Hindawi
Computational Intelligence and Neuroscience
Volume 2022, Article ID 4043309, 11 pages
https://doi.org/10.1155/2022/4043309

Research Article

A Semi-Self-Supervised Intrusion Detection System for Multilevel Industrial Cyber Protection

Fuchuan Ye[1] **and Weiqiong Zhao**[2]

[1]*Information and Educational Technology Center, Southwest Minzu University, Chengdu 610041, China*
[2]*School of Intelligent Technology, Geely University of China, Chengdu 641423, China*

Correspondence should be addressed to Fuchuan Ye; fuchuan_ye@163.com

Received 30 June 2022; Revised 12 August 2022; Accepted 12 August 2022; Published 21 September 2022

Academic Editor: Konstantinos Demertzis

Industry 4.0 affects all components of the modern industry value chain. The accelerating use of the Internet and the convergence of industrial and operational networks constantly increase the need for secure industrial communication solutions. Therefore, "multilevel industrial cyber protection" is critical to Industry 4.0. In general, industrial protection refers to safeguarding information and data and the intellectual property rights of production processes related to the overall industry environment. The availability, integrity, and confidentiality of systems must be maintained. The goal challenge is the best possible protection from attacks and threats which create immediate financial damage and other risks in the industry (reputation, etc.). Based on the Defense-in-Depth strategy, a holistic, multilayered, and in-depth protection of industrial systems is developed in this paper. Specifically, a Semi-Self-Supervised Intrusion Detection System (S3IDS) is proposed, which combines advanced machine learning techniques for industrial data noise reduction to automate the discovery and separation of classes, which are essentially equivalent to cyber-related anomalies. As demonstrated by a mathematical simulation based on computational number theory and specifically on the concept of the single object, the proposed S3IDS learns to accurately reconstruct samples to predict the nature of an anomaly created directly by the industrial ecosystem.

1. Introduction

Historically, industrial companies worldwide have approached cybersecurity in their Information Technology (IT) and Operational Technology (OT) networks very differently [1]. Most companies have already implemented technological infrastructures for detecting and dealing with network threats, but, for their industrial (OT) systems, coping with cyber threats is usually limited to isolating the relevant procedures from the rest of the network. Industries are constantly being "digitized" by investing more and more in intelligent technologies, new automation systems, and other applications that promote productivity growth or improve many other indicators of interest to the organization [2]. This rapidly equates IT systems with OT systems, making the latter more vulnerable to attacks that formerly solely affected the former [3].

Cyberattacks on industrial organizations are considered a perilous threat [4], as they have the potential to cause significant material losses and lead to disruption of the production cycle of the entire system [5]. They target, among others, industrial control and data collection systems (ICSs, SCADA) [6, 7]. In addition, due to the sensitive information available to industrial organizations, they are usually an attractive target for attackers [3].

The situation to date focuses on the human aspect, experience, and expert opinion, using assistive technology to analyze and reduce risks and dangers to industrial infrastructure [8, 9]. For optimal results with this methodology, there should be up-to-date threat intelligence, incident reports, and vulnerability warnings, which will feed indefinitely the power grid monitoring tools and in-depth human oversight and intervention from cybersecurity staff [10].

The above passive function, in combination with the new class of requirements in cybersecurity, leads to the logic of adopting solutions that include fully automated security methods based on advanced techniques of artificial intelligence [11], with the parallel minimization of human intervention [12]. The idea of getting rid of the constant surveillance and direct presence of people is related to advanced attacks like Stuxnet and BlackEnergy, where it turned out that it just needed an infected USB stick or open a phishing e-mail to allow the attacker to access an isolated industrial network [5]. In addition, throughout the last several months, we have witnessed, in many cases, highly specialized attacks on systems and infrastructures that use industrial protocols [4, 13].

A great example is the largest colonial gas pipeline in the USA, which was shut down for several days after a malicious cyberattack and attributed to the shadow criminal group DarkSide [14–16]. Also, in the first quarter of 2021, in the city of Oldsmar in Florida, there was an attack on the government infrastructure responsible for the city's water supply. In essence, a remote attempt was made to change the mix of related chemicals with water disinfection, resulting in the mass poisoning of consumers [7, 15, 17, 18].

Just months ago, another cyberattack occurred in a public hospital in Israel. The specific attack created significant problems in the smooth provision of services of the organization, while it required the payment of a certain amount of money (ransomware). Hospital services reacted by using alternate resources and support systems. Fortunately, there was no loss of life in this case, as unfortunately happened at a similar point in a hospital in Germany a few months ago [4, 6, 15].

To deal with these offensive techniques, the research community has proposed various solutions in which machine learning systems operate with self-adaptation procedures and rearrange their mode of operation, depending on the algorithms' hyperparameters that most often specify their mode of operation.

2. Literature Review

Several researches have presented adaptive cyberattack detection algorithms to fulfill the requirement for continuing learning paradigm changes [19–21]. Still, they have failed to establish a more comprehensive system of knowledge for detection performance and their evaluation practice [22–24].

In their review of earlier work for threat detection techniques in industrial control systems, Kaouk et al. [3] underlined the difficulties and advantages of putting such solutions into practice. Such information, in our opinion, will be helpful for future studies in manufacturing security. ICS intrusion detection technology is evolving swiftly, but there is still room for improvement. The integration of IDS with ICS will face a variety of risks. Most methods used in the literature are anomaly based, meaning that they look for any notable departure from the norm. To enhance how IDS can react to alarms, techniques that can tell the difference between a flaw and a threat are desperately needed. Another difficulty is that the vast majority of existing IDS are network based and cannot access encrypted data because of this. For instance, encryption use is hampered by hardware limitations. However, new parts of the ICS have begun to offer encryption because of advancements in hardware computation capabilities. IDS must therefore rely on data sources other than Internet activity. The operation of IDS should also be taken into account as ICS grow in size and complexity and comprise geographically dispersed systems. Alternative technologies decentralized and collaborative IDS must therefore be created. Such information, in our opinion, will help advance future studies on the integrity of ICS.

Hu et al. [25] went into more detail on ICS's attributes and security needs in 2018. They proposed a taxonomy of IDS for industrial control systems based on three techniques: protocol analysis, traffic mining, and control process analysis. They also examined the benefits and drawbacks of various IDS categories. They concluded that, despite the rapid advancement of ICS technology, there is still much opportunity for ICS IDS development. It was crucial to construct dispersed and collaborative IDS due to the scattered structure of ICS subsystems. Evaluating associations between distributed IDS, fusing a group of dispersed and potentially contradictory detection findings, and obtaining accurate and real-time complete detection results are a novel and intriguing subject. How to react to warnings is a major problem for ICS IDS. In specific control systems, simply notifying administrators of the alarm may be considered sufficient; nevertheless, automatic reaction mechanisms must be taken into account to ensure the protection and reliability of ICS. How to automatically improve intrusion detection algorithms while they are being used is a crucial topic. To maintain a satisfactory detection accuracy, intrusion detection algorithms must automatically optimize their judgments of changing contexts. ICSs typically need to operate continuously, and the system parameters (such as durable components, access controls, and system constraints) of an objective ICS may change over time. These days, ICSs are internet-accessible, and ICS security concerns are increasingly becoming more critical. Traditional IDS created for IT platforms cannot function well on ICS because of its uniqueness. It can assist ICS in identifying various intrusions and lowering the frequency of industrial mishaps caused by malicious attacks.

