Cyber Attacks Simulation for Modern Energy Facilities

Oleksii Novikov[1], Mariia Shreider[1], Iryna Stopochkina[1] and Mykola Ilin[1]

[1]National Technical University of Ukraine "Igor Sikorsky KPI", Beresteiskyi Ave, 37, Kyiv, 03056, Ukraine

Abstract

This work focuses on enhancing the toolkit for simulating cyber attacks on energy facilities. The paper examines models of typical attacks on energy systems, specifically accounting for an attacker's ability to distort control system signals, manipulate control measurements, and alter measurement signals related to the state of the facility. A threats model for a critical infrastructure energy facility is proposed that refers to attack techniques. The approach considers integrity-breaking attacks expression as a function dependent on unknown parameters. Criteria are introduced to enable parametric identification of integrity compromising attack parameters, based on measurement data and constraints on process behavior. Stability conditions for a typical automatic gain control system under cyber attack are analyzed. An algorithm for identifying attack parameters is proposed. Computer simulations of facility processes under various attack types were conducted, appropriate software was developed, and conclusions were drawn regarding the impact of attacks on facility resilience.

Keywords

energy facilities, cybersecurity attacks, FDI attacks, models, resilience

1. Introduction

The AGC system is highly dependent on open communication infrastructure, such as the SCADA system, which increases its operational efficiency and responsiveness, but at the same time makes it more vulnerable to cyber attacks. Network technologies have many advantages, but all their defects – insufficient security, outdated protocols and software, and weak authentication mechanisms – create new opportunities for attackers. Therefore, the vulnerable points of the system are the inputs and outputs of the control center, that is, the communication channels through which data is transmitted [1].

Due to the need for rapid operation, the system does not employ complex algorithms for verifying and evaluating measurement data. Attackers can exploit this to manipulate data without sophisticated calculations. By knowing certain characteristics, an adversary can identify other unknown parameters of the system. In this paper, we demonstrate how this can be done, based on principles described in [2, 3].

Moreover, high coordination between interconnected control zones enhances productivity but also means that a sufficiently powerful cyberattack on one zone can adversely impact the entire power system.

Cyber attacks on energy supply facilities amplify and deepen the effects of physical attacks for maximum destructive impact. Understanding the limits of resilience to cyber influences is crucial in developing effective protective mechanisms and preventive measures. However, existing research [4-7] provides insufficient attention to the assessment of attack features or parameters.

The cyber vulnerabilities of AGC systems stem from data transfer mechanisms and protocol weaknesses. A taxonomy of these attacks was proposed in [8-10]. The paper [11] provides a detailed description of existing attack types on the advanced measurement infrastructure of smart grids, focusing on both IT (Information Technology) and OT (Operational Technology) systems. We

ITS-2023: Information Technologies and Security, November 30, 2023, Kyiv, Ukraine

o.novikov@kpi.ua (O. Novikov); marshr-ipt23@lll.kpi.ua (M.Shreider); i.stopochkina@kpi.ua (I.Stopochkina); m.ilin@kpi.ua (M.Ilin)

0000-0001-5988-3352 (O. Novikov); 0009-0006-8621-5521 (M. Shreider); 0000-0002-0346-0390 (I. Stopochkina); 0000-0002-1065-6500 (M. Ilin)

consider the entire AGC cyber-physical system, with particular emphasis on its OT features, and consider these attacks in terms of the necessary knowledge about cyber-physical system parameters.

The main classes of cyber threats for AGC system of energy facility are:

1. DoS (Denial of Service), DDoS (Distributed Denial of Service), and time delay attacks (targeting availability) [4, 5].
2. Replay attacks (targeting integrity) [6].
3. FDI (false data injection) and covert attacks (targeting integrity) [6,7].

In wartime, these cyber attacks are often combined with physical attacks on critical infrastructure facilities [12]. Developing algorithms for calculating attack parameters remains a crucial task for understanding the resilience limits of the facility and for investigating cyber incidents.

The findings of this work will contribute to more accurately fulfilling the guidelines of document [13] regarding the identification of adversary tactics, techniques, and procedures used to circumvent controls, along with other cybersecurity objectives.

2. Cyber attack models in AGC systems

Paper [1] examines a two-area power system and its dynamic model equations, demonstrating system behavior under abnormal conditions and analyzing the types of attacks that can disrupt the power system.

In paper [4], a dynamic model of a single-area load-frequency control (LFC) system is presented, focusing on the principles of sustainable operation. The study addresses time-delay attacks and DoS (Denial of Service) attacks, providing equations for the main system components under DoS attack conditions.

Paper [5] expands on DoS attacks by exploring data integrity attacks as well. It proposes a multi-area scheme with a control center, presenting detailed LFC equations and describing the main types of attacks.

Paper [6] discusses power grid control strategies, with particular emphasis on time-delay threats and replay attacks. The authors derive stability bounds for systems subjected to these attacks.

In paper [7], a different class of cyber attacks is explored: robust stealth covert attacks. The study includes a simulation example and uses a mathematical approach to calculate attack parameters for adversaries.

Paper [8] addresses cyber-physical reliability using game theory, incorporating probability factors into the calculations.

Paper [9] focuses on technical aspects of cyber attacks, reviewing examples, countermeasures, and a taxonomy of attack types. A section is dedicated to the use of machine learning algorithms for attack detection.

In paper [11], a detailed taxonomy of IT (Information Technology), OT (Operational Technology), and AMI (Advanced Metering Infrastructure) attacks is provided, along with an overview of papers that propose approaches to counter these attacks.

