Trustworthiness in Industrial IoT Systems Based on Artificial Intelligence

Zhihan Lv, *Senior Member, IEEE*, Yang Han, Amit Kumar Singh, *Senior Member, IEEE*, Gunasekaran Manogaran, and Haibin Lv

***Abstract*—The intelligent industrial environment developed with the support of the new generation network cyber-physical system (CPS) can realize the high concentration of information resources. In order to carry out the analysis and quantification for the reliability of CPS, an automatic online assessment method for the reliability of CPS is proposed in this article. It builds an evaluation framework based on the knowledge of machine learning, designs an online rank algorithm, and realizes the online analysis and assessment in real time. The preventive measures can be taken timely, and the system can operate normally and continuously. Its reliability has been greatly improved. Based on the credibility of the Internet and the Internet of Things, a typical CPS control model based on the spatiotemporal correlation detection model is analyzed to determine the comprehensive reliability model analysis strategy. Based on this, in this article, we propose a CPS trusted robust intelligent control strategy and a trusted intelligent prediction model. Through the simulation analysis, the influential factors of attack defense resources and the dynamic process of distributed cooperative control are obtained. CPS defenders in the distributed cooperative control mode can be guided and select the appropriate defense resource input according to the CPS attack and defense environment.**

***Index Terms*—Against and attack, cyber-physical system (CPS) and artificial intelligence (AI), industrial environments, Internet of things, trustworthiness model.**

Manuscript received March 29, 2020; accepted April 12, 2020. Date of publication May 14, 2020; date of current version November 18, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61902203 and in part by the Key Research and Development Plan, Major Scientific and Technological Innovation Projects of ShanDong Province under Grant 2019JZZY020101. Paper no. TII-20-1592. *(Corresponding author: Zhihan Lv.)*

Zhihan Lv is with the School of Data Science and Software Engineering, Qingdao University, Qingdao 266071, China (e-mail: lvzhihan@gmail.com).

Yang Han is with the College of Metallurgy and Energy, North China University of Science and Technology, Tangshan 063009, China (e-mail: hanyang@ncst.edu.cn).

Amit Kumar Singh is with the Department of Computer Science and Engineering, National Institute of Technology Patna, Bihar 800005, India (e-mail: amit_245singh@yahoo.com).

Gunasekaran Manogaran is with the University of California, Davis, CA 95616 USA, and also with the College of Information and Electrical Engineering, Asia University, Taichung City 41354, Taiwan (e-mail: gmanogaran@asia.edu.tw).

Haibin Lv is with the North China Sea Offshore Engineering Survey Institute, Ministry of Natural Resources North Sea Bureau, Beijing 266061, China (e-mail: lvhaibin@ncs.mnr.gov.cn).

Color versions of one or more of the figures in this article are available online at https://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2020.2994747

I. Introduction

THE PROPOSAL and development of cyber–physical systems (CPSs) provide a new idea and approach for the deep integration of a computer system, information system, and communication system. CPS is an intelligent system in which the computing units and physical objects are highly integrated and interacted in a network environment [1]–[3]. CPS, which includes the Internet of Things (IoT), information physical integration energy system or energy Internet, smart grid, intelligent transportation system, intelligent manufacturing system, intelligent logistics systems, etc., has become the core technology to support and lead a new round of industrial transformation [4], [5]. In artificial intelligence and machine learning, the data and application scenarios are very important. If the application scenario can be abstracted into a model, the data can be used, and the correct algorithm can be combined, then the information trusted application will greatly improve the credibility of the network information, which has become one of the key issues in the development and application of the Internet. Whether artificial intelligence can be successfully applied to the field of credibility, including several measurable factors, such as adaptability, interpretability and the feasibility of algorithm training Enforceability [6]–[10]. The traditional physical systems (such as industrial control systems, drive systems, and medical devices) are relatively difficult to access, penetrate, and enforce attacks by the attackers because of their relatively isolated operating environment and the use of dedicated channels for communication. Due to the need of informatization and intelligence, the traditional physical systems have gradually evolved into CPS, and the operating environment of the system has been opened and interconnected through closure and isolation [11]–[13]. While improving operational efficiency, it also provides new attack channels for the attackers, making CPS more vulnerable to internal or external attacks. For example, Stuxnet attacks invade nuclear facility control systems through U disk ferries, while WindShark directly accesses through physics. In a manned wind farm control system, once an attacker breaks through the CPS network boundary and enters the internal network, the attack success rate may be higher than the Internet attack, and the threat may be greater [14]–[17]. Most physical systems design fault diagnosis and safety emergency measures from an engineering safety perspective. For example, the Stuxnet attack tampers with the reported system measurement data, making the control center unable to detect system anomalies, black energy destroys the

system communication module and uses the denial of service attacks interfere with grid calls [18]–[20]. The service system makes the system's default emergency response and defense strategies not implemented effectively. The common features of these attacks are formulating the corresponding attack strategies for specific business processes and the trusted plans of physical systems, using a network attack technology to implement coordinated attacks on multiple targets, and bypassing physical trusted systems to disrupt the normal operation of the system, even destroying physical equipment [21]–[24]. In the future, based on artificial-intelligence-based autonomous learning and powerful database analysis capabilities, people can anticipate dangers in advance and truly kill threats in the cradle, thereby greatly improving the agility of network physical credibility. The credible control of CPS is shifting from physical credibility to physical and social factors integrated disaster prevention [25]–[28]. However, due to the characteristics of CPS, such as large-scale data processing, continuous online operation of the system, the operators can only conduct closed feedback, and so on, it is urgent to realize the real-time reliability evaluation of CPS. To promote the in-depth integration of informatization and industrialization is to profoundly grasp the characteristics of the era of global informationization and the process of accelerating the convergence of industrialization.