Adversarial Machine Learning, often known as cyberattacks over neural network models on Engineering IDS, was examined by Anthi et al. [4]. By constructing adversarial samples and evaluating classification patterns, they studied how adversarial learning may be used to target supervised models. As adversaries could be able to get beyond the defenses, such attacks could have dire effects on ICS systems. This can result in delayed assault detection, which might harm the infrastructure, cause financial loss, or even result in fatalities. An actual electric grid data set was utilized for training and evaluating commonly used unsupervised feature learning classifiers in support of the studies described here. The investigation also studies how adversarial training on such sets can enhance the resilience of supervised models. Using the testing data, adversarial samples were created with

various combinations that changed the model's interference and complexity.

According to Ayodeji et al. [4], in 2020, the failure to recognize and distinguish between the intrinsically identical signatures that define normal transients typical of complex systems contributes significantly to false alarms in ICS systems. The majority of machine learning-based detection techniques created for Scada Systems (ICSs) are taught on network packet logs and solely rely on network layer traffic monitoring to identify intrusions. They looked at the most current developments in malware detection algorithms, their shortcomings, difficulties, and the state of their use in crucial infrastructures. Additionally, they started a conversation about the parallels and differences between the growth of computational skills and equipment for classification and hacking in defense of complex systems and the requirement to distinguish between them clearly. They used nuclear energy controllers as a case study to demonstrate the challenges to a smooth changeover of security algorithms. To significantly reduce the number of annoyance warnings generated, they suggested a method that considers the subtleties in the data utilized in creating machine learning algorithms. The current findings and recommended course of action lay the groundwork for creating robust intrusion detection systems that significantly reduce the problem of false alarms that plague existing intrusion detection systems.

The transition of the ICS from isolated systems to virtualized platforms was closely examined by Bhamare et al. [1], who also noted the considerable efforts made by both business and technology to construct secure ICSs and the relevance of machine learning approaches for ICS cybersecurity. ICS security remains a concern despite the recent popularity of big data insights and cloud computing. Cloud platforms will eventually help ICSs and industries. Still, inadequate security in cutting-edge multicloud platforms could result in expensive security breaches in real-time industry platforms. It is incredibly challenging to prevent and identify assaults at the ICS component level due to the sophistication of emerging viruses attacking control systems, including rootkits and zero-day attacks. New intrusion detection strategies for ICS devices at the production control level are thus required. Additionally, they said that a testbed might help with the difficulties of safeguarding an industrial process by offering more information about how the method is managed with the aid of sensors and control laws and comprehension of the security needs, mainly to handle control using cloud-based services.

An examination of the development and usefulness of security mechanisms that have been put out in both industry and academia was presented by Rubio et al. in [2]. In the past several years, there has been a tremendous advancement in the design of security methods for industrial environments [1]. Advanced solutions like honeypot systems and data correlation systems are integrated into commercially accessible products, but innovative detection techniques and architectures are also created in academia [19]. Research is still needed in several areas, including the viability and incorporation of proactive defenses, the deployment of defensive mechanisms in the IIoT and cloud computing, and the emergence of Industry 4.0[26]. Furthermore, to validate defense mechanisms against Advanced Persistent Threats (APTs) and make them more integrable and usable so they can be readily integrated into more crucial infrastructures, it is vital to take into account existing APTs and APT phases [15, 16, 21].

In this spirit, an approach is needed that with minimal configuration and the necessary training samples each time will be able to create a generalized framework for detecting known and unknown attacks on a network. Based on the above challenge and the Defense-in-Depth strategy, in general, S3IDS is proposed that should be applied in the industry. Using advanced machine learning methods automates recognizing anomalies related to cyberattacks [27]. To prove the applicability, we used mathematical simulation based on computational number theory. Mathematical simulation is a process to identify and predict the behavior, performance, and optimization of some physical or abstract systems corresponding to various scientific and engineering applications.

3. Proposed S3IDS

Given the general issues of machine learning systems to deal against serious cyberattacks effectively and with minimal human intervention, this work proposes the creation of an innovative computer intelligence system [28, 29], with minimal human intervention [30–32], significantly strengthening the security mechanisms of network infrastructure [12, 23, 33]. In particular, S3IDS is proposed, an advanced cyber threat detection system, which is a highly innovative tool for operational security. Specifically, we implement a semi-self-adapted machine learning methodology [9–11] based on Semi-Self-Supervised Learning, which may determine the sort of attack based on generic reshaping characteristics generated directly from the unknown online environment and web data [22, 34, 35].

The proposed system's major innovation is based on computational number theory, notably the idea of a monoid object in a category. Monoids are semigroups that have an identity. A monoid is a set containing an associative binary operation and an identity member in abstract algebra, a field of mathematics. For example, nonnegative integers with addition form a monoid, with 0 as the identity member. Such algebraic structures may be found in many disciplines of mathematics. In terms of function composition, the functions from a set create a monoid. In general, the morphisms of an item form a monoid in category theory; conversely, a monoid may be considered a category containing a single entity.

Many abstract data types in computer science may have a monoid structure. A succession of monoid components is "folded" or "stacked" to generate a final value in a recognizable pattern. Many iterative algorithms, for example, must update some "current set" at each iteration. A monoid function may be used to represent this pattern cleanly. In particular, the proposed methodology ensures that the correlation of monoid operations can be predicted using a correlation algorithm, effectively using multiple cores [36, 37].

In particular, if A is a nonempty set, the operation on A for any representation of the form $f : A \times A - \longrightarrow A$; e.g., addition and multiplication are operations on Z. The value of f in the pair (a, b) will be denoted by afb. A pair $(G, *)$, where G is a set and $*$ is one operation on G, is called a monoid if the following properties are valid [36, 38, 39]:

$$x * (y * z) = (x * y) * z, \tag{1}$$

such as

$$x * e = x = e * x. \tag{2}$$

If there is another element $k \in G$with the above property, then for every $x \in G$ we have

$$k * x = x = x * k. \tag{3}$$

Thus, we get $k = e * k$ and $e * k = e$, from where $e = k$. Therefore, the element e is unique and is called the neutral element of G. If $x * y = y * x$ is also valid for every $x, y \in G$, then the monoid $(G, *)$is permutable. So, the pairs $(N, +)$, $(Z, +)$, $(Q, +)$ are substitutively monosyllabic with neutral element 0 and the pairs $(N, \cdot)$, $(Z, \cdot)$, $(Q, \cdot)$ are substitutive monoid.