Paper [12] examines DoS and DDoS models, emphasizing that these attacks may have different impacts when combined with physical attacks by adversaries during wartime.

Simulation models of cascading effects in power grids under cyber attack are discussed in paper [14].

Paper [15] investigates various attack strategies, mathematical models, and methods for assessing system vulnerabilities.

The authors of paper [16] delve into the interconnected AGC systems and existing frequency deviations, advancing the study in this area.

Existing research reveals a gap in deterministic mathematical approaches, based on control theory methods, for not only identifying stability bounds but also uncovering unknown attack parameters. The current work aims to address this gap by developing relevant algorithm.

Paper [13] provides guidelines and compliance directions for reporting cyber incidents in critical infrastructure. This document offers guidance that could be reinforced by mathematical analyses and studies, particularly in the field of restoring attack parameters. The findings of the current study could provide the necessary numerical data for addressing these challenges.

3. Cyber threats to the AGC system

Let us examine the structural features of the AGC (Automatic Gain Control) system that make it susceptible to attacks. The AGC system operates within a communications infrastructure, facilitating data transmission between control centers and control zones. Sensor measurement data is sent to the control center, where an error signal is generated and then transmitted back to the control area. The local controller subsequently calculates the power control signal.

Real-time data collection can be achieved through remote terminal units (RTUs) or intelligent electronic devices (IEDs) positioned at critical locations (such as power stations and substations) within the control zone.

The SCADA (Supervisory Control and Data Acquisition) system collects and aggregates this data and relays it to the control center via communication channels using various protocols, such as DNP3 (Distributed Network Protocol), IEC 61850, and IEC 60870-5-104. Similarly, signals from the control center are transmitted back to the control zone. A general diagram of a single-area power zone under DDoS attack conditions is presented in [15], with specific points highlighted where other types of attacks (particularly FDI attacks) could be applied (Fig. 1).

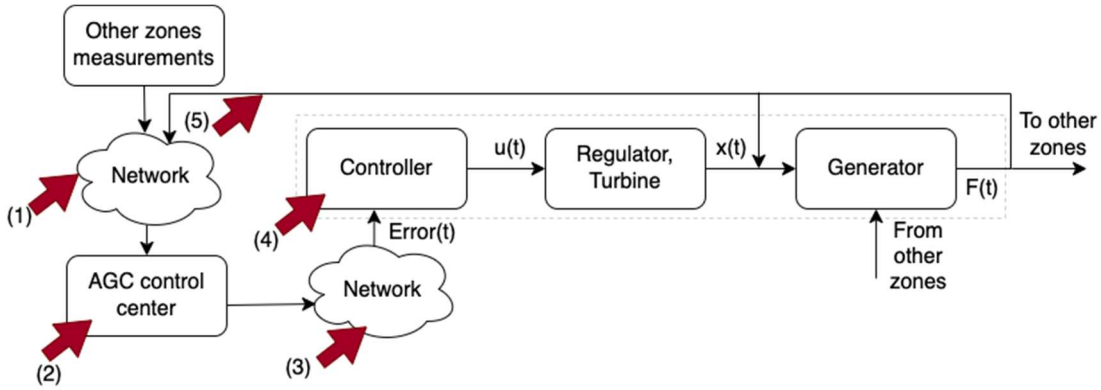

Figure 1: AGC System with External Communications. Arrows (1), (2), (3), (4), and (5) indicate points where a cyberattack can be applied. Potential targets include the communication network (1) and (3), internal communication lines (5), the AGC control center (2), and the programmable logic controller (4). An adversary could impact measurements (5), control signals $u(t)$, and the system state $x(t)$.

Let us compile a list of common attacks on the AGC system, linking specific attack types to technique classes from the MITRE ATT&CK® Matrix for ICS, as shown in Table 1. In Table 1, CIA refers to confidentiality, integrity, and availability, respectively.

Table 1
Energetic facility cyber attacks

Attack, technique ID	Affected (CIA)	Description	Attack pre-conditions	Information gathering	Target	Sub-system
DoS (T0814)	A	Data flood of the internal	Partial knowledge about	—	Channels for measurements and	IT, OT, AMI

Attack, technique ID	Affected (CIA)	Description	Attack pre-conditions	Information gathering	Target	Sub-system
		network and services	software, hardware versions, open interfaces		commands, system services	
DDoS (T0814)	A	DoS from multiple sources	Partial knowledge about software, hardware versions, open interfaces	—	Channels for measurements and commands, system services	IT, AMI, OT
FDI (T0836, T0868, T0830)	IA	False data injection	Normal mode features and anomaly ranges knowledge	System reactions and measurements	Measurement transmission channels	OT, IT, AMI
Replay (T0856, T0830)	CIA	Replaying real data	Partial knowledge about protocol timelines	Sensors and signals data	Measurement and control signals transmission channels	OT, AMI
Covert (T0836, T0868, T0830)	IA	Hidden attack	System full knowledge	Sensors and actuators data	Channels for measure-ments and commands	OT
Time Delays (T0814, T0830)	A	Introducing time delays	Partial knowledge about protocol timelines	—	Channels for measurements, control signals, and commands	OT, AMI
Physical attacks (T0879)	CIA	Destroying infrastructure, intercepting control under biometrical features, controlling the locks and other physical objects	Partial knowledge about system	Gathering all the data using social engineering, geolocation detection	Physical parts of critical infrastructure facility	OT, IT, AMI
Spoofing (T0856, T0830)	CI	Identity spoofing due to lack of authentication	Network protocols knowledge, access to transmitted data	—	IoT devices, PLCs, control center, network objects	IT, OT, AMI

