Federated-SRUs: A Federated-Simple-Recurrent-Units-Based IDS for Accurate Detection of Cyber Attacks Against IoT-Augmented Industrial Control Systems

Izhar Ahmed Khan, Dechang Pi, Muhammad Zahid Abbas, Umar Zia, Yasir Hussain, and Hatem Soliman

***Abstract*—The security of industrial control systems (ICSs) against cyber-attacks is essential in modern era since ICSs are vital constituent of modern societies and smart cities. However, the augmentation of legacy ICS networks with smart computing and networking technologies [such as Internet of Things (IoT)] has intensely enlarged the surface of attacks against these critical infrastructures. This augmentation makes these networks more vulnerable to cyber-attacks and despite the current security solutions, attackers still find ways to proliferate these networks. The intrusion detection system (IDS) is one of the key security aspect to prevent these networks from contemporary cyber-attacks. Therefore, this article proposes a new IDS model named federated-simple recurrent units (SRUs) for the security of IoT-based ICSs. Specifically, the federated-SRUs IDS model uses an improved simple recurrent units architecture to reduce computational cost and alleviate the gradient vanishing issue in recurrent networks. Then, it performs data aggregation through several communication rounds in the federated architecture which allows multiple ICS networks and stakeholders to build a comprehensive IDS model in a privacy-preserving manner. The performance of the federated-SRUs IDS model is validated through experiments using real-world gas pipeline-based ICS network data, which indicates that it is able to accurately detect intrusions in real time without compromising privacy and security. Experiments also verify that the federated-SRUs model outperforms existing state-of-the-art approaches and thus can serve as a viable IDS method in IoT-based ICS networks.**

***Index Terms*—Cyber-attacks, industrial control systems (ICSs), industrial networks, intrusion detection, Internet of Things (IoT).**

Manuscript received 19 March 2022; revised 5 June 2022 and 12 July 2022; accepted 15 August 2022. Date of publication 19 August 2022; date of current version 9 May 2023. This work was supported by the National Science and Technology Innovation 2030 through the Key Project of "New Generation Artificial Intelligence" under Grant 2021ZD0113103. *(Corresponding authors: Dechang Pi; Izhar Ahmed Khan.)*

Izhar Ahmed Khan, Dechang Pi, Yasir Hussain, and Hatem Soliman are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: izhar@nuaa.edu.cn; dc.pi@nuaa.edu.cn; yaxirhuxxain@nuaa.edu.cn; hatem@nuaa.edu.cn).

Muhammad Zahid Abbas is with the Department of Computer Science, COMSATS University Islamabad (Vehari Campus), Vehari 61100, Pakistan (e-mail: zahidabbas@cuivehari.edu.pk).

Umar Zia is with the Department of Computer Science, COMSATS University Islamabad (Attock Campus), Attock 43600, Pakistan (e-mail: muftiumar@gmail.com).

Digital Object Identifier 10.1109/JIOT.2022.3200048

I. Introduction

INDUSTRIAL control systems (ICSs) comprises of controllers, tools, devices, and distributed networks, to facilitate the development of automating industrial processes. These kinds of systems are generally governed by the distributed control systems (DCSs) or supervisory control and data acquisition (SCADA) systems. The ICSs are extensively utilized in diverse critical infrastructures, for example, gas pipeline systems, manufacturing plants, water systems, and power grids [1]. With the rise of Industry 4.0 concept in recent years, the incorporation of traditional ICSs with emerging technologies, such as Internet of Things (IoT) or Industrial IoT, artificial intelligence (AI), and software defined networking (SDN), the architecture of which is displayed in Fig. 1, has facilitated the industries and corporations to take advantage of different services, such as remote control, and prompt monitoring of these critical infrastructure networks. Since ICSs are vital elements of smart cities, they have a huge influence on the broader society which, if interrupted, could lead to overwhelming outcomes [2].

However, the integration of IoT and ICSs has intensified the attack surfaces and the risk of malicious activities and cyber-attacks. Hackers and state sponsored actors are continuously targeting the networks of ICSs in the past years through cyber-attacks. Which is why the cybersecurity and infrastructure security agency (CISA) [3] announced their report on "One CISA" initiative, with the aim to persuade the operators and owners of ICSs and critical infrastructures to develop security capabilities that directly authorize the stakeholders of ICS networks to protect their systems against ICS threats.

Usually, the transmission or communication between the components of ICSs is founded on the principles of remote connectivity and current information technology (IT) stack. This communication dependency on the transmission networks might upsurge the likelihood of deliberate cyber-attacks against physical infrastructures. Traditionally, the traffic of communication network is protected through the usage of different security techniques for example, encryption, authentication, and data integrity methods. However, these techniques are unable to entirely defend the complete stages of ICSs against varied kinds of malicious actions. Various attacks in the past years such as *BlackEnergy* [4], *Duqu* [5], *Flame* [6],