II. Research Status of Trustworthiness Model

The security status of the industrial IoT system is a comprehensive appearance under time accumulation. It is necessary to comprehensively consider the time-series evolution law of the key parameters of system operation. Therefore, the space-time analysis method is needed, and the differential equation model is an effective model to describe the spatiotemporal distribution of parameters. In order to eliminate the limitations of the differential equation model for only a few key parameters, this article considers the other redundant parameters of the system state to correct and compensate for the deficiency of the differential equation model [29], [30]. The accuracy of system security status detection and forecasting is measured by the length of equipment life extension and system operation stability. These two indicators are also the most critical indicators for the practical application of the industrial IoT.

A. Typical CPS Control Model

The ability to defend against information network attacks is an important measure of CPS performance. Trustworthiness events in recent years have shown that the means of attack faced by CPS are constantly being updated, and its diversity and concealment greatly increase the cost and cost of defense. Successful systems defense is usually based on a thorough understanding of the system architecture and attacks. Therefore, it is necessary to fully understand the characteristics of CPS attacks and existing anomaly detection mechanisms and develop the targeted defense strategies to enhance system trustworthiness [29], [30].

The coupling of the information system and the physical system is the difficulty in establishing the CPS-integrated trustworthiness model. It not only analyzes the topology structure and

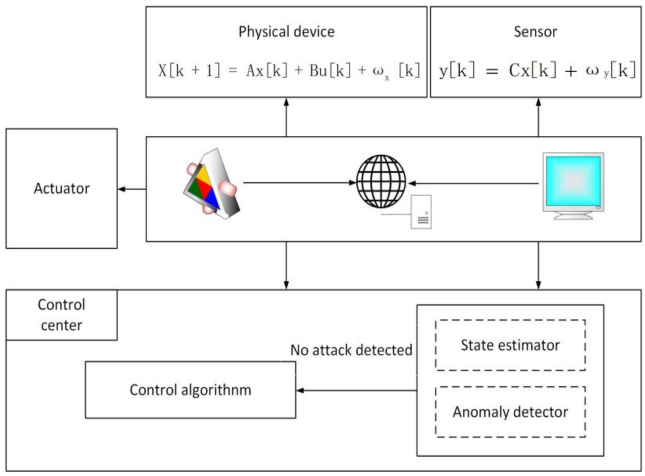


Fig. 1. CPS control model.

the trustworthiness monitoring mechanism of different systems but also considers the information exchange mode and trustworthiness vulnerability between the two types of systems. First, the existing CPS system model and the trustworthiness monitoring mechanism are summarized. Combined with the existing threat model, focusing on the concealment of control signals and measurement signals tampering attacks, a CPS-integrated trustworthiness model is proposed.

A typical CPS consists of a control center, physical equipment, actuators, and sensors, as shown in Fig. 1. When the system state changes within a certain range, CPS can be approximated as a linear system. Since linear system theory and methods are mature, the linear time-invariant (LTI) system of discrete time is taken as an example to illustrate the modeling ideas and methods. Specifically, the LTI model can be described as follows:

$$x[k+1] = Ax[k] + Bu[k] + \omega_y[k] \tag{1}$$

$$y[k] = Cx[k] + \omega_y[k]. \tag{2}$$

$x[k]$ and $y[k]$ represent the state of the system and the measured value of the sensor, $u[k]$ represents the control signal, and A, B, and C represent the system matrix, the control matrix, and the measurement matrix respectively. $\omega_x[k]$ and $\omega_y[k]$ both obey the zero-mean multivariate Gaussian distribution, and the covariance matrix is Q and R. Both the control signal $u[k]$ and the sensor measurement $y[k]$ are transmitted by the communication network.

The formula, as presented in Fig. 1, is a schematic diagram of the time-domain equation of physical parameters as a function of time. In practical applications, these formulas use Laplace transform to transform differential equations into algebraic equations, and the solution process and analysis process can be realized.

The measurement data of the industrial control system need to be preprocessed, mainly including a state estimator and an anomaly detector. The detection principle can be divided into two categories: one is to use the state of the system past time to predict the future state, and according to the prediction results, measurement results were used to detect the abnormal data.

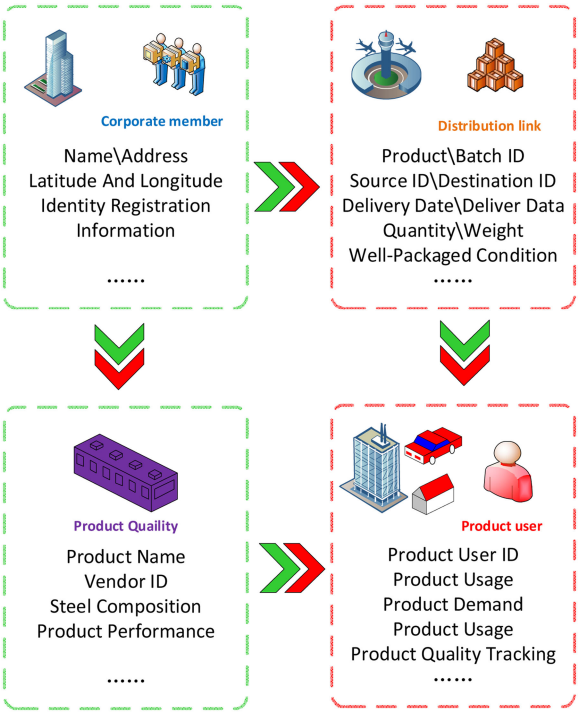


Fig. 2. Partial data structure relationship between the iron and steel product traceability system.

In this article, it was named as time-dependent detection. The other type was cross validation through the correlation between the different sensors of the system to achieve the purpose of detection. This article named it space-dependent detection.

B. Time-Dependent and Spatially Correlated Detection Models

The main members of the steel product traceability system include steel companies, distribution, product users, quality regulators, etc. The information of the steel enterprise includes the registration information of the enterprise, the raw material composition of the product, the production process data of the product, the product quality data, etc. The information of the distribution includes the batch, packaging, and the logistics date, etc. The information on the product user includes the required product category, use, quality traceability, etc. As a result, a part of the data relationship between the product manufacturing chain and the supply chain is shown in Fig. 2, and when a problem occurs at a node, the impact can be traced back or up along the data relationship network to improve the product supply and related dispute resolution efficiency.