Attack, technique ID	Affected (CIA)	Description	Attack pre-conditions	Information gathering	Target	Sub-system
Sniffing (T0842, T0887, T0801, T0830)	C	Access to data transfer nodes to sniff	Access to the network channels	Obtaining any usable data for further intrusion	Network channels	IT, AMI
TSA (time synchronization attack) (T0868)	IA	Synchronizing signal delay (replaying signals)	Knowledge about protocol peculiarities	Obtaining local time on target object	Channels of signals transmission	OT
Malware (TA0108, TA0104, TA0110, TA0111, TA0103, TA0102, TA0109, TA0100, TA0101, TA0107, TA0106, TA0105)	CIA	Taking control under controllers and other cyber-physical elements, or software of critical infrastructure facility. Can realize all types of possible techniques	Full knowledge of object architecture, and partial knowledge of system vulnerabili-ties	Keylogging and gathering all accessible data	Software and hardware of critical infrastructure facility	IT, OT, AMI

4. AGC mathematical models

In this section, we present generalized mathematical models in state space, building on previous works [5,6]. The primary vectors under consideration include malicious intrusion into the system state via control parameters and measurement parameters (see Fig. 1). We then focus on the FDI (False Data Injection) class of attacks and develop an algorithm to identify attack parameters under certain assumptions. Additionally, we discuss the adversary's potential extended knowledge of the system.

1.1. Initial undisturbed system model

We consider an initial undisturbed system with control, which is described by equations system in state space:

$$x'(t) = Ax(t) + kBu(t) + F, \quad (1)$$

where x is system state; u is control; F is source function (energy supply from/to neighboring zones); k is a parameter of control influence intensity.

We have to notice, that in the general description, state vector $x(t)$ can contain the components of frequency deviation Δf_i, regulator, turbine, and tie-line power deviations as it was proposed in [6]. But we consider the scalar values.

If the control depends on y measurements:

$$u(t) = -C_1 y(t),$$

where measurements depend on the state:

$$y(t) = C_2 x(t).$$

Then:

$$x'(t) = (A - kB_1)x(t) + F, \tag{2}$$

where $B_1 = BC_1C_2$.

$$B_1 = BC_1C_2.$$

For stability, the matrix $A - kB_1$ has to be negatively defined or at least, non-positively defined. This depends on eigenvalues λ_i of this matrix that can be defined from equation $\det(A - kB_1 - \lambda I) = 0$. Suppose that $A - kB_1$ is negatively defined for a sufficiently large k. Then the necessary condition that this property becomes invalid at some k_0, i. e., the largest eigenvalue changes its sign $\lambda(k_0) = 0$ is

$$\det(A - kB_1) = 0. \tag{3}$$

That can be used to find a critical value k_0.

1.2. Attack on system measurements and instability conditions determining

Let $\xi(t)$ be the distortion introduced to the measurements by an attacker. The measurements are given by

$$y(t) = C_2x(t) + \xi(t),$$

then

$$u(t) = -C_1y(t) = -C_1\big(C_2x(t) + \xi(t)\big) = -C_1C_2x(t) - C_1\xi(t).$$

Thus, equation (1) takes the form

$$x'(t) = \big(A - kB_1 - kB_2\xi(t)\big)x(t) + F, \tag{4}$$

where

$$B_2 = BC.$$

If $x(t)$ is known, identifying the attacker's intervention $\xi(t)$ becomes a standard fitting problem. Otherwise, it is necessary to determine $x(t)$ simultaneously with $\xi(t)$ when $y(t)$ is known.

The problem can be simplified if we know etalon values x^*, y^*, which allow us to eliminate F:

$$\begin{gathered} z(t) \equiv x(t) - x^*(t); \\ z'(t) = Dz(t) + D_1\xi(t)x(t); \\ \xi(t) = y(t) - C_2\big(x^*(t) + z(t)\big). \end{gathered} \tag{5}$$

From here:

$$z'(t) = Dz(t) + D_1\{y(t) - C_2[x^*(t) + z(t)]\}[x^*(t) + z(t)], \tag{6}$$

or

$$z'(t) = Dz(t) + f(t) + D_1\{y(t) - C_2[x^*(t) + z(t)]\}z(t), \tag{7}$$

where

$$f(t) = D_1\{y(t) - C_2[x^*(t)]\}x^*(t).$$

Assuming the effect of disturbances is small, successive approximations can be considered for equation (6). For the zero approximation, we set

$$\begin{gathered} \xi(t) = 0; \\ z(t) = 0. \end{gathered}$$

In the first approximation, we neglect the quadratic term by z :

$$z'(t) = Dz(t) + f(t) + D_1\{y(t) - C_2x^*(t)\}z(t),$$

$$z'(t) = \widetilde{D}z(t) + f(t), \tag{8}$$

where

$$\widetilde{D} = D + D_1\{y(t) - C_2x^*(t)\}.$$

Assuming

$$z(0) = 0,$$

we can find $z^1(t)$ by numerically solving the linear equation.

Given a set of measurements:

$$y^*(t) = C_2 x^*(t),$$

which characterizes normal process flow (solution of equation (2) or (4) when $\xi(t) \equiv 0$), we assume the adversary aims to maximize damage, causing $x^*(t)$ becomes unstable. The control problem for critical infrastructure systems is to prevent such scenarios through control measures and by comparing $y(t)$ and $y^*(t)$.

