as $\tilde{x}(t)$, the restriction $y_a(t) \in Y_{\text{acc}}$ is taken into account at the final stage of the algorithm. We form Lagrange functional as

$$L(\tilde{x}(t), y_a, \lambda) = \int_{t_0}^{t_k} \left\{ [C\tilde{x}(t) - y(t)]^T Q^{-1} [C\tilde{x}(t) - y(t)] + \lambda^T(t) \left[ A\tilde{x}(t) - K(t)[C\tilde{x}(t) - y(t) - y_a(t)] - \frac{d\tilde{x}(t)}{dt} \right] \right\} dt. \quad (11)$$

Based on the variational principle, provided $\partial L/\partial \tilde{x} = 0$, we get a conjugate equation

$$\frac{d\lambda(t)}{dt} = -(A - KC)^T \lambda(t) - 2C^T Q^{-1} [C\tilde{x}(t) - y(t)] \quad (12)$$

$$\lambda(t_k) = 0. \quad (13)$$

Required optimality conditions concerning unknown parameter $y_a \in Y_{\text{acc}}$ take a form of

$$\delta L(y_a) \frac{dL}{dy_a} = \delta y_a = 0 \; \forall y_a \in Y_{\text{acc}}. \quad (14)$$

We supplement condition (14) for $y_a \notin Y_{\text{acc}}$

$$\frac{dL}{dy_a} = 0 \; \forall y_a \notin Y_{\text{acc}} \quad (15)$$

By varying the functional (11), it can be shown that in (14)

$$\frac{dL}{dy_a} = \int_{t_0}^{t_k} K^T \lambda(t). \quad (16)$$

We will determine harmful measurements $y_a(t)$ that minimize the indicator of the fault detector $J(y_a)$ using the gradient method of the fastest descent:

$$y_a^{i+1} = Pr\{y_a^i - \alpha \frac{dL^i}{dy_a}\}, \quad (17)$$

where $Pr\{\bullet\}$ is projection of the solution $y_a^{i+1}(t)$ onto the area $y_a(t) \in Y_{\text{acc}}$, $i$ is the number, and $\alpha$ is the known step of the gradient cycle, $y_a^0$ is also known.

The determination of unknown variable $y_a$ based on the gradient procedure (17) is completed if following condition is true:

$$| J^i - J^{i+1} | / J^i \quad (18)$$

where ε is known accuracy.

## 2.2. Algorithm for determining falsified measurements that minimize the fault detector index

Combining relations (8)-(10), (12)-(16) with (17), (18), we formulate an algorithm for determination $y_a$, that minimize the indicator of the fault detector $J(y_a)$:

1. For $i$ = 0, where $i$ is the iteration number of the gradient cycle, we give starting value $y_a^0$ and step $\alpha$.
2. For iteration $i$+1 based on relations (16) and (15) we calculate $dL^i/dy_a$, where Lagrange multiplier λ(t) and distorted state estimate $\tilde{x}(t)$, which is under the influence of malicious data $y_a(t)$, are determined by relations (12), (13) and (9), (10), respectively.
3. Using (17), we determine $y_a^{i+1}(t)$.
4. We calculate (8) and check (18). If it is fulfilled, then we complete the algorithm, otherwise, we go to point 2.

In order to analyze the efficiency of the developed method and algorithm for determining the parameters of a cyberattack on the measuring systems of a critical infrastructure object, let's consider an example.

## 2.3. Analysis of computational experiment results

Consider the technical system described by relations (1), (2). A linear filter that calculates the optimal estimate of the state $\hat{x}(t)$ of (1) based on measurements $y(t)$ of the sensors of the measuring system. It is necessary to solve the problem of finding the parameters of stealth attacks on the Kalman linear filter data measurement system that bypasses the standard fault diagnosis detector according to proposed algorithm. The initial values are given in Table 1.

**Table 1**
Initial values

| Parameter | Denotation | Value |
|---|---|---|
| Matrices dimension | $A, C, K, P, Q$ | $2 \times 2$ |
| Dimensions of column-vectors | $\hat{x}(t), \tilde{x}(t), \lambda(t), y(t), \tilde{y}(t), y_a(t)$ | $2 \times 1$ |
| Kalman filter coefficients matrix | $A$ | 0; 1; -6; -5 |
| Matrix of measurements | $C$ | 1; 0; 0; 1 |
| Weight matrices of the filter and fault detector | $R, Q$ | 1; 0; 0; 1 |
| Filter initial conditions | $\tilde{x}(0)$ | 0; 1 |
| Final conditions of conjugate equation | $\lambda(t_k)$ | 0; 0 |
| Period and discretization time step | $t_0, t_k, \Delta t$ | 0; 4; 0.01 |
| Step size, number of steps | $\alpha$, N | 0.1; 40 |
| Search parameter start value | $y_a^0$ | 1; 1 |

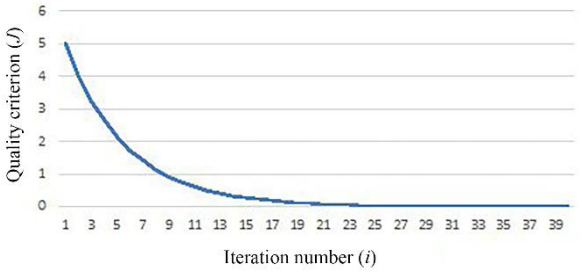


**Figure. 2:** Quality criterion behavior depending on the algorithm iterations

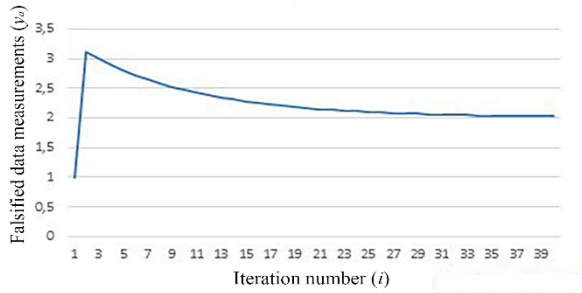


**Figure. 3:** Falsified data determination

The results (Figure 2) show the quality criterion $J(y_a)$, which decreases with each iteration, and approaches zero. The behavior of false measurement data $y_a$ during the gradient procedure, which are determined and entered by the attacker in order to minimize the detector criterion and make its actions invisible to the monitoring system, is presented in Figure 3.