At present, the graph computing system generally adopts the node centered programming model or interface. The user can customize a node update function update(), and the graph calculation system will complete the calculation task by calling the update function for each node. This function describes the operations that a node needs to perform in a round of iterations. A typical node update function update() follows the pseudocode, as shown in Fig. 3.

```
Update (vertex) begin
    x ← read values of in- and out-edges of vertex.
    vertex. Value ← f(x).
    for each edge of vertex do
    edge. Value ← g(vertex. value).
    end
end
```

Fig. 3. Typical node update function update() follows a logical algorithm.

C. Comprehensive Trustworthiness Model Analysis

This section mainly analyzes the CPS-integrated trustworthiness model in the attack scenario. In order to ensure the universality of the analysis, the attack scenario where the control signal and the measured value are simultaneously tampered with is considered. The symbol is used to indicate the normal state. The physical quantity underneath define the amount of attack the system suffers as follows (to facilitate the omission of the time index in the definition of this section).

1) Use x_a and y_a to represent the actual system status and the sensor measurements, respectively.
2) Use $\hat{x}_m, \hat{y}_m, \varepsilon_m,$ and u_m to represent the state estimation value of the control center, the predicted value of the sensor measurement, the anomaly detector residual, and the determined control signal, respectively.
3) Use u_a to indicate the actual control signal received by the actuator. This signal has been tampered with by the attacker. Specifically, $u_a = u_m + a_u$, where u_m is the control command issued by the control center, and $a_u \in R^l$ is the malicious injected in u_m control signal.
4) Use y_m to indicate the measured value of the sensor received by the control center. This signal has also been tampered with by the attacker. Specifically, $y_m = y_a + a_y$, where $a_y \in R^m$ represents the malicious signal injected into the actual measured value of the sensor y_a.

From the point of view of the controlled physical system, the control commands received by the physical system have been tampered with by the attacker, and the physical system makes an erroneous decision based on the instruction. At the same time, the measured value of the current state is also tampered with by the attacker.

At time k, the control commands, the system state equations, and the measurement results accepted can be expressed as follows:

$$u_a = u_m + a_u \tag{3}$$

$$x_a[k+1] = Ax_a[k] + Bu_a[k] + w_x[k] \tag{4}$$

$$y_a[k] = Cx_a[k] + w_y[k]. \tag{5}$$

From the control center, the measured value y_m received by the attacker has been changed to the data under normal conditions. The control center detects the abnormality of the

received measured value. After it is determined to be normal, it will send according to the control policy control signal u_m. At time k, the tamper-received sensor measurement will affect the system state estimation process, as shown in (2) and (3), and the anomaly detection, as shown in (4) and (5). The control algorithm can be expressed as

$$\hat{x}_m[k+1] = L_1(\hat{X}_m[k], U_m[k], Y_m[k+1]) \tag{6}$$

$$\hat{y}_m[k+1] = L_2(\hat{X}_m[k], U_m[k], Y_m[k]) \tag{7}$$

$$\varepsilon_m[k+1] = y_m[k+1] - \hat{y}_m[k+1] \tag{8}$$

$$u_m[k+1] = L_3(\hat{X}_m[k+1], U_m[k], Y_m[k+1], x_0). \tag{9}$$

The differential equation model with the same structure is adopted, and the model analysis method is consistent with the previous method. The threat model under a time-varying attack is taken as an example of the analysis. In addition, the threat model under the space-related data integrity attack is similar to the threat model under the time-related attack, which is not described here.

The core idea of the CPS-integrated trust model designed in this article is to abstract the possible situation of CPS attack from the perspective of the system administrator, abstract it into the system control flow, and adaptively modify the measured values and the control commands. The constructed credibility constraint space is a defensive measure that CPS can take during the attack. It provides a feasible theoretical way for researchers and managers to understand the comprehensive reliability of CPS.

III. Research Status of Safety Performance Analysis

This article proposes a new power CPS security risk assessment framework, as shown in Fig. 4, which aims to estimate the credibility of industrial IoT systems and provide guidance for the prevention of system security risks. The preferred subject model of attack prediction and defense strategy is the differential equation model. To compensate for the defects of the main model, the data-driven and artificial intelligence algorithms are used to correct the attack prediction model to obtain the optimal defense strategy. The typical case has three parts. Part A introduces the case application framework, Part B introduces the key principles and theoretical analysis, and Part C shows the practical application effect, and the three progress in layers to clearly define trustworthiness in industrial IoT systems based on the artificial intelligence.

A. CPS Trustworthiness Robust Control Strategy

For the attack defense trustworthiness problem of the CPS network, this section proposes a virtual power plant (VPP) distributed economic scheduling strategy. Based on the designed strategy, even in the network attack threat environment, the microincreasing rate of the distribute generation (DG) group can converge to the optimal solution of the optimization problem according to the convergence rule. In this strategy, a distributed credibility manager is set up to monitor the status of the credibility information of each DG through the designed observing

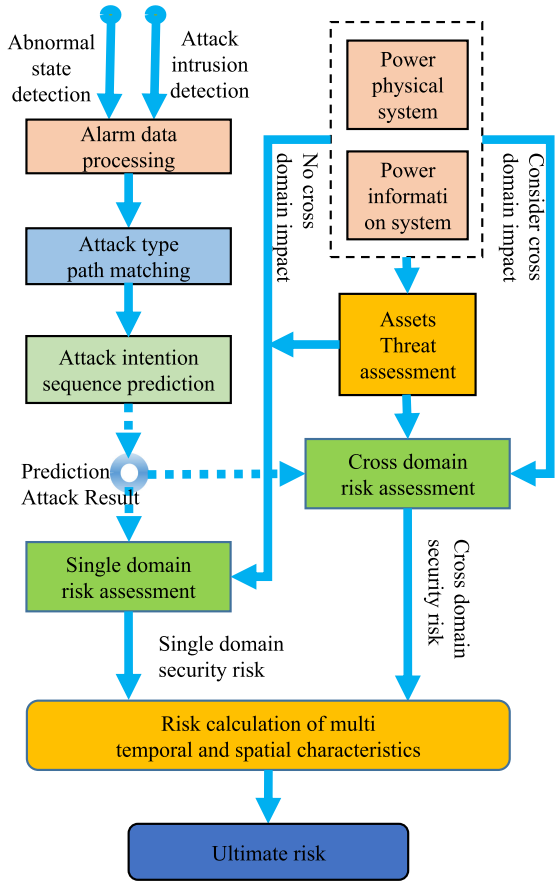


Fig. 4. Overall frame map.