To detect intrusions caused by additional adversarial distortions, an additional criterion can be added to the measurement system to identify deviations from the normal process flow (e.g., electricity supply):

$$J(y) = \int_0^T (y(t) - y^*(t))^2 dt \to min.$$

Let $\mathcal{E}_{cr}$ be threshold such that

$$J(y) \geq \mathcal{E}_{cr},$$

signals abnormal system behavior. For discrete measurements:

$$J = \sum_{i=1}^{n} (y(t_i) - y^*(t_i))^2 \longrightarrow min. \tag{9}$$

Next, let us determine $\xi(t)$ that leads to system instability. Such a problem can arise in cyber incident investigation, especially when trying to uncover adversarial actions aimed at destabilizing the system. We can use

$$\det(D + D_1 \xi(t)) = 0 \tag{10}$$

where

$$D \equiv A - kB_1,$$
$$D_1 \equiv -kBC_1.$$

This allows us to define $\xi(t)$.

In equation (8), the addition of

$$y(t) - C_2 x^*(t)$$

is small because the distortions introduced by the adversary are minor and can be neglected in the first approximation.

Thus, from (8) we can write:

$$z^{(1)}(t) \approx e^{Dt} \int_0^t e^{-Dt' f(t')}\, dt';$$
$$\xi(t) \approx y(t) - C_2 \left[x^*(t) + e^{Dt} \int_0^t e^{-Dt'} f(t') dt' \right].$$

In the next approximation, we substitute $z^{(1)}(t)$ in the last term of (7).

As Table 1 shows, some attacks require knowledge of system functioning and parameters. Using the principles of parametric identification outlined above, and having access to measurement data, an adversary can infer unknown parameters (e.g., $A, k,$ or B from (1)). Thus, intercepting measurement information may enable more dangerous attacks, such as covert attacks.

1.3. Attack on system state and parameter identification

Let us consider a typical attack on the system state that involves false data injection (FDI) by manipulating system control parameters.

In FDI attacks, a scaling parameter ξ is used to alter control [5], allowing the adversary to influence system requlation.

Under attack, system (1) takes the form:

$$x'(t) = Ax(t) + kBu(t) + \xi u(t) + F, \quad (11)$$
$$u(t) = -C_1 y(t),$$
$$y(t) = C_2 x(t).$$

Rewriting the state equation, we have:

$$x'(t) = \big(A - k\tilde{A}\big)x(t) + \xi u(t) + F,$$

where

$$\tilde{A} \equiv BC_1C_2.$$

Let us rewrite the state equation in the form

$$x'(t) = Mx(t) + F,$$

where

$$M \equiv A - k\tilde{A} - \xi C_1 C_2.$$

From the condition

$$\det M = 0\,, \quad (12)$$

we can obtain the critical value of ξ that leads to system instability.

To illustrate the process of restoring attack parameters of the cyber incident, let us consider the generalized case of the system (11):

$$x'(t) = \mathrm{A}x(t) + \mathrm{B}(\xi)u(t) + F; \quad (13)$$
$$x(0) = x_0, \quad (14)$$

where x represents the system state, u is the control function, F is the source function, and ξ describes the intensity of adversary's intrusion. The dependency B on ξ is assumed to be known.

Suppose the adversary's goal is defined by the criterion under conditions (13), where $P(t), Q(t)$ are given functions, $x_m(t)$ represents the process state boundaries, and $u_m(t)$ is the desired control target of the adversary. We assume that $x_m(t)$ and $u_m(t)$ are known:

$$J(u) = \int_0^T [P(t)\big(x(t) - x_m(t)\big)^2 + Q(t)(u(t) - u_m(t))^2]dt \to min. \quad (15)$$

Setting $z(t) = x(t) - x_m(t);\ v(t) = u(t) - u_m(t)$, equation (1) can be reformulated as:

$$z'(t) = Az(t) + B(\xi)v(t), \quad (16)$$
$$z(0) = z_0, \quad (17)$$

where $z_0 = x_0 - x_m(0)$ and ideally

$$F(t) + Ax_m(t) + B(\xi)u_m(t) - x'_m(t) = 0. \quad (18)$$

Then, expression (15) is transformed to

$$J = \int_0^T [P(t)\big(z(t)\big)^2 + Q(t)(v(t))^2]dt. \quad (19)$$

The objective is to determine the feedback between $z(t)$ and $v(t)$ that the attacker introduces into the system to achieve the goal (15). This enables: 1) predicting the magnitude of adversary actions to train anomaly detection systems, and 2) recovering details of adversary actions from known incident characteristics $(x_m(t), u_m(t))$.