Respectively, if $(G_i, *_i)$ is a monoid with neutral element $e_i\,(i = 1, \ldots, k)$, the set $G_1 \times \cdots \times G_k$ is a monoid with the operation [40, 41]:

$$(x_1, \ldots, x_k) * (y_1, \ldots, y_k) = (x_1 *_1 y_1, \ldots, x_k *_k y_k). \tag{4}$$

Its neutral component is

$$(e_1, \ldots, e_k). \tag{5}$$

If we have a function with the field of definition, the set of positive integers, and a field of values, the set of complex numbers (numerical function), then we denote by A the set of numerical functions, while the numerical function calculates the exponential product of f and g [36, 37, 39]:

$$f * g\colon N\setminus\{0\} \longrightarrow C, n \longmapsto (f * g)(n) = \sum_{ab=n} f(a)g(b), \tag{6}$$

where the pairs (a, b) run through all the natural whose product is equal to n. The correspondence $(f, g) - \longrightarrow f * g$ defines an operation on A, which is called associative multiplication since the pair $(A, *)$ is a permutable monoid. If $g, h \in A$, then for every natural $n > 0$ we have [36, 37, 39]

$$\begin{aligned}[f * (g * h)](n) &= \sum_{ab=n} f(a)(g * h)(b) \\ &= \sum_{ab=n} f(a) \sum_{c\,d=b} g(c)h(d) \\ &= \sum_{ac\,d=n} f(a)g(c)h(d).\end{aligned} \tag{7}$$

Similarly, we get

$$[(f * g) * h](n) = \sum_{ac\,d=n} f(a)g(c)h(d). \tag{8}$$

Therefore, for every natural $n > 0$ it holds

$$[f * (g * h)](n) = [(f * g) * h](n). \tag{9}$$

And so

$$f * g = g * f. \tag{10}$$

Next, consider the numerical function ϵ defined by the relations:

$$\begin{aligned}\epsilon(1) &= 1, \\ \epsilon(n) &= 0.\end{aligned} \tag{11}$$

For every $f \in A$ and natural $n > 1$, we have

$$(f * \epsilon)(n) = \sum_{ab=n} f(a)\epsilon(b) = f(n), \tag{12}$$

where $f * \epsilon = f$. As the operation $*$ is transitive, the relation $\epsilon * f = f$ is also valid. So, the function ϵ is the neutral element for associative multiplication. Therefore, the pair $(A, *)$ is a permutable monoid. $(G, *)$ is a monoid and ϵ its neutral element. A subset H of G is called a submonoid of G if $\epsilon \in H$ and for every $x, y \in H$ it holds $x * y \in H$; that is, the pair $(H, *)$ is also a monoid with a neutral element ϵ.

Based on the above view, $(A, *)$, $(B, \diamond)$, an d $(C, \triangleright)$ are monoids with neutral elements eA, eB, and eC, respectively, and $f : A \longrightarrow B$, $f : B \longrightarrow C$ are monoid morphisms. We will show that the expression $g \circ f$ is a monoid morphism since, for every $x, y \in A$, the composition of two morphisms of monoids is a monoid morphism, which is proved by the following relation [36, 38, 41]:

$$\begin{aligned}(g \circ f)(x * y) &= g(f(x * y)) \\ &= g(f(x) \diamond f(y)) \\ &= g(f(x)) \triangleright g(f(y)) \\ &= (g \circ f)(x) \triangleright (g \circ f)(y).\end{aligned} \tag{13}$$

Also, it holds

$$\begin{aligned}(g \circ f)(e_A) &= g(f(e_A)) \\ &= g(e_B) \\ &= e_C.\end{aligned} \tag{14}$$

This hypothesis creates a process where the data in a machine learning system is predicted with high accuracy (any anomalies are recognized) even when they come slightly modified [22, 33]. The output of the intelligent mechanism can now be considered as a recognition of the input data's shifted prediction, based on the isomorphism of monoids that may appear in the unknown data set (assuming a uniform distribution which, although unknown, includes properties of monoid theory). That is, the output of the intelligent mechanism approaches the displaced version of the input as the intelligent system is trained. The machine learning system learns to distinguish displaced samples using this approach, resulting in highly generalized algorithmic frameworks for detecting abnormalities [19, 20].

Given that the synthesis of two monoid morphisms is a monoid morphism, proving that the inverse representation of a monoid isomorphism is likewise a monoid isomorphism suffices for implementing this mechanism [36, 37, 39].

So, considering $(M, *)$, $(N, \diamond)$ monoids and $f : M \longrightarrow N$ isomorphism of monoids, if $y_1, y_2 \in N$, then there exist $x_1, x_2 \in M$ with $y_1 = f(x_1)$ and $y_2 = f(x_2)$. The above formulation is related to the hypothesis of a supervised learning problem, where a set of training with N samples, $\{X,Y\} = \{x_i, y_i\}_{i=1}^{N}$, where xi $\in R^{n_i}$, y_i is a no-dimensional binary vector with only one input (corresponds to the class x_i) equal to a multidimensional categorization process, where ni and no are the input and output dimensions, respectively. Unlabeled data helps study the data structure of the accessible data set, but classified data aids in learning. With this in mind, we have [36, 41]

$$\begin{aligned} f^{-1}(y_1 \diamond y_2) &= f^{-1}(f(x_1) \diamond f(x_2)) \\ &= f^{-1}(f(x_1 * x_2)) \\ &= \left(f^{-1} \circ f\right)(x_1 * x_2) \\ &= I_G(x_1 * x_2) \\ &= x_1 * x_2 \\ &= f^{-1}(y_1) * f^{-1}(y_2). \end{aligned} \tag{15}$$

If eM and eN are the neutral elements of M and N, respectively, then $f(e_M) = e_N$ and therefore $f^{-1}(e_N) = e_M$ and f^{-1} is a monoid morphism.

$(G, *)$ is a monoid with neutral elements e and $x \in G$. Assume that $y \in G$ exists such that

$$x * y = e = y * x. \tag{16}$$

In this case, the element y is unique because if y' is another element with this property, then

$$\begin{aligned} y &= y * e \\ &= y * (x * y') \\ &= (y * x) * y' \\ &= e * y' \\ &= y'. \end{aligned} \tag{17}$$

So, the element y is symmetric to x. Also, the symmetric of y is x.

But since in a monoid each element does not always have a symmetric, then f must be calculated which has a symmetric element g (associative inverse of f). If and only if $g * f = e$, which is equivalent to $(1)f(1) = 1$, then [38, 40]

$$\sum_{ab=n} g(a) f(b) = 0, \tag{18}$$

for every natural $n > 1$. In general, for every natural $n > 1$, it applies

$$f^*(n) = -\frac{1}{f(1)} \sum_{st=n, t<n} f(s) f^*(t). \tag{19}$$

Therefore, f has an associative inverse if and only if $f(1) \neq 0$. The derivative of the function is

$$f_\Delta(t) = \frac{\mathrm{d}u_\Delta(t)}{\mathrm{d}t} = \{0, ttn < q0h_{1x}/\Delta C, ; 0 \leq t \leq \Delta\, 0, t \geq \Delta. \tag{20}$$