communication network. When the confidence value of a DG falls below the set threshold, it will be isolated by the modulation and coding scheme (NCS) to avoid the serious harm that the malicious attack may cause to the CPS.

1) Network to CPS Attack Detection Process: To achieve the detection of maliciously following DG n, DG $i \in N_n^+$ uses the N^- information from the observed communication subnetwork to assess the feasible range of its rate of increase. In terms of implementation technology, this information can be added to the transmission packet of DG n and sent to all DG $i \in N_n^+$.

When certain conditions are met, by virtue of the information in the observed communication subnetwork, any DG $i = \{1, \ldots, N\}$ can estimate the upper and lower limits of the microincrement rate in the next iteration of any of its receiving neighbors DG $j = N_n^-$.

where η is the lower bound of the adjacency matrix element; $\lambda_j^{\max}(k)$ and $\lambda_j^{\min}(k)$ are defined as $\lambda_j^{\max}(k) = \max\{\lambda_l(k) | \forall l \in N_j^-\}$, $\lambda_j^{\min}(k) = \min\{\lambda_l(k) | \forall l \in N_j^-\}$, and $D_j^{\max}(k)$ $D_j^{\min}(k)$ are defined as $D_j^{\max}(k) = \{l | \lambda_l(k) = \lambda_j^{\max}(k)\}$, $D_j^{\min}(k) = \{l | \lambda_l(k) = \lambda_j^{\min}(k)\}$.

The above formula gives the maximum and minimum possible reasonable values of the microincrement rate of DG j in the next iteration period. If the following inequality is not true

$$\lambda_j^{\text{low}}(k) \le \lambda_j(k) \le \lambda_j^{\text{up}}(k) \tag{10}$$

then DG i suspects that it receives the neighbor DG j state abnormality, determines that it is a suspicious node, and prepares to initiate the edge and isolation process; where $\lambda_j(k)$ is the microincrement rate information that DG i receives from DG j. This establishes a detection mechanism for the following DG.

On the other hand, in the mechanism for detecting the malicious leader DG, one of the neighboring DGs of the leader DG can be set as the agent leader DG, which is responsible for direct monitoring. Take the denial of service (DoS) attack and the replay attack as an example. The detection principle is as follows.

If the following relationship is true

$$\lambda_l(k) = \phi. \tag{11}$$

It is believed that the leading DG is attacked by DoS.

If the following relationship is true

$$|P_{\text{ref}} - P_{\text{out}}| > e_{\text{conv}} \text{ and } \lambda_l(k) \equiv \lambda_l(k+1). \tag{12}$$

The leader DG is considered to be replayed, where e_{conv} is a reasonable convergence error constant.

2) Marginalization and Isolation Process: The deep&cross network link manager (DCLM) located local to each DG information terminal can maintain and update the confidence $c_{ij}(k)(\forall j \in N_i^-)$ of its receiving neighbor DG j. The initial value of the credibility is set to $c_{ij}(0)$. The credibility update principle is

$$c_{ij}(k+1)\begin{cases}\max\{c_{ij}(k)-1,0\}, & \text{If DG } j \text{ Suspicious}\\ c_{ij}(k), & \text{Other}\end{cases}. \tag{13}$$

For the malicious DG j that is attacked, the sending neighbor DG i will gradually reduce its credibility. At this time, DG j will be gradually marginalized according to the following rules:

$$a_{ij}(k) = \frac{[c_{ij}(k)]^+_{\tilde{c}_{ij}}}{\sum_{l\in N_i^-}[c_{ij}(k)]^+_{\tilde{c}_{ij}}} \tag{14}$$

where c_{ij} is the threshold of credibility $c_{ij}(k)$, which can be calculated by an advanced intelligent algorithm; $[x]^+_y$ is defined as

$$[x]^+_y = \begin{cases} x, & \text{if } x \geq y\\ 0, & \text{other}\end{cases}. \tag{15}$$

Assume that the adjacency matrix $\mathbf{A}(k)$ is a row random. According to the above principle, the weight of the control communication line connected to the malicious DG will gradually decrease until it is finally 0 in which time its transmitting neighbor will disconnect all communication lines between them. When all relevant control communication lines are removed, the malicious DG is also completely isolated from the communication network. At the same time, the malicious DG is also physically isolated from the distribution network, completely preventing its possible harmful effects on the entire system.

It should be noted that in the case where the leader DG is attacked when the original malicious leader DG is isolated, the agent leader DG starts to receive the information $P_{\text{ref}} = P_{\text{out}}$ delivered by the superior express mail service (EMS) so that the VPP is again controllable.

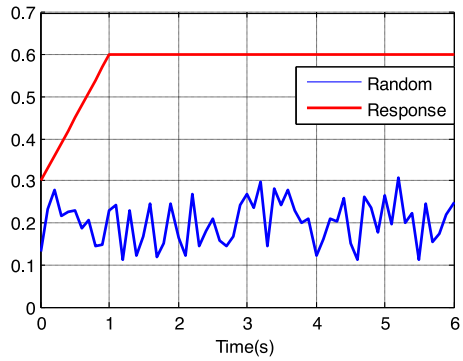


Fig. 5. Robustness.