Introducing Lagrange multiplier, we have:

$$\delta J = \int_0^T \{2P(t)z(t)\delta z(t) + 2Q(t)v(t)\delta v(t) \quad (20)$$
$$+ \lambda(t)[\delta z'(t) - A\delta z(t) - B(\xi)\delta v(t)]\}dt =$$
$$= \int_0^T \{[2P(t)z(t) - \lambda'(t) - \lambda(t)A]\delta z(t) + [2Q(t)v(t) - \lambda(t)B(\xi)]\delta v(t)\}dt +$$
$$+\lambda(T)\delta z(t).$$

From the condition $\delta J = 0$, we select $\lambda(t)$ so that:

$$\lambda'(t) = 2P(t)z(t) - \lambda(t)A, \tag{21}$$
$$\lambda(T) = 0 \tag{22}$$

and

$$2Qv = \lambda B, \tag{23}$$
$$v = \frac{1}{2}Q^{-1}\lambda B. \tag{24}$$

Equation (16) then becomes:

$$z' = Az + \frac{1}{2}BQ^{-1}\lambda B, z(0) = z_0.$$

This problem is reduced to equations (21) and (25). However, this system is inconvenient because the conditions apply for $t = T$ and $t = 0$, respectively. To simplify it, substitute $\lambda = Lz$, where $L(T) = 0$:

$$z' = Az + \frac{1}{2}BQ^{-1}LBz, z(0) = z_0. \tag{25}$$

$$L'z = 2Pz - ALz - L\left[Az + \frac{1}{2}BQ^{-1}LBz\right].$$

$$L' = 2P - AL - L\left[A + \frac{1}{2}BQ^{-1}LB\right], L(T) = 0 \tag{26}$$

Thus, we can solve (26) for L numerically with the condition at $t = T$ and then, with the known L, solve equation (25) to find $v = \frac{1}{2}Q^{-1}LBz$. This solution minimizes the expression (19).

Given v and z, we can investigate the minimal values of $\mathcal{J}$ with respect to the attack parameter ξ using equations (15) and (16).

Applying the gradient method allows for a more efficient parameter identification process compared to a "brute force" calculation approach. The convergence ratio for the gradient procedure is estimated in [17]. Additionally, the conjugate gradient method [18] can be used as an alternative in step 6 of the algorithm.

The algorithm steps are as follows:

1. Set an initial arbitrary value ξ_0.
2. Find L from equation (26).
3. Determine z using equation (25).
4. Calculate v from equation (24), with $\lambda = Lz$.
5. Calculate $J(\xi_i)$ using equation (19), with $v = \frac{1}{2}Q^{-1}LB(\xi_i)z$.
6. Update ξ_{i+1}: $\xi_{i+1} = \xi_i + \tau\left(\frac{\partial J}{\partial \xi}\right)_i$.
7. If $\frac{|J^{i+1} - J^i|}{J^{i+1}} \leq \mathcal{E}$, proceed to step 8. Otherwise, return to step 2 for the next iteration.
8. The parameter value ξ_i will then satisfy (15) with precision $\mathcal{E}$.

A similar algorithm can also be used by a malicious actor to identify unknown parameters of the system. For this, only system state measurements are needed.

5. Computer simulation results

Using the presented models, we generated dynamics graphs of FDI attacks. For the simulations, we developed a Python software package.

1.4. Stability violation features

In Fig. 2, we illustrate the normal situation for the AGC. Here, we consider a one-component state *x*, representing frequency deviation *Δf*, and constant values of *ξ*, which could generally time-dependent. For a two-component state, see the example in Fig. 3.

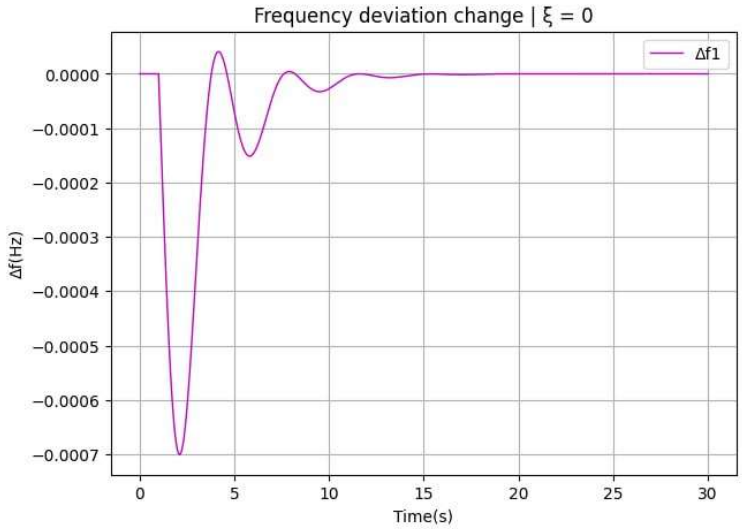

Figure 2: Undisturbed system, $\xi = 0$

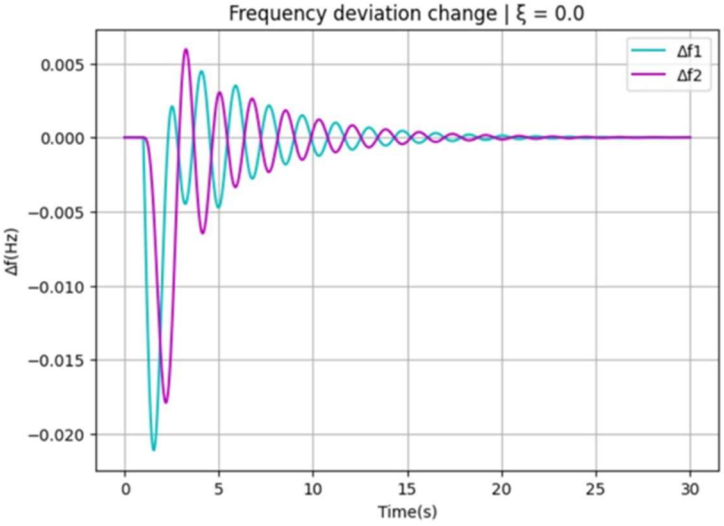

Figure 3: Undisturbed system, two-component state, $\xi = 0$

To identify the parameter ξ that meets a certain criterion $\mathcal{J}$ (see Fig.4), the proposed algorithm can be applied. In certain cases, some $\mathcal{J}$ samples may not contribute to the rapid convergence of the algorithm. However, in a significant number of cases, the proposed algorithm proves to be numerically efficient.