## Conclusion

A cyber attack from the class of stealth attacks on the data measurement system of the linear Kalman filter, which bypasses the standard fault diagnosis detector, is considered. A method and a corresponding algorithm for searching for malicious measurements using the variational optimization method and the gradient method of the fastest descent are proposed. A computational experiment was conducted, the quantitative characteristics of the algorithm were obtained and analyzed. The analysis of the results confirmed the efficiency of the developed method and algorithm.

## References


[1] H.S. Sánchez, D. Rotondo, T. Escobet, V. Puig, J. Quevedo, “Bibliographical review on cyber attacks from a control oriented perspective”, Annual Reviews in Control, Vol. 48, 2019, pp.103-128. doi: 10.1016/j.arcontrol.2019.08.002.

[2] O. Novikov, M. Shreider, I. Stopochkina, M. Ilin, ”Cyber Attacks Simulation for Modern Energy Facilities”, CEUR Workshop Proceedings. Selected Papers of the XXIII International Scientific and Practical Conference "Information Technologies and Security" (ITS 2023), Vol. 3887, Pp.35-49. URL: https://ceur-ws.org/Vol-3887/.

[3] Y. Wang, M. C. Vuran, S. Goddard, “Cyber-physical systems in industrial process control”, ACM Sigbed Review, 5 (1), 2008, pp.1-2. doi: 10.1145/1366283.1366295.

[4] О. Новіков, І. Стьопочкіна, М. Ільїн, М. Овчарук. Визначення параметрів непомітних кібератак на системи керування об’єктів критичної інфраструктури // Наукові вісті КПІ, No 1, с. 69–75, 2025. doi: 10.20535/kpisn.2025.1.322905.

[5] L. Alekseichuk, O. Novikov, A. Rodionov, D. Yakobchuk, “Cyber Security Logical and Probabilistic Model of a Critical Infrastructure Facility in the Electric Energy Industry”, Theoretical And Applied Cybersecurity - Vol. 5, No. 1, 2023, pp. 61-66. doi:10.20535/tacs.2664-29132023.1.287365.

[6] M. Syfert, A. Ordys, J. Maciej Koscielny, P. Wnuk, J. Mozaryn and K. Kukielka, “Integrated Approach to Diagnostics of Failures and Cyber-Attacks in Industrial Control Systems”, MDPI Energies, Vol. 15 (17), 6212, 2022, pp.1-24. doi: 10.3390/en15176212.

[7] A. Sztyber-Betley, M. Syfert, J. Maciej Koscielny and Z. Gorecka, “Controller Cyber-Attack Detection and Isolation”, MDPI Sensors, Vol. 23 (5), 2778, 2023, pp.1-27. doi: 10.3390/s23052778.

[8] A. Cooper, A. Bretas and S. Meyn, “Anomaly Detection in Power System State Estimation: Review and New Directions”, MDPI Energies, Vol. 16, Issue 18, 6678, 2023, pp.1-15 URL: https://www.mdpi.com/1996-1073/16/18/6678.

[9] Y. Mo, B. Sinopoli, “Secure Control Against Replay Attacks”, 47th Annual Allerton Conference on Communication, Control, and Computing, Sept. 30 – Oct. 2, 2009, IEEE Xplore, pp. 911-918. doi: 10.1109/Allerton16076.2009.

[10] A. A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, Sh. Sastry, “Challenges for Securing Cyber Physical Systems”. URL: https://ptolemy.berkeley.edu/projects/chess/pubs/601.html.

[11] Kalman R.E., Bucy R.S., “New results in linear filtering and prediction theory”, Trans. ASME, Ser. D. J. Basic Eng., 83 (March 1961), pp. 95-108. URL: https://www.semanticscholar.org/reader/5c2f635fd11d2d001b7f9921007c6d3cf201eebf.

[12] N. F. I.Gulcharan, N. M. Nor, T. Ibrahim, H. Daud, “Power System State Estimation Bad Data Detection and Identification: A Review on Issues and Alternative Formulations”, Indonesian Journal of Electrical Engineering and Computer Science, Vol. 8, No. 1, 2017, pp. 122 – 128. URL: https://www.researchgate.net/publication/322030877_Power_System_State_Estimation_Bad_Data_Detection_and_Identification_A_Review_on_Issues_and_Alternative_Formulations.

[13] K.J. Astrom, Introduction to Stochastic Control Theory, Academic Press, New York, 1970. URL: https://www.amazon.com/Introduction-Stochastic-Control-Electrical-Engineering/dp/0486445313.

[14] F. Pasqualetti, F. Dörfler, F. Bullo, “Attack detection and identification in cyber-physical systems”, IEEE Transactions on Automatic Control 58 (2013), pp. 2715–2729. URL: https://arxiv.org/pdf/1202.6049.

[15] A. Teixeira, D. Perez, H. Sandberg, K. H. Johansson, “Attack models and scenarios for networked control systems”, Proceedings of the 1st international conference on High Confidence Networked Systems, ACM, Beijing, China, pp.55–64. URL: https://www.researchgate.net/publication/254008495_Attack_models_and_scenarios_for_networked_control_systems.