The VPP distributed economic scheduling strategy robust to network attacks proposed by this research has strong convergence characteristics and can resist various types of network attacks, including noncollusion and collusion network attacks, and various types of communication failures. Baton. When a malicious DG violates the above conditions, it will be gradually marginalized and isolated from the NCS. When the output random disturbance is added to the input, the output is still stable, robustness' intuitive expression is shown in Fig. 5.

B. CPS Trustworthiness Assessment Method

The traditional grid risk assessment is generally based on historical data or current data. The results obtained are "static" results in the current state of a single time scale, which are transient short-term scale risks, while multitime scale risk assessments are required. Consider the risk of the system for a period of time in the future, namely the long-term scale risk. The power CPS trustworthiness assessment proposed in this article comprehensively considers the single-domain trustworthiness and the cross-domain trustworthiness. The former represents the direct loss caused by the attack on single information and physical systemand only considers the current vulnerability of the system and the risk of being attacked. The static short-time scale risk can guide the defender to protect the relevant equipment in time; the latter considers the fault propagation process caused by the attack, and studies the system risk in the future period after the attack occurs, which belongs to the long-term scale risk, and the power CPS. The replanning of the safety structure has important guiding significance. Therefore, the combination of the two can more fully understand the trustworthiness of CPS.

The safety risk calculation method in this article adopts the classic risk assessment model—the Rand risk assessment model: RT = PVC. RT is the expected value of the ultimate loss of the target, representing the risk; P is the probability of a target being attacked, "threatening"; V is the probability that a target will eventually be destroyed after being attacked, "vulnerability"; C is a target. The magnitude of the damage caused by the attack, and eventually being destroyed, represents the consequences.

The cross-domain trustworthiness assessment method adopted in this article first establishes an information system network model $N_1 = (V_1, \alpha_1)$ and a power system network model $N_P = (V_P, \alpha_P)$ based on the complex network theory. Among

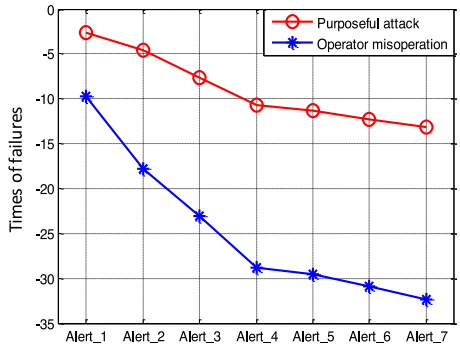


Fig. 6. Attack experimental warning.

them, V_1 is a set of middle nodes in N_1, where N_1 satisfies the idempotent rate, that is, the proportion of nodes whose degree is k in the network N_1 is $P_1(k)$. Then, using the attack prediction result as input, combined with the power CPS network topological degree distribution characteristics and the information physical interaction system cross-domain connection characteristics, the probability model of the attack fault cross-domain propagation process is established, and the dynamic fault probability is calculated. The times of successes are positive, and the times of failures are negative. The attack experimental warning is shown in Fig. 6.

The online queuing algorithm is used to predict the system state in real time. The online queuing algorithm can evaluate CPS transparently and automatically. It can work in parallel with CPS without any influence. It has high feasibility. The whole evaluation process works online and circularly, keeping synchronous with the operation of the CPS system, which can effectively improve the system reliability. The workflow steps are as follows.

Step 1: Obtain system status information and component experience data set information from CPS.

Step 2: According to the collection and information, the online queuing algorithm is used to build the prediction queue to evaluate the reliability of the system.

Step 3: As the evaluation result, the predicted queue is fed back to the operator's visual interface in the form of an alarm.

Step 4: The operator shall refer to the alarm information and take corresponding improvement measures for specific components of the components that may trigger the next system fault to ensure that the CPS operates without fault.

In CPS, a physical system node is called a power element, an information system node is called an information element, a power element that provides power support for an information element is called a supporting power element, and an information element that provides control support for a power element is called a supporting information element. Define the meaning of the power element/information element fault to avoid confusion with the traditional faults.

The pseudocode of the online queuing algorithm is shown in Fig. 7.

Because the system will not fail only if its own node is not faulty and at least one neighbor node is not faulty, that is, the

OnlineRank($L, \Delta, \gamma, \theta, \tau, \lambda$)

1 $L \leftarrow \{\}$

2 while {true}

3 do At time t equals t

4 Ranking $_i$ (t) $\leftarrow m_i$ (Capture(t)) for $m_i \in M$

5 $\Lambda \leftarrow$ The top Δ models in model queue Q

6 Ranking(t) Weighed_average($\{\omega_i \times$ Ranking(t) $\mid m_i \in \Lambda\}$)

7 if new fault occurs Faults(n, t)

8 then z_i=Estimate$_{mi}$(n, t) for $m_i \in M$

9 $z_{best} \leftarrow \max(z_1, z_2, \bullet\bullet\bullet, z_{|M|})$

10 $z_{worst} \leftarrow \max(z_1, z_2, \bullet\bullet\bullet, z_{|M|})$

11 for $m_i \in M$

12 do $d_i = (z_{best} - z_i)/(z_{best} - z_{worst})$

13 $\omega_i = \omega_i \times \gamma^{di}$

14 if elapsed time interval λ

15 then add a new model $m_{|M|+1}$ to M

16 $\omega_{new} \leftarrow \omega_{min} + \tau \times (\omega_{max} - \omega_{min})$

17 $\omega_{|M|+1} \leftarrow \omega_{new}$

18 $M \leftarrow M \cup \{m_{|M|+1}\}$

19 if $|M| > L$

20 then delete the worst model according to $\omega_i \times \theta^{age(mi)}$

Fig. 7. Pseudocode of the online queuing algorithm.