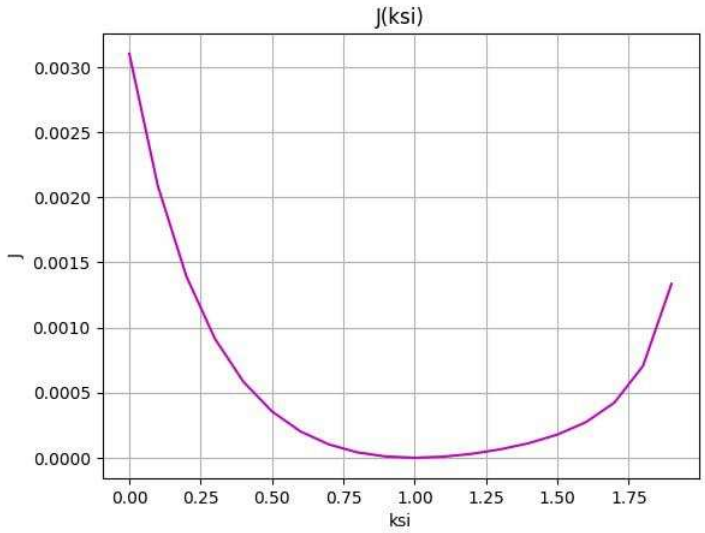

Figure 4: Criterion *J* sample with a minimum at $\xi = 1$

Fig. 5 shows that with small values of attack parameter, malicious influence may be subtle, making these attacks difficult for anomaly detection systems to detect. Such attacks typically target the software components of cyber-physical systems, aiming to insert false data into monitoring systems.

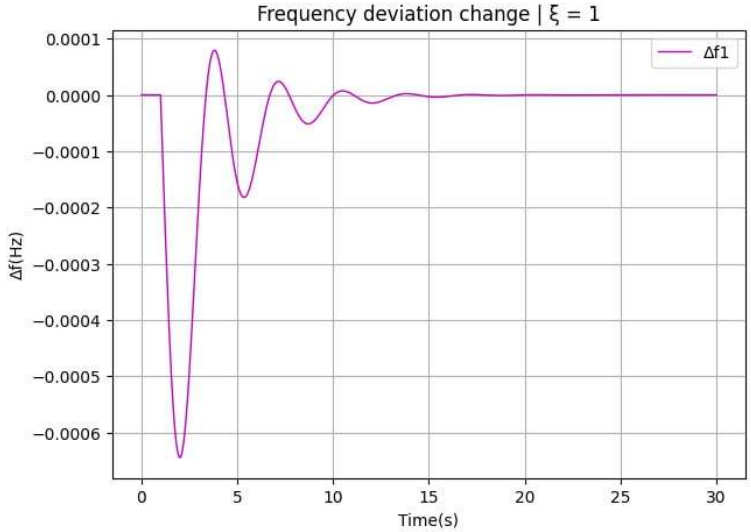

Figure 5: Malicious influence with $\xi = 1$

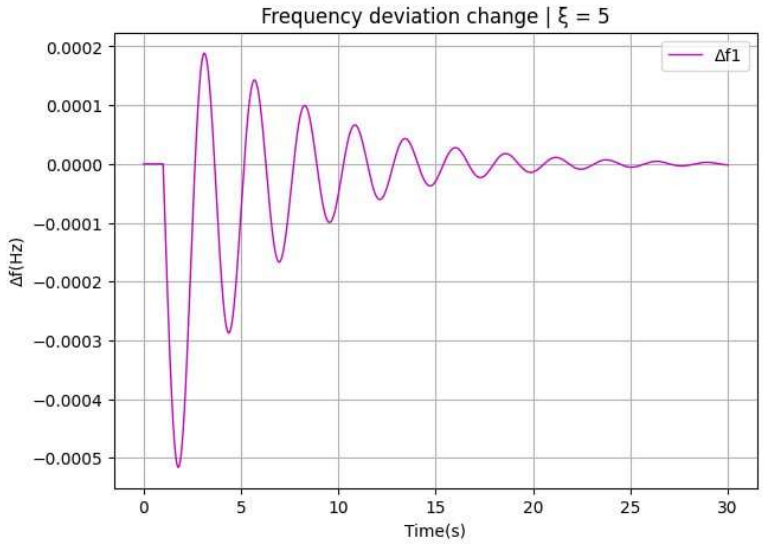

Figure 6: Malicious influence with $\xi = 5$

Attacks with larger values of scaling attack parameter can be detected effectively by monitoring systems due to noticeable changes in state pattern. For such attacks, cyber defenders should not only detect but also react quickly to mitigate potential damage. High values of the scaling parameter can pose risks to hardware components by threatening system stability. As shown in Figs. 7-9, with certain values of ξ, system state becomes unstable. The threshold value $\xi = 9.4$ (corresponding to the Fig. 8) can be calculated with necessary accuracy from (12). In the case of measurement intrusion, the stability boundary is determined by (10).