So, the data set is obtained as a subscale of the signal processing process for analyzing and manipulating the physical quantities that define the given problem of information systems security [42]. Thus, when $\Delta \longrightarrow 0$, the duration of the pulse decreases and its height increases, but the area remains constant and equal to the unit. So, we study the function $f(t)$ as an operator that acts on other functions that are smooth at points 0. Thus, we can express the function $f(t)$ as [43–45]

$$\int_{-\infty}^{+\infty} f(t)\varphi(t)\mathrm{d}t = \varphi(0), \tag{21}$$

where $\varphi(t)$ is a test function, for $f(t) = 0$ and $t \neq 0$. So, the above process can be generalized to describe the time-shifted data expressed by the functionf (t-t_0) ADDIN ZOTERO_ITEM CSL_CITATION {"citationID":"l8CicIUS","properties":{"formattedCitation":" [46]\\uc0\\u8211{┤} [47]","plainCitation":" [46–48]","noteIndex":0},"citationItems":[{"id":47,"uris":["http://zotero.org/users/local/knpFELzr/items/RISE.2017.8378144","event":"2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE)","page":"153–156","source":"IEEE Xplore","title":"Moving object detection using self adaptive Gaussian Mixture Model for real-time-applications","author":[{"family":"Ali","given":"Syed Tariq"},{"family":"Goyal","given":"Kalpana"},{"family":"Singhai","given":"Jyoti"}],"issued":{"date-parts":[["2017",7]]}}},{"id":277,"uris":["http://zotero.org/users/local/knpFELzr/items/ICOMSSC45026.2018.8941982","event":"2018 International Computers, Signals and Systems Conference (ICOMSSC)","page":"378–381","source":"IEEE Xplore","title":"A Method of Fast Extract Signal Subspace Based on the Householder Transformation","author":[{"family":"Chang","given":"Yu"},{"family":"Wan","given":"Qun"},{"family":"Xia","given":"Changxiong"},{"family":"Wan","given":"Yihe"}],"issued":{"date-parts":[["2018",9]]}}},{"id":279,"uris":["http://zotero.org/users/local/knpFELzr/items × 32K complex points) image, achieving real-time-performance.","container-title":"2016 IEEE 13th International Conference on Signal Processing (ICSP)","DOI":"10.1109/ICSP.2016.7877887","event":"2016 IEEE 13th International Conference on Signal Processing (ICSP)","note":"ISSN: 2164–5221","page":"513–517","source":"IEEE Xplore","title":"Design of a flexible high-performance real-time SAR signal processing system","author":[{"family":

"Jin","given":"Ting"},{"family":"Wang","given":"Hongxian"},{"family":"Liu","given":"Hongwei"}],"issued":{"date-parts":[["2016",8]]}}},"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"}
[46–48]:

$$\int_{-\infty}^{+\infty} f(t-t_0)\varphi(t)\mathrm{d}t = \int_{-\infty}^{+\infty} f(t-t_0)\varphi(t_0)\mathrm{d}t = \varphi(t_0)\int_{-\infty}^{+\infty} f(t-t_0)\mathrm{d}t = \varphi(t_0). \tag{22}$$

The above relation describes the mathematical model of the sampling process applied during the application of the semisupervised learning technique of the proposed machine learning model [49–51].

For $\varphi(t) = 1$, we have [52–55]

$$\int_{-\infty}^{+\infty} f(t)\mathrm{d}t = \int_{0^-}^{0^+} f(t)\mathrm{d}t = 1. \tag{23}$$

And so

$$\int_{t_1}^{t_2} f(t-t_0)\varphi(t)\mathrm{d}t = \{\varphi(t_0), t_1 < t_0 < t_2\ 0, t_0 < t_1, t_0 > t_2. \tag{24}$$

However,

$$\int_{t_1}^{t_2} f(\tau - t) f(t-t_0)\mathrm{d}t = f(t-t_0), t_1 < t_0 < t_2. \tag{25}$$

According to the logic presented by the system under consideration, the error function is defined as the integral [56]:

$$erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\mathrm{d}t, x > 0. \tag{26}$$

Also, the complementary error function is defined as the integral [57]:

$$erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^{\infty} e^{-t^2}\mathrm{d}t, x > 0. \tag{27}$$

So, the error function and the complementary error function satisfy the following equation:

$$er\ f(x) + er\ fc(x) = 1. \tag{28}$$

The above hypotheses are proved based on the observation that

$$erf(x) + erfc(x) = \frac{2}{\sqrt{\pi}}\int_0^{\infty} e^{-t^2}\mathrm{d}t. \tag{29}$$

For the calculation of the integral [58],

$$\int_0^{\infty} e^{-t^2}\mathrm{d}t. \tag{30}$$

We consider

$$t^2 = u \Rightarrow 2t\,dt = du. \tag{31}$$

Since the integration ends are the same, we have

$$\int_0^{\infty} e^{-t^2}\mathrm{d}t = \int_0^{\infty} e^{-u}u^{-1/2}\frac{du}{2} = \frac{1}{2}\int_0^{\infty} e^{-u}u^{1/2-1}\mathrm{d}u = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = {}^{(2.9)}\frac{\sqrt{\pi}}{2}. \tag{32}$$

This fact proves the above hypotheses about the relationships of the error functions.

Finally, a self-supervised learning methodology [17, 59, 60] is an unsupervised learning method where supervised learning work is created from unlabeled input data. Simple supervised learning usually requires a lot of labeled data. Obtaining good quality labeled data is a costly and time-consuming task, especially for a complex task such as detecting anomalies. On the other hand, unlabeled information is readily available in abundance. So, the motivation behind the self-supervised learning methodology is to learn useful representations of industrial data from an unlabeled data pool using the semisupervised process and then refine the few-tagged representations for the supervised work.

The implementation of the self-supervised learning methodology will require the reconstruction loss function, which is responsible for capturing the essential features of the context of the complete categorization process. The loss function used to train an undercomplete autoencoder is called reconstruction loss, as it is a check of how well the image has been reconstructed from the input [54, 61, 62]:

$$L_{rec}(x) = \| \widehat{M} \odot (x - F((1-\widehat{M})\odot x))\|_2^2, \tag{33}$$

and the adversarial loss which models the latent data entry space of the monoid morphisms in which the following is trained:

$$L_{adv} = \max_D E_{x\in X}[\log\ (D(x)) + \log\ (1 - D(F((1-\widehat{M})\odot x)))]. \tag{34}$$

Joint loss was used to implement the proposed template utilizing the combination of the above functions as follows:

$$L = \lambda_{rec}L_{rec} + \lambda_{adv}L_{adv}. \tag{35}$$

To develop representations encapsulating the underlying standard information across various regions of the data while rejecting low-level information and noise that is a local phenomenon, we use the Contrastive Predictive Coding technique [63–66]:

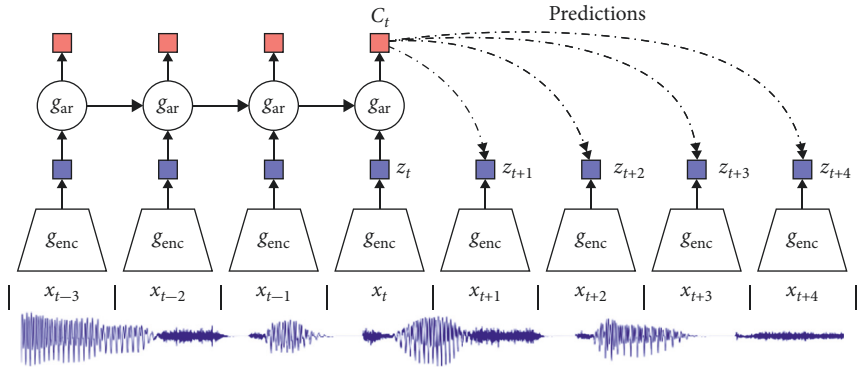

Figure 1: Contrastive Predictive Coding example (https://anilkeshwani.github.io/CPC/).

$$L_{\theta^a,\theta^+,\{\theta^-\}} = -\log \frac{\exp \exp\left(\theta^a \cdot \left(\theta^+/k\right)\right)}{\exp\left(\theta^a \cdot \left(\theta^+/k\right)\right) + \sum_{\theta^-} \exp\left(\theta^a \cdot \left(\theta^+/k\right)\right)}. \tag{36}$$

For example, given a lack of information, Figure 1 depicts the Contrastive Predictive Coding network, where "x" is a time series signal, data for which is available until time "t" and the model must predict the signal by the time "$t+4$." Here, "g_{enc}" is an integration network that extracts "z_t" attributes from the "x_t" signal and "g_{ar}" is a self-regression model that summarizes all the $z \leq t$ in the integration space to produce a latent representation of the environment $c_t = g_{ar}$ $(z \leq t)$ [67]. This composite representation is used to model a density ratio that maintains the mutual information between the predicted signal and the aggregate environment "c_t" [68–71].

Thus, in the proposed system, we combine future observation predictions with a likely loss linked to whether each monoid element is always symmetric [72].

So, for a 2-class (binary) classification problem, we have [69, 71] the following.

Activation function:

$$y = \sigma(a) \equiv \frac{1}{1 + \exp(-a)}. \tag{37}$$

Probability:

$$p(t \mid x, w) = \prod_{n=1}^{N} y(x_n, w)^{t_n} \{1 - y(x_n, w)\}^{1-t_n}. \tag{38}$$

Error function:

$$E(w) = -\sum_{n=1}^{N} \{t_n \ln y_n + (1 - t_n) \ln (1 - y_n)\}. \tag{39}$$

Moreover,

$$\frac{\partial E}{\partial a_k} = y_k - t_k. \tag{40}$$

However, for classification with k-classes (multiclass), we have the following.

Activation function:

$$y_k(x, w) = \frac{\exp \quad \exp(a_k(x, w))}{\sum_j \exp\left(a_j(x, w)\right)}. \tag{41}$$

Probability:

$$p(T \mid W) = \prod_{n=1}^{N} \prod_{k=1}^{K} y_{nk}^{t_{nk}}. \tag{42}$$

Error function:

$$E(w) = -\sum_{n=1}^{N} \sum_{k=1}^{K} t_{kn} \ln y_k(x_n, w). \tag{43}$$

Furthermore,

$$yy_k = a_k \Rightarrow \frac{\partial E}{\partial a_k} = y_k - t_k. \tag{44}$$

When an anomaly is run through the model, it will not recreate it since it is taught only to reproduce standard data, resulting in a considerable Mean Absolute Percentage Error (MAPE) [73, 74]:

$$\frac{100}{n} \sum_{i}^{n} \frac{y_i - \widehat{y}_i}{y_i}. \tag{45}$$

The comparison and the final categorization are achieved by defining a threshold value for MAPE, which is not sensitive to extreme values. At the same time, its values are normalized based on the actual observation, so it predicts the sample's class with high precision and recall [75].

This procedure may be repeated multiple times if it makes sense; that is, the reconstruction at each phase is adequate, implying that the new objectives are not too challenging. When a sample goes to a displaced region, it is always conceivable that it may end up in a zone with more opponents than previously. Furthermore, even if his aim puts him in a better position than before, there is a potential that the sample will be rebuilt in a worse situation. In these circumstances, repeating the operation for the troublesome pieces each time seems reasonable. That is, rediscover the problematic samples as they emerge from the categorizer's reconstruction of the details, to use the procedure again to locate the inverse of the function and to begin the process

from where the previous phase of categorizing had ended. So, if the reconstruction is good enough, the procedure can detach the network from local minimums and may be done numerous times. The suggested method's most significant novelty is the simple confirmation of the results of assigning classes to an unknown collection of values using quantifiable criteria [64, 76].

Finally, we have a reduction in data dimension, clear separation of classes, and self-adaptation with this method, as the proposed system learns to reconstruct the wrong samples in the supersphere defined by computational number theory and precisely the concept of the single object to perceive the nature of an unknown state based on generalized reshaped characteristics that come directly from the unknown environment.

4. Conclusions

Attempting to comment on the proposed system, it is a sophisticated practice that solves an essential problem of information systems security with great accuracy and reliability. With the proposed methodology currently presented and simulated by mathematical modeling, the artificial intelligence algorithm leads to a high learning rate, which is determined by how fast the industrial system converges. In general, self-adaptation and self-learning functionality enable identifying and maintaining fundamental characteristics of complex patterns that grow and contribute to the timely and accurate forecast of circumstances completely relevant to the industrial environment.

The proposed technique significantly strengthens the methodology because, in this problem of high complexity under consideration, the results of the prediction eliminate the variability, which is attributed to the sensitivity of industrial data. This complicated connection identifies and captures the minute distinctions that set them out amid the chaotic din. The suggested technique assures that the correlation of monoid operations can be anticipated using an intelligent correlation algorithm, efficiently using multiple learning cores and matching machine learning algorithmic structures with the single-mode process.

Furthermore, an additional benefit derived from the suggested function is that it provides better prediction and a more stable categorization rate since the general behavior of the model minimizes the overall probability of an awful decision that may be associated with occurrences such as this notion. This is because modern industrial data generators generate data in huge quantities and at high speed. The result is an increase in flow data. Extracting useful information from flow data is a challenge because its nature imposes constraints that cannot be satisfied by classical learning algorithms. Stream data is infinitely large, so it is not stored in memory, and each snapshot is usually and only accessible once. So, the snapshots are not available from the beginning as they arrive at a fast pace. Also, every snapshot is processed within a short time, and access to the actual price is limited. Most important, however, is the possibility of a change in the essential data production function, which is predicted with high reliability by the proposed system.

It is critical to stress that the quality of model adaption is interpreted as a percentage of "prediction self-improvement" owing to the higher rate of categorization accuracy fluctuation using this approach. The high percentages of accuracy reached after the general convergence of the reconstructed samples represent the temporal bias induced towards the dynamics of a model at a particular moment.

Another critical interpretation that emerges from the proposed algorithm's methodology is the characteristics of the relatively low rate of "mutation" in the changes that characterize the data shift, which allows the discovery of local extremes that may be included in a learning context, given the exploration of new areas of the multidimensional solution space. On the other hand, if the rate of "mutation" was too great, it might restrict the utilization of regions of high appropriateness in the solution space and imprison the system in nongeneralizable solutions.