nonfault probability is $1 - \gamma = (1 - \theta)(1 - \rho)$, so the average failure probability of the node is

$$\gamma = 1 - (1 - \theta)(1 - \rho) = \theta + \rho + \theta\rho. \tag{16}$$

After the above attack failure cross-domain propagation algorithm, the final power CPS failure probability is obtained (γ_1, γ_p). Since the fault loss is the product of the node average failure probability γ and the total node value C, the final cross-domain trustworthiness can be calculated according to the power element and information element value and the vulnerability index C_i^p, V_i^p and C_j^i, V_j^i, obtained from the single-domain trustworthiness assessment phase.

$$R_c = \gamma_1 \sum C_i^I V_i^I + \gamma P \sum C_j^P V_j^P. \tag{17}$$

The security protocol research includes two aspects: node security and information security between the nodes. The main solutions for current node security include trusted root-based identity authentication, which uses the trusted platform module as a built-in trusted root and uses this as a starting point, combined with secure boot, data storage protection, integrity metrics, and trusted delivery. And other technologies adopt step-by-step measurement, verification, and delivery methods to complete node credibility evaluation. In this article, a secure

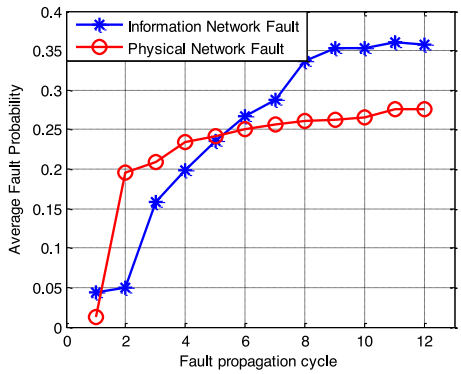


Fig. 8. Cascade mean failure probability change caused by the attack.

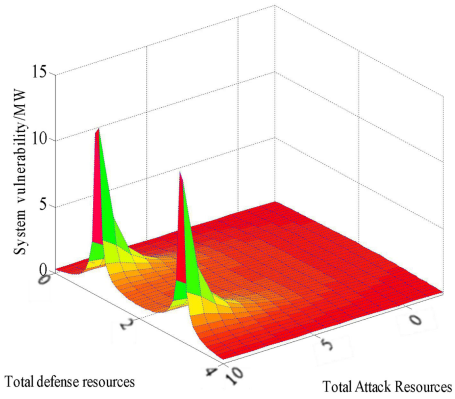


Fig. 9. 3-D map of the impact of changes in total offensive and defensive resources on system vulnerability.

link state routing protocol based on the node credibility is applied. By introducing trust and complex process capability index (CPK)-based security authentication mechanisms in the network, the security of information interaction is enhanced on the one hand, and the node is considered comprehensively in routing decision. Credibility and path credibility are the paths to choose high-security credibility for information transmission.

C. Analysis of Safety Risk Assessment Results

In the experiment, the appsim simulation platform is used, the MATLAB/Simulink software is integrated, the data classification, redundancy check, and trend monitoring are used to build the model, and self-tuning is realized by this method. For the acquired system state information, 80% of the data are used for classification model training, and 20% of the data are used for a blind test. In order to check the redundancy, first, the history and test system state information instances are transformed into vector groups by using a vector space model.

Fig. 8 shows the cascade mean failure probability change caused by the attack. It is known that θ is the probability of failure caused by the failure of the node itself in the system, ρ is the failure probability caused by all the failures of the neighbor nodes in the system, and γ is the probability that any node in the system will leave the maximum connectivity graph of the network due to the failure, that is, the average failure of the node probability.

The total number of offensive and defensive resources S^A and S^D are two key factors affecting the level of system vulnerability during the game phase. The sensitivity analysis of them is shown in Fig. 9.

The simulation results and the analysis conclusions of this section can help guide the CPS defenders of the distribution network with distributed collaborative control mode and choose how to invest the appropriate amount of defense resources according to the CPS attack and defense environment.

In the game phase, in addition to the influencing factors of the network attack and defense, the dynamic convergence process of the distributed collaborative control itself will also affect the system vulnerability assessment results. The following is a description of the influence mechanism through the simulation example and then the sensitivity analysis.

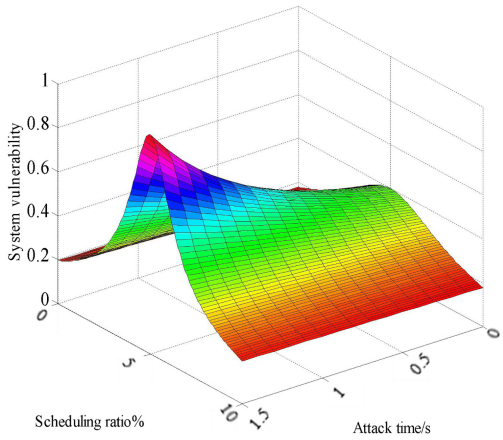


Fig. 10. 3-D map of the impact of collaborative control dynamic processes on system vulnerability.

The simulation analysis of the impact of network attack time and control scheduling commands on the vulnerability of distribution network CPS.

The simulation results are shown in Fig. 10, where the scheduling ratio is defined as the ratio of the reference power P_{ref} to the maximum allowable output of the VPP, which shows that the VPP dynamically adjusts according to each control scheduling command, and the attack and defense game is performed at each moment. Corresponding system vulnerability: It can be seen from the figure that in the process of VPP accepting a certain instruction P_{ref} scheduling, the system vulnerability changes obviously at the beginning stage and gradually becomes fixed after a period of time; meanwhile, as the P_{ref} increases, the system vulnerability gradually increases. Upgrade: This is because under the new scheduling command, each DG will initially adjust its own output quickly, and P_{ref} is positively correlated with the value of the scheduling command. According to this, the grid defender can prearm according to the VPP scheduling situation and reasonably plan the investment timing of the defense resources according to the vulnerability of the system.

It can be seen that the control system has strong fault tolerance under a certain (structure and size) parameter perturbation and maintains the system's strong defensiveness. According to the different definitions of performance, it can be divided into stable robustness and performance robustness. The defense robustness obtained in this article not only has its own stable robustness but also has the performance robustness.