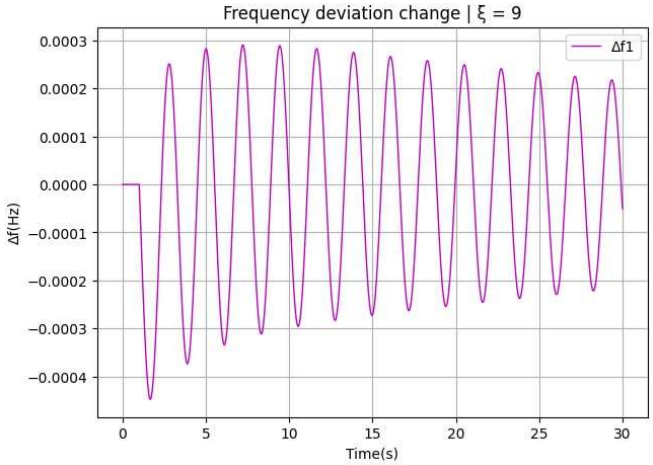

Figure 7: System state remains stable

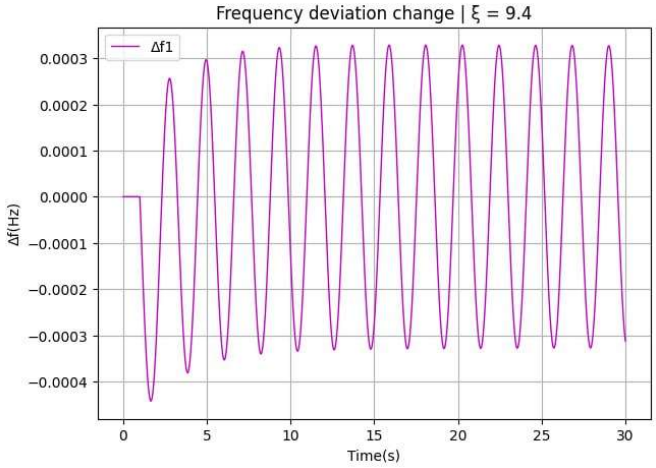

Figure 8: Attack with threshold value of the parameter

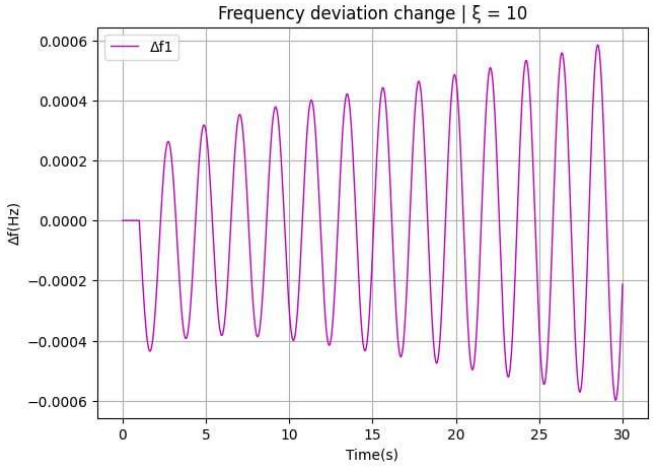

Figure 9: System state loses stability

1.5. Illustration of malicious activity at a specific time

Figures 10 and 11 illustrate scenarios where malicious influence "activates" at a specific time rather than initially. We observe a minor spike with low scaling parameter values (Fig. 10) and a clear change in the pattern with more significant influence intensity (Fig. 11).

Depending on the attacker's goal, small impacts can also lead to serious consequences as a result of tampering, affecting intrusion detection systems.

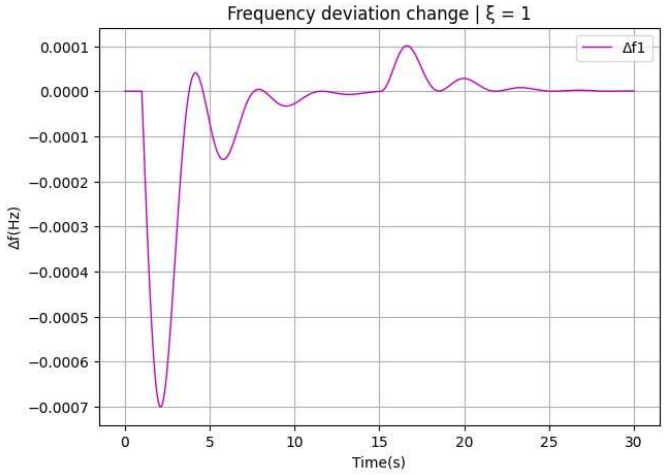

Figure 10: Frequency deviation pattern under attack parameter $\xi = 1$

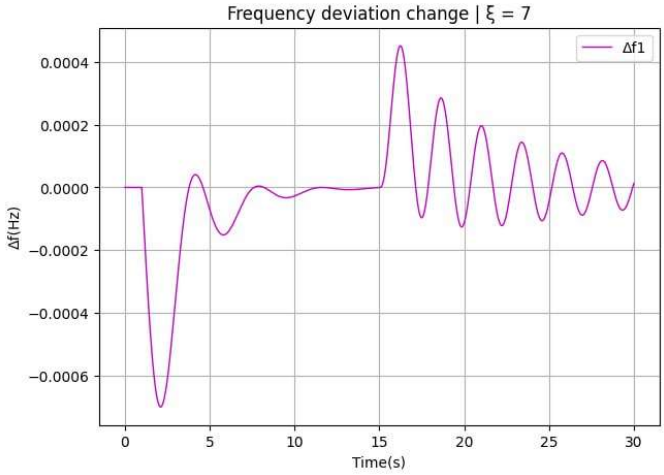

Figure 11: Frequency deviation pattern under attack parameter $\xi = 7$

6. Conclusions

Computer simulation results indicated that attacks with low values of the scaling parameter are not a threat to system stability but are challenging for anomaly detection systems to detect. Such attacks could be used by malicious actors to incrementally falsify historical data or poison machine-learning-based modules.