The basic model is high speed, owing to the limits on the connections between the hidden and visible units that make it up, as mentioned above. Because of the algorithm hidden layer's function, where the teams of one level rely solely on branches of the other level, it also efficiently and precisely detects high-level correlations in data sets. Another significant feature of the proposed method is its ability of separating and rejecting random noise in the training set. The addition of automation to reclassifying complex data as a future extension of the proposed system is essential. This is the most realistic way of operating and using intelligent systems in the operational security of modern industrial infrastructures and systems.

Suggestions for future development and enhancements to this system should also concentrate on further improving the settings of the heuristic approach of redefining and rearranging the issue samples utilized to obtain an even more efficient, accurate, and quicker classification process. Finally, it is critical to investigate the expansion of this algorithm for the analysis and classification of real-time data presented in streams so that it can completely automate identifying even stealth zero-day attack types.[77]

Data Availability

The data can be obtained from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] D. Bhamare, M. Zolanvari, A. Erbad, R. Jain, K. Khan, and N. Meskin, "Cybersecurity for industrial control systems: a survey," *Computers & Security*, vol. 89, Article ID 101677, 2020.

[2] J. E. Rubio, C. Alcaraz, R. Roman, and J. Lopez, "Current cyber-defense trends in industrial control systems," *Computers & Security*, vol. 87, Article ID 101561, 2019.

[3] M. Kaouk, J.-M. Flaus, M.-L. Potet, and R. Groz, "A review of intrusion detection systems for industrial control systems," in *Proceedings of the 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, pp. 1699–1704, Paris, France, April 2019.

[4] E. Anthi, L. Williams, M. Rhode, P. Burnap, and A. Wedgbury, "Adversarial attacks on machine learning cybersecurity defences in Industrial Control Systems," *Journal of Information Security and Applications*, vol. 58, Article ID 102717, May 2021.

[5] K. Tsiknas, D. Taketzis, K. Demertzis, and C. Skianis, "Cyber threats to industrial IoT: a survey on attacks and countermeasures," *IoT*, vol. 2, no. 1, p. 1, 2021.

[6] I. Zerdazi and M. Fezari, "SCADA attack modeling using bond graph," in *Proceedings of the 2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*, pp. 1-2, Paris, France, September 2019.

[7] C. Kaura, N. Sindhwani, and A. Chaudhary, "Analysing the impact of cyber-threat to ICS and SCADA systems," in *Proceedings of the 2022 International Mobile and Embedded Technology Conference (MECON)*, pp. 466–470, March 2022.

[8] W. Li, L. Xie, D. Liu, and Z. Wang, "False logic attacks on SCADA control system," in *Proceedings of the 2014 Asia-Pacific Services Computing Conference*, pp. 136–140, Fuzhou, China, September 2014.

[9] A. Babay, T. Tantillo, T. Aron, M. Platania, and Y. Amir, "Network-attack-resilient intrusion-tolerant SCADA for the power grid," in *Proceedings of the 2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pp. 255–266, Luxembourg, Luxembourg, June 2018.

[10] R. Atat, L. Liu, H. Chen, J. Wu, H. Li, and Y. Yi, "Enabling cyber-physical communication in 5G cellular networks: challenges, spatial spectrum sensing, and cyber-security," *IET Cyber-Physical Systems: Theory & Applications*, vol. 2, no. 1, pp. 49–54, 2017.

[11] A. M. Kanca and Ş. SAĞIROĞLU, "Sharing cyber threat intelligence and collaboration," in *Proceedings of the 2021 International Conference on Information Security and Cryptology (ISCTURKEY)*, pp. 167–172, Ankara, Turkey, September 2021.

[12] M. A. Umer, K. N. Junejo, M. T. Jilani, and A. P. Mathur, "Machine learning for intrusion detection in industrial control systems: applications, challenges, and recommendations," *International Journal of Critical Infrastructure Protection*, Article ID 100516, 2022.

[13] H. Yang, L. Cheng, and M. C. Chuah, "Deep-learning-based network intrusion detection for SCADA systems," in *Proceedings of the 2019 IEEE Conference on Communications and Network Security (CNS)*, pp. 1–7, Washington, DC, USA, June 2019.

[14] F. A. Alhaidari and E. M. Al-Dahasi, "New approach to determine DDoS attack patterns on SCADA system using machine learning," in *Proceedings of the 2019 International Conference on Computer and Information Sciences (ICCIS)*, pp. 1–6, Sakaka, Saudi Arabia, Apil 2019.

[15] S.-P. Hong, C.-H. Lim, and H. J. Lee, "APT attack response system through AM-HIDS," in *Proceedings of the 2022 24th International Conference on Advanced Communication Technology (ICACT)*, pp. 271–274, PyeongChang Kwangwoon_Do, Republic of Korea, 2022.

[16] P. V. S. Charan, P. Mohan Anand, S. K. Shukla, N. Selvan, and H. Chunduri, "DOTMUG: a threat model for target specific APT attacks–misusing google teachable machine," in *Proceedings of the 2022 10th International Symposium on Digital Forensics and Security (ISDFS)*, pp. 1–8, Istanbul, Turkey, June 2022.

[17] Y. Xue, Q. Zhang, and F. Neri, "Self-Self-Adaptive Particle Swarm Optimization-Based Echo State Network for Time Series Prediction.daptive particle swarm optimization-based echo state network for time series prediction," *International Journal of Neural Systems*, vol. 31, no. 12, Article ID 2150057, 2021.

[18] S.-X. Lin, Z.-J. Li, T.-Y. Chen, and D.-J. Wu, "Attack tactic labeling for cyber threat hunting," in *Proceedings of the 2022 24th International Conference on Advanced Communication Technology (ICACT)*, pp. 34–39, PyeongChang Kwangwoon_Do, Republic of Korea, October 2022.

[19] L. Xing, K. Demertzis, and J. Yang, "Identifying data streams anomalies by evolving spiking restricted Boltzmann machines," *Neural Computing and Applications*, vol. 32, no. 11, pp. 6699–6713, 2020.

[20] K. Demertzis, L. Iliadis, E. Pimenidis, and P. Kikiras, "Variational restricted Boltzmann machines to automated anomaly detection," *Neural Computing and Applications*, vol. 34, 2022.

[21] K. Demertzis, D. Taketzis, V. Demertzi, and C. Skianis, "An Ensemble Transfer Learning Spiking Immune System for Adaptive Smart Grid Protectionnsemble transfer learning spiking immune system for adaptive smart grid protection," *Energies*, vol. 15, no. 12, p. 4398, 2022.

[22] J. Guo and Y. Shen, "Online Online Anomaly Detection of Industrial IoT Based on Hybrid Machine Learning Architecturenomaly detection of industrial IoT based on hybrid machine learning architecture," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8568917, 10 pages, 2022.

[23] X. Zheng and X. Yin, "A A Privacy-Preserved Variational-Autoencoder for DGA Identification in the Education Industry and Distance Learningrivacy-preserved variational-autoencoder for DGA identification in the education industry and distance learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7384803, 8 pages, 2022.