IV. Conclusion

In this article, the reliability of CPS was studied, and an automatic online evaluation method of CPS reliability based on machine learning was proposed. In one step, the article mainly focused on the CPS model, algorithm, and implementation tools.

With the continuous development of CPS in recent years, the development of high real-time, automatic, and intelligent CPS had gradually become the common demand of various industries. With the application of information network technology, such as sensor, embedded processing, digital communication, artificial intelligence, and so on, the various vulnerabilities and vulnerabilities in the information network seriously affected the safe operation of CPS and caused great losses while improving the performance and efficiency of existing systems. With the maturity of artificial intelligence technology, the application of artificial intelligence in the field of network physical space credibility not only improves the response speed and the response speed of various threats in cyberspace but also improves the predictability and accuracy of risk prevention in an all-round way. Therefore, artificial intelligence technology had been fully applied in the field of CPS and played a great potential in dealing with all kinds of human integrity problems in the intelligent era.

References

[1] Y. Chu, Q. Chen, Z. Fan, and J.-J. He, "Tunable V-cavity lasers integrated with a cyclic echelle grating for distributed routing networks," *IEEE Photon. Technol. Lett.*, vol. 31, no. 12, pp. 943–946, Jun. 15, 2019.

[2] M. Muhammad, R. Talat, A. H. Sodhro, and S. Pirbhulal, "A multi-sensor data fusion enabled ensemble approach for medical data from body sensor networks," *Inf. Fusion*, vol. 53, pp. 155–164, 2020.

[3] C. Lu, W. Liang, M. Gao, S.-N. Luo, and Y. Lin, "Terahertz transmittance of cobalt-doped VO $_2$ thin film: Investigated by terahertz spectroscopy and effective medium theory," *IEEE Trans. THz Sci. Technol.*, vol. 9, no. 2, pp. 177–185, Mar. 2019.

[4] Y. C. Lin *et al.*, "Development of advanced manufacturing cloud of things (AMCoT)—A intelligence manufacturing platform," *IEEE Robot. Autom. Lett.*, vol. 2, no. 3, pp. 1809–1816, Jul. 2017.

[5] C.-C. Chen, M.-H. Hung, P.-Y. Li, Y.-C. Lin, Y.-Y. Liu, and F.-T. Cheng, "A novel automated construction scheme for efficiently developing cloud manufacturing services," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1378–1385, Jul. 2018.

[6] S. Shin, H. Wang, and G. Gu, "A first step toward network security virtualization: From concept to prototype," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 10, pp. 2236–2249, Oct. 2015.

[7] R. Talat, M. S. Obaidat, M. Muzammal, A. H. Sodhro, Z. Luo, and S. Pirbhulal, "A decentralised approach to privacy preserving trajectory mining," *Future Gener. Comput. Syst.*, vol. 102, pp. 382–392, Jan. 2020.

[8] H. Xie, Z. Yan, Z. Yao, and M. Atiquzzaman, "Data collection for security measurement in wireless sensor networks: A survey," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2205–2224, Apr. 2019.

[9] H.-M. Wang, T.-X. Zheng, J. Yuan, D. Towsley, and M. H. Lee, "Physical layer security in heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1204–1219, Mar. 2016.

[10] Y. Yang, H. Peng, L. Li, and X. Niu, "General theory of security and a study case in internet of things," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 592–600, Apr. 2017.

[11] Z. H. Lv *et al.*, "Intelligent security planning for regional distributed energy internet," in *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3540–3547, May 2020.

[12] Y. Gao, H. Ma, D. Abbott, and S. F. Al-Sarawi, "PUF sensor: Exploiting PUF unreliability for secure wireless sensing," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 64, no. 9, pp. 2532–2543, Sep. 2017.

[13] A. Kyriacou, M. P. Michaelides, V. Reppa, S. Timotheou, C. G. Panayiotou, and M. M. Polycarpou, "Distributed contaminant detection and isolation for intelligent buildings," *IEEE Trans. Control Syst. Technol.*, vol. 26, no. 6, pp. 1925–1941, Nov. 2018.

[14] X. Wang, W. Yu, X. Fu, D. Xuan, and W. Zhao, "iLOC: An invisible localization attack to internet threat monitoring systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 20, no. 11, pp. 1611–1625, Nov. 2009.

[15] P. Fairley, "Blockchain world—Feeding the blockchain beast if bitcoin ever does go mainstream, the electricity needed to sustain it will be enormous," *IEEE Spectr.*, vol. 54, no. 10, pp. 36–59, Oct. 2017.

[16] Y. Han, C.-J. Zhang, L. Wang, and Y.-C. Zhang, "Industrial IoT for intelligent steel making with converter mouth flame spectrum information processed by deep learning," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2640–2650, Apr. 2020.

[17] Z. Lv, W. Kong, X. Zhang, D. Jiang, H. Lv, and X. Lu, "Intelligent security planning for regional distributed energy internet," *IEEE Trans. Ind. Inform.*, vol. 16, no. 5, pp. 3540–3547, May 2020.

[18] L. Zhang, G. Zhou, Y. Han, H. Lin, and Y. Wu, "Application of Internet of Things technology and convolutional neural network model in bridge crack detection," *IEEE Access*, vol. 6, pp. 39442–39451, 2018.

[19] H. Rastegarfar, D. C. Kilper, M. Glick, and N. Peyghambarian, "Cyber-physical interdependency in dynamic software-defined optical transmission networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 7, no. 12, pp. 1126–1134, Dec. 2015.

[20] L. Ribeiro and M. Bjorkman, "Transitioning from standard automation solutions to cyber-physical production systems: An assessment of critical conceptual and technical challenges," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3816–3827, Dec. 2018.

[21] D. Roy, L. Zhang, W. Chang, S. K. Mitter, and S. Chakraborty, "Semantics-preserving cosynthesis of cyber-physical systems," *Proc. IEEE*, vol. 106, no. 1, pp. 171–200, Jan. 2018.