We derived the conditions for stable system operation based on the values of the attack parameter. Additionally, an algorithm was proposed for estimating the control intensity of FDI attacks, enabling the collection of quantitative data on malicious strategies to support system resilience.

An analysis of typical attack patterns in modern energy facilities showed that certain classes of attacks require full knowledge of the system. This information (e.g., system parameters) can be indirectly recovered using control theory principles, similar to the algorithm proposed in this paper for identifying unknown attack parameters. This highlights the risks posed by "sniffing" as a method for gathering measurement data, underscoring the need for preventive measures to prevent sniffing. Most data transfer protocols in AGC systems lack confidentiality by default, making them vulnerable.

The proposed approach and algorithm can be used for numerical incident investigations, providing solid foundations for response strategies. Future research could focus on studying combined attack types and enhancing detection methods.

References

[1] M. Vrakopoulou, P.M. Esfahani, K. Margellos, J. Lygeros, G. Andersson. Cyber-Attacks in the Automatic Generation Control. In: Khaitan, S., McCalley, J., Liu, C. (eds) Cyber Physical Systems Approach to Smart Electric Power Grid. Power Systems (2015). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-662-45928-7_11.

[2] P. S. Kundur, O. P. Malik, Power system stability and control, 2-nd ed., McGraw Hill Education, New York, 2022. URL: https://lccn.loc.gov/2021062400.

[3] P. L. Goethals, N, M. Scala, D. T. Bennett, Mathematics in Cyber Research, 2022. doi: 10.1201/9780429354649

[4] Y. Shen, M. Fei, D. Du, Cyber security study for power systems under denial of service attacks, 2017, SageJournals, Volume 41, Issue 6. doi: 10.1177/0142331217709528.

[5] A.M Mohan, N. Meskin, H. Mehrjerdi, A Comprehensive Review of the Cyber-Attacks and Cyber-Security on Load Frequency Control of Power Systems, 2020, Energies, 13(15), 3860. doi: 10.3390/en13153860.

[6] C. Moya, On Cyber-Attacks against Modern Power Grids, Ph.D. Thesis, The Ohio State University, 2020. URL: https://dl.acm.org/doi/10.5555/AAI28890224.

[7] X.Li, P.Zhang, H.Dong, Robust Stealthy Covert Attacks on Cyber-Physical Systems, 2022, IFAC-PapersOnLine, Volume 55, Issue 6, P. 520-525. doi: 10.1016/j.ifacol.2022.07.181.

[8] A. Ashok, A. Hahn, M. Govindarasu, Cyber-Physical Security of Wide-Area Monitoring, Protection and Control in a Smart Grid Environment, Journal of Advanced Research (2013), doi: http://dx.doi.org/10.1016/j.jare.2013.12.005.

[9] J. Ding, A. Qammar, Z. Zhang, A. Karim, H. Ning, CyberThreats to Smart Grids: Review, Taxonomy, Potential Solutions, and Future Directions. Energies, 2022, 15, 6799. doi: 10.3390/en15186799.

[10] P. S. Kundur, O. P. Malik, Power system stability and control, 2-nd ed., McGraw Hill Education, New York, 2022. URL: https://lccn.loc.gov/2021062400.

[11] Kim, Yoonjib & Hakak, Saqib & Ghorbani, Ali. (2022). Smart grid security: Attacks and defense techniques, IET Smart Grid, 6. doi: 10.1049/stg2.12090.

[12] M. Ovcharuk, M. Ilin, Models of Denial of Service Attacks on Cyber-Physical Systems, 2023, Theoretical and Applied Cybersecurity, Vol. 5, No2 (2023). doi: 10.20535/tacs.2664-29132023.2.289459.

[13] Cyber Incident Reporting For Critical Infrastructure Act of 2022, Subtitle D – Cyber Incident reporting. URL: https://www.cisa.gov/sites/default/files/2023-01/Cyber-Incident-Reporting-ForCriticalInfrastructure-Act-o-f2022_508.pdf .

[14] G. Vedmedenko, I. Stopochkina, O. Novikov, M. Ilin, Cascading effects simulation for cyber attacks on the power supply network // XXI International Scientific and Practical Conference "Information Technologies and Security" (ITS-2021), 09.12.2021. URL: http://ceur-ws.org/Vol-3241/.

[15] S. Saxena, S. Bhatia, R. Gupta, Cybersecurity analysis of load frequency control in power systems: A survey. Designs, 2021, 5(3), 52. doi: 10.3390/designs5030052.

[16] G. S. Kamboj, R. Dhiman, R. Choudhary, 2015, Automatic Generation Control of Two Areas in Interconnected Power System, International Journal of Engineering Research & Technology, (IJERT) NCETEMS – 2015, Vol.3, Issue 10. URL: https://www.ijert.org/automatic-generation-control-of-two-areas-in-interconnected-power-system.

[17] Boyd, S., & Vandenberghe, L. (2004). Convex Optimization. Cambridge: Cambridge University Press. URL: https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf .

[18] M.R. Hestenes, E. Stiefel, Methods of Conjugate Gradients for Solving Linear Systems. Journal of Research of the National Bureau of Standards, 1952. 49 (6): 409. doi:10.6028/jres.049.044.