[24] W. Jiang, "Machine Learning Methods to Detect Voltage Glitch Attacks on IoT/IIoT Infrastructuresearning methods to detect voltage glitch attacks on IoT/IIoT infrastructures," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6044071, 7 pages, 2022.

[25] Y. Hu, A. Yang, H. Li, Y. Sun, and L. Sun, "A survey of intrusion detection on industrial control systems," *International Journal of Distributed Sensor Networks*, vol. 14, no. 8, Article ID 155014771879461, 2018.

[26] A. Facchini, "Semilocal categories and modules with semilocal endomorphism rings," 2022, https://link.springer.com/book/10.1007/978-3-030-23284-9.

[27] F. Zhao, H. Zhang, J. Peng, X. Zhuang, and S.-G. Na, "A semi-self-taught network intrusion detection system," *Neural Computing and Applications*, vol. 32, no. 23, pp. 17169–17179, 2020.

[28] S. Algarni, F Eassa, K Almarhabi et al., "Blockchain-Blockchain-Based Secured Access Control in an IoT Systemased secured access control in an IoT system," *ApplApplied Sciences*, vol. 11, no. 4, p. 1772, 2021.

[29] X. Li, "A A Blockchain-Based Verifiable User Data Access Control Policy for Secured Cloud Data Storagelockchain-based verifiable user data access control policy for secured

cloud data storage," *Computational Intelligence and Neuroscience*, vol. 2022, pp. Apr–12, Article ID 2254411, 2022.

[30] M. R. Naeem, R. Amin, S. S. Alshamrani, and A. Alshehri, "Digital Digital Forensics for Malware Classification: An Approach for Binary Code to Pixel Vector Transitionorensics for malware classification: an approach for binary code to pixel vector transition," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6294058, 12 pages, 2022.

[31] Z. Ma, J. Li, Y. Song, X. Wu, and C. Chen, "Network Network Intrusion Detection Method Based on FCWGAN and BiLSTMntrusion detection method based on FCWGAN and BiLSTM," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6591140, 17 pages, 2022.

[32] W. Zhang, Y Zhang, C Guo et al., "Certificateless Certificateless Hybrid Signcryption by a Novel Protocol Applied to Internet of Thingsybrid signcryption by a novel protocol applied to internet of things," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3687332, 7 pages, 2022.

[33] W. Jiang, "A A Machine Vision Anomaly Detection System to Industry 4.0 Based on Variational Fuzzy Autoencoderachine vision anomaly detection system to industry 4.0 based on variational fuzzy autoencoder," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1945507, 10 pages, 2022.

[34] D. Li, J. Wang, Z. Tan, X. Li, and Y. Hu, "Differential privacy preservation in interpretable feedforward-designed convolutional neural networks," in *Proceedings of the 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, pp. 631–638, Guangzhou, China, September 2020.

[35] M. A. Albahar, M. S. ElSayed, and A. Jurcut, "A A Modified ResNeXt for Android Malware Identification and Classificationodified ResNeXt for android malware identification and classification," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8634784, 20 pages, May 2022.

[36] Y. Dandan, V. Gould, M. Hartmann, N. Ruškuc, and R.-E. Zenab, "Coherency and Coherency and Constructions for Monoidsonstructions for monoids," *The Quarterly Journal of Mathematics*, vol. 71, no. 4, pp. 1461–1488, 2020.

[37] H. Machida, "Centralizing monoids, majority operations and the słupecki clone," in *Proceedings of the 2022 IEEE 52nd International Symposium on Multiple-Valued Logic (ISMVL)*, pp. 62–67, Dallas, TX, USA, February 2022.

[38] M. Behrisch, "Centralising monoids with conservative majority operations as witnesses," in *Proceedings of the 2021 IEEE 51st International Symposium on Multiple-Valued Logic (ISMVL)*, pp. 56–61, Nur-sultan, Kazakhstan, February 2021.

[39] T. Ishida and S. Inokuchi, "Commutativity of composition of some elementary cellular automata on monoids," in *Proceedings of the 2020 Eighth International Symposium on Computing and Networking (CANDAR)*, pp. 128–133, Naha, Japan, August 2020.

[40] H. Machida and I. G. Rosenberg, "Centralizing monoids on a three-element set related to binary idempotent functions," in *Proceedings of the 2016 IEEE 46th International Symposium on Multiple-Valued Logic (ISMVL)*, pp. 84–89, Sapporo, Japan, February 2016.

[41] H. Machida and I. G. Rosenberg, "Centralizing monoids and the arity of witnesses," in *Proceedings of the 2017 IEEE 47th International Symposium on Multiple-Valued Logic (ISMVL)*, pp. 236–241, Novi Sad, Serbia, February 2017.

[42] M. Behrisch and V. García, "Centralising Monoids with Low-Arity Witnesses on a Four-Element Set," vol. 13, no. 8, p. 1471, 2022, https://www.mdpi.com/2073-8994/13/8/1471.

[43] A. Hassanpour, M. Moradikia, H. Adeli, S. R. Khayami, and P. Shamsinejadbabaki, "A novel end-to-end deep learning scheme for classifying multi-class motor imagery electroencephalography signals," *Expert Systems*, vol. 36, no. 6, Article ID e12494, 2019.

[44] W. Ruan, J. Zhao, and X. Bai, "Block backtracking-based matching pursuit for arbitrary block sparse signal recovery," in *Proceedings of the 2018 International Conference on Signals and Systems (ICSigSys)*, pp. 209–212, Bali, Indonesia, February 2018.

[45] K. Yan, H.-C. Wu, H. Xiao, and X. Zhang, "Novel robust band-limited signal detection approach using graphs," *IEEE Communications Letters*, vol. 21, no. 1, pp. 20–23, 2017.

[46] S. T. Ali, K. Goyal, and J. Singhai, "Moving object detection using self adaptive Gaussian Mixture Model for real time applications," in *Proceedings of the 2017 International Conference on Recent Innovations in Signal processing and Embedded Systems (RISE)*, pp. 153–156, Bhopal, India, July 2017.

[47] T. Jin, H. Wang, and H. Liu, "Design of a flexible high-performance real-time SAR signal processing system," in *Proceedings of the 2016 IEEE 13th International Conference on Signal Processing (ICSP)*, pp. 513–517, Chengdu, China, August 2016.

[48] Y. Chang, Q. Wan, C. Xia, and Y. Wan, "A method of fast extract signal subspace based on the householder transformation," in *Proceedings of the 2018 International Computers, Signals and Systems Conference (ICOMSSC)*, pp. 378–381, Dalian, China, September 2018.

[49] J. E. van Engelen and H. H. Hoos, "A survey on semi-supervised learning," *Machine Learning*, vol. 109, no. 2, pp. 373–440, 2020.

[50] Y. Ouali, C. Hudelot, and M. Tami, *An Overview of Deep Semi-Supervised Learning*, 2020, https://arxiv.org/abs/2103.00550.

[51] Y.-F. Li and D.-M. Liang, "Safe semi-supervised learning: a brief introduction," *Frontiers of Computer Science*, vol. 13, no. 4, pp. 669–676, 2019.