[22] C. Lu *et al.*, "Real-time wireless sensor-actuator networks for industrial cyber-physical systems," *Proc. IEEE*, vol. 104, no. 5, pp. 1013–1024, May 2016.

[23] H. Ye, Q. Mou, X. Wang, and Y. Liu, "Eigen-analysis of large delayed cyber-physical power system by time integration-based solution operator discretization methods," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 5968–5978, Nov. 2018.

[24] S. Xin, Q. Guo, H. Sun, B. Zhang, J. Wang, and C. Chen, "Cyber-physical modeling and cyber-contingency assessment of hierarchical control systems," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2375–2385, Sep. 2015.

[25] U. Adhikari, T. Morris, and S. Pan, "WAMS cyber-physical test bed for power system, cybersecurity study, and data mining," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2744–2753, Nov. 2017.

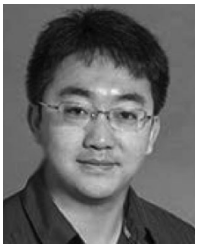
[26] S., Xin, Q. Guo, H. Sun, C. Chen, J. Wang, and B. Zhang, "Information-energy flow computation and cyber-physical sensitivity analysis for power systems," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 2, pp. 329–341, Jun. 2017.

[27] M. U. Khan, S. Li, Q. Wang, and Z. Shao, "CPS oriented control design for networked surveillance robots with multiple physical constraints," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 35, no. 5, pp. 778–791, May 2016.

[28] K. L. Keller, "Leveraging biologically inspired models for cyber-physical systems analysis," *IEEE Syst. J.*, vol. 12, no. 4, pp. 3597–3607, Dec. 2018.

[29] R N. M. Watson *et al.*, "Fast protection-domain crossing in the CHERI capability-system architecture," *IEEE Micro*, vol. 36, no. 5, pp. 38–49, Sep./Oct. 2016.

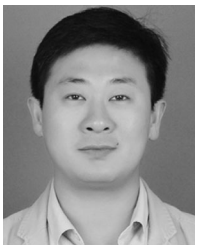
[30] X. Liu, M. Shahidehpour, Z. Li, X. Liu, Y. Cao, and Z. Li, "Power system risk assessment in cyber attacks considering the role of protection systems," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 572–580, Mar. 2017.



Zhihan Lv (Senior Member, IEEE) received the Ph.D. degree in computer applied technology from Paris Diderot University, Paris, France, and the Ocean University of China, Qingdao, China, in 2012.

He was a Research Associate with University College London. From 2012 to 2016, he was an Assistant Professor with the Shenzhen Institutes of Advanced Technology, *Chinese Academy of Sciences*. He was a Research Engineer with Centre National de la Recherche Scientifique, Paris, France, a Postdoctoral Research Fellow with Umea University, Umea, Sweden, and an Experienced Researcher with Fundación FIVAN, Valencia, Spain. He is currently an Associate Professor with Qingdao University, Qingdao, China.

Dr. Lv is a Program Committee Member of ACM IUI 2015-2020, IEEE BIGDATA4HEALTH Workshop 2016, IEEE/CIC WIN Workshop 2016, IIKI 2016, and WASA 2016. He has been an Associate Editor for the *PLOS One*, since 2016, the IEEE ACCESS, since 2016, the *Neurocomputing*, since 2016, and the *IET Image Processing*, since 2017, and is a Lead Guest Editor for the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE NETWORK, and IEEE SENSORS.



Yang Han was born in Jizhou, Hebei, China, in 1987. He received the B.S. and M.S. degrees in science and mathematics from the North China University of Science and Technology, Qinhuangdao, China, in 2010 and 2018, respectively, where he is currently working toward the Ph.D. degree in engineering in iron and steel metallurgy.

His main research interests include numerical computation, iron and steel big data, and intelligent computation.

Amit Kumar Singh (Senior Member, IEEE) received the Ph.D. degree in computer engineering from the National Institute of Technology Kurukshetra, Haryana, India, in 2015.

He is currently an Assistant Professor with the Department of Computer Science and Engineering, National Institute of Technology Patna, Bihar, India. He has authored more than 100 peer-reviewed journals, conference publications, book chapters, and three books, and has edited four books with internationally recognized publishers, such as Springer and Elsevier. His research interests include data hiding, image processing, biometrics, and cryptography.

Dr. Singh is the Associate Editor for the IEEE ACCESS, since 2016, *IET Image Processing*, since 2020, and the former member of the Editorial Board of Multimedia Tools and Applications, Springer, from 2015 to 2019. He has edited various international journal special issues as a Lead Guest editor. He has obtained the memberships from several international academic organizations, such as ACM and IEEE.

Gunasekaran Manogaran received the bachelor of engineering and the master's of technology degrees from Anna University, Chennai, India, and Vellore Institute of Technology, Vellore, India, in 2010 and 2012, respectively, and the Ph.D. degree in information technology from the Vellore Institute of Technology, Vellore, India, in 2017.

He is currently a Big Data Scientist with the University of California, Davis, CA, USA, an Adjunct Assistant Professor with the Department of Computer Science and Information Engineering, Asia University, Taichung City, Taiwan, an Adjunct Faculty with the School of Computing, SRM Institute of Science and Technology, Kattankulathur, India, and a Visiting Researcher/Scientist with the University of La Frontera, Colombia, and the International University of La Rioja, Logroño, Spain. He is the author/co-author of more than 100 papers. His current research interests include big data analytics, Internet of Things, and soft computing.

Dr. Manogaran is a member of the IEEE Society and International Society for Infectious Diseases and Machine Intelligence Research labs.

Haibin Lv graduated and Master degree in marine geology from the First Institute of Oceanography, State Oceanic Administration, Qingdao, China, in 1990.

He is currently a Professor with the North China Sea Offshore Engineering Survey Institute, Ministry of Natural Resources North Sea Bureau, Beijing, China.