[52] Q. Liu, X. Liao, and L. Carin, "Semi-supervised life-long learning with application to sensing," in *Proceedings of the 2007 2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pp. 1–4, St. Thomas, VI, USA, September 2007.

[53] P. K. Mallapragada, R. Jin, A. K. Jain, and Y. Liu, "SemiBoost: SemiBoost: Boosting for Semi-Supervised Learningoosting for semi-supervised learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, 2009.

[54] M. Soliman, C. Lehman, and G. AlRegib, "S6: semi-supervised self-supervised semantic segmentation," in *Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP)*, pp. 1861–1865, Abu Dhabi, United Arab Emirates, July 2020.

[55] C. Wei, C. Guo, and W. Yan, "Forest fire risk forecast method with pseudo label based on semi-supervised learning," in *Proceedings of the 2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, pp. 36–39, Taiyuan, China, September 2021.

[56] J. G. Altonji and R. L. Matzkin, "Cross section and panel data estimators for nonseparable models with endogenous regressors," *Econometrica*, vol. 73, no. 4, pp. 1053–1102, 2005.

[57] "Co-simulation study on vibration control of multistage gear transmission system based on multiple control algorithms | IEEE Conference Publication | IEEE Xplore," https://ieeexplore.ieee.org/abstract/document/8316474.

[58] K. Demertzis, L. S. Iliadis, and V.-D. Anezakis, "Extreme deep learning in biosecurity: the case of machine hearing for

marine species identification," *J. Inf. Telecommun.*vol. 2, no. 4, pp. 492–510, 2018.

[59] A. J. Kulkarni, I. P. Durugkar, and M. Kumar, "Cohort Cohort Intelligence: A Self Supervised Learning Behaviorntelligence: a self supervised learning behavior," in *Proceedings of the 2013 2013 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1396–1400, Manchester, UK, July 2013.

[60] K. M. O. Vale, A. C. Gorgônio, F. D. L. E. Gorgônio, and A. M. D. P. Canuto, "An An Efficient Approach to Select Instances in Self-Training and Co-Training Semi-Supervised Methodsfficient approach to select instances in self-training and Co-training semi-supervised methods," *IEEE Access*, vol. 10, pp. 7254–7276, 2022.

[61] L. Song and W. Luo, "Self-supervised learning of visual odometry," in *Proceedings of the 2020 International Conference on Information Science, Parallel and Distributed Systems (ISPDS)*, pp. 5–9, Xi'an, China, December 2020.

[62] L. Zhou, X. Ling, S. Zhu, Z. Sun, and J. Yang, "An self-supervised learning & self-attention based method for defects classification on PCB surface images," in *Proceedings of the 2021 2nd International Conference on Electronics, Communications and Information Technology (CECIT)*, pp. 229–234, Sanya, China, September 2021.

[63] L. Chen, X. Liang, Y. Feng, L. Zhang, J. Yang, and Z. Liu, "Online Online Intention Recognition With Incomplete Information Based on a Weighted Contrastive Predictive Coding Model in Wargame.ntention recognition with incomplete information based on a weighted contrastive predictive coding model in wargame," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–14, 2022.

[64] J. Ebbers, M. Kuhlmann, T. Cord-Landwehr, and R. Haeb-Umbach, "Contrastive predictive coding supported factorized variational autoencoder for unsupervised learning of disentangled speech representations," in *Proceedings of the ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3860–3864, Toronto, Canada, June 2021.

[65] X. Zhu, H. Dong, P. S. Rossi, and M. Landrø, "Time-Time-Frequency Fused Underwater Acoustic Source Localization Based on Contrastive Predictive Codingrequency fused underwater acoustic source localization based on contrastive predictive coding," *IEEE Sensors Journal*, vol. 22, no. 13, pp. 13299–13308, 2022.

[66] R. Qiu, Z. Huang, and H. Yin, "Memory Memory Augmented Multi-Instance Contrastive Predictive Coding for Sequential Recommendationugmented multi-instance contrastive predictive coding for sequential recommendation," in *Proceedings of the 2021 2021 IEEE International Conference on Data Mining (ICDM)*, pp. 519–528, Auckland, New Zealand, September 2021.

[67] X. Zhu, H. Dong, P. S. Rossi, and M. Landrø, "Self-supervised Self-supervised Underwater Source Localization based on Contrastive Predictive Codingnderwater source localization based on contrastive predictive coding," in *Proceedings of the 2021 IEEE Sensors*, pp. 1–4, Sydney, Australia, July 2021.

[68] Y. Chen, J. Zhao, W. Wang et al., "SEQ-CPC: sequential contrastive predictive coding for automatic speech recognition," in *Proceedings of the ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3880–3884, Toronto, Canada, June 2021.

[69] C. Wan, T. Zhang, Z. Xiong, and H. Ye, "Representation learning for fault diagnosis with contrastive predictive coding," in *Proceedings of the 2021 CAA Symposium on Fault Detection, Supervision, and Safety for Technical Processes (SAFEPROCESS)*, pp. 1–5, Chengdu, China, September 2021.

[70] J. Nistal, C. Aouameur, S. Lattner, and G. Richard, "VQCPC-GAN: variable-length Adversarial audio synthesis using vector-quantized contrastive predictive coding," in *Proceedings of the 2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pp. 116–120, Paltz, NY, USA, July 2021.

[71] A. van den Oord, Y. Li, and O. Vinyals, *Representation Learning with Contrastive Predictive Coding*, 2019, https://arxiv.org/abs/1807.03748.

[72] E. Anthi, L. Williams, A. Javed, and P. Burnap, "Hardening Machine Learning Denial of Service (DoS) Defences against Adversarial Attacks in IoT Smart home Networks - ScienceDirect,"vol. 108 https://www.sciencedirect.com/science/article/pii/S0167404821001760002520.

[73] J. Gui, Z. Sun, Y. Wen, D. Tao, and J. Ye, "A review on generative adversarial networks: algorithms, theory, and applications," 2020, http://arxiv.org/abs/2001.06937.

[74] Handbook of Statistics, *Bayesian Thinking, Modeling and Computation - PDF Free Download*, vol. 25 https://epdf.tips/handbook-of-statistics-volume-25-bayesian-thinking-modeling-and-computation.html.

[75] K. Demertzis, L. S. Iliadis, and V. Anezakis, *A Dynamic Ensemble Learning Framework for Data Stream Analysis and Real-Time Threat Detection*, https://www.springerprofessional.de/en/a-dynamic-ensemble-learning-framework-for-data-stream-analysis-a/16154694.

[76] S. V. Boštjančič Rakas and M. D. Stojanović, "A Review of Research Work on Network-Based SCADA Intrusion Detection Systems," *IEEE Journals & Magazine | IEEE Xplore*, vol. 8 https://ieeexplore.ieee.org/author/37086663933https://doi.org/10.1109/ACCESS.2020.2994961https://ieeexplore.ieee.org/document/9094250.

[77] A. Ayodeji, Y. k. Liu, N. Chao, and L. q. Yang, "A new perspective towards the development of robust data-driven intrusion detection for industrial control systems," *Nuclear Engineering and Technology*, vol. 52, no. 12, pp. 2687–2698, 2020.