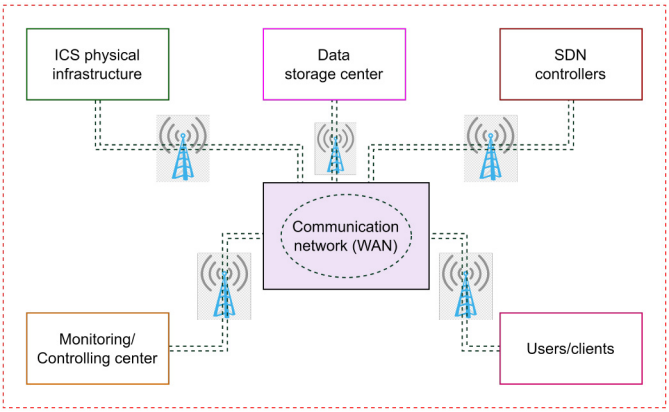


Fig. 1. General architecture of ICSs.

and *Seismic* threats [7] targeted ICS networks and proved to be of great danger by causing excessive damage to common informations (CIs) in many countries.

The intrusion detection system (IDS) has been broadly explored in the recent years for the protection of ICSs against cyber-attacks. Numerous researchers have proposed security methods based on the IDS model using recurrent neural network (RNN). However, these algorithms suffer from: 1) due to their shallow architecture nature, they are unable to correctly detect complex features in the instances of minority class; 2) as the layers in the neural network increases, the performance of the model is degraded due to the vanishing gradient problem, thus leading to low accuracy scores; and 3) RNN variants, such as GRU and LSTM are difficult to scale since the state calculations are dependent on time (i.e., the calculation of every next phase is postponed till the entire completing of the preceding phase). Hence, the training process becomes slow limiting the parallelizability of the model.

Therefore, this article proposes a decentralized architecture for the detection of intrusions from ICS networks. The phenomenon of simple recurrent unit (SRU) is improved to resolve issues stated above, i.e., skip connections are utilized to mitigate the disappearing gradient problem, and the bidirectional architecture is applied to expands the performance of IDS model in ICS environments. The main contributions of this study are as follows.

1) An efficient IDS model is proposed for the privacy and security of critical infrastructures, such as ICSs. This IDS model is vastly effective in spotting numerous kinds of cyber-attacks against ICS networks, for example, DoS attacks, command and response injection threats, and reconnaissance attacks.
2) A federated learning-based distributed model is developed, which, on one hand, supports the processing of data at the premises of each ICS networks, enabling the privacy of the data and ICS networks to be preserved. On other hand, the federated model empowers multiple owners/stakeholders of ICS networks to build a comprehensive IDS model for detection of evolving cyber-attack vectors.
3) Comprehensive experiments using real-scale data from the ICS network indicates that the proposed federated-model performed better than state-of-the-art approaches. Furthermore, this proposed model is proved to be robust against imbalanced data.

The remaining parts of this article are organized as follows. Section II discusses the background and related literature. Section III discusses the proposed IDS model, while Section IV presents the evaluation framework and results. Last, we conclude our study in Section V.

II. Background and Literature Review

This section conveys a brief background of SCADA systems and ICSs, as well as related literature concerning intrusion detection. These studies are briefly summarized and compared in Table I to highlight their main characteristics and limitations.

A. SCADA Systems in ICS Networks

Generally speaking, SCADA operates as a system that controls automation and consists of several core functions, such as controlling of physical apparatuses, configuration, and monitoring of processes, such as pump pressure [2], [8] and power relays [9]. It utilizes mechanisms based on computerized methods to control the operations and functioning of the network, such as human–machine interface (HMI). In the context of ICS networks, the responsibility of SCADA technology is to send control commands and govern the process of collecting measurement data through the sensors and actuators. The main modules used in this governing process involves the remote terminal units (RTUs), and the programmable logic controllers (PLCs), that assist as a processing gateway for the attained data.

Existing networks based on the SCADA technology are capable of network connectivity functions through the Ethernet protocols, e.g., *DNP*3 and *IEC*61850, which are accustomed to function using the Modbus TCP/IP technique. The interaction and communication among the IT, subsystems, and SCADA networks are presented in Fig. 2. ICSs within SCADA networks become more exposed to vulnerabilities and cyber-attacks due to their Internet connectivity and remote monitoring of the IoT devices. To expand the security of these networks, IDSs models have been widely employed (as discussed in next section) to cope with the evolving cyber-attack scenarios.

B. Intrusion Detection in ICS and IIoT Networks

IDSs have been widely adapted for investigating the network data traffic of ICSs and IoT networks to identify intrusive activities. For example, Gu *et al.* [10] developed an IDS model named *DEIDS* for the detection of intrusive actions in ICS networks. They proposed an optimized SMOTE method to label the data set and utilized the CNN architecture which managed to achieved 97% accuracy rate. Although this accuracy rate is fairly acceptable but the detection rate (recall) is less than accuracy. Similarly, Wang *et al.* [11] proposed the transfer learning-based model for the detection of abnormalities in ICSs. They utilized the Adaboost architecture to develop their model and claimed that their model is able to learn using few abnormal data. Although they managed to achieved good accuracy rates but they did not discuss the potential of their work in IoT or IIoT perspective.

To develop an IDS model for gas pipeline-based ICSs, the writers of [12] developed an RNN-based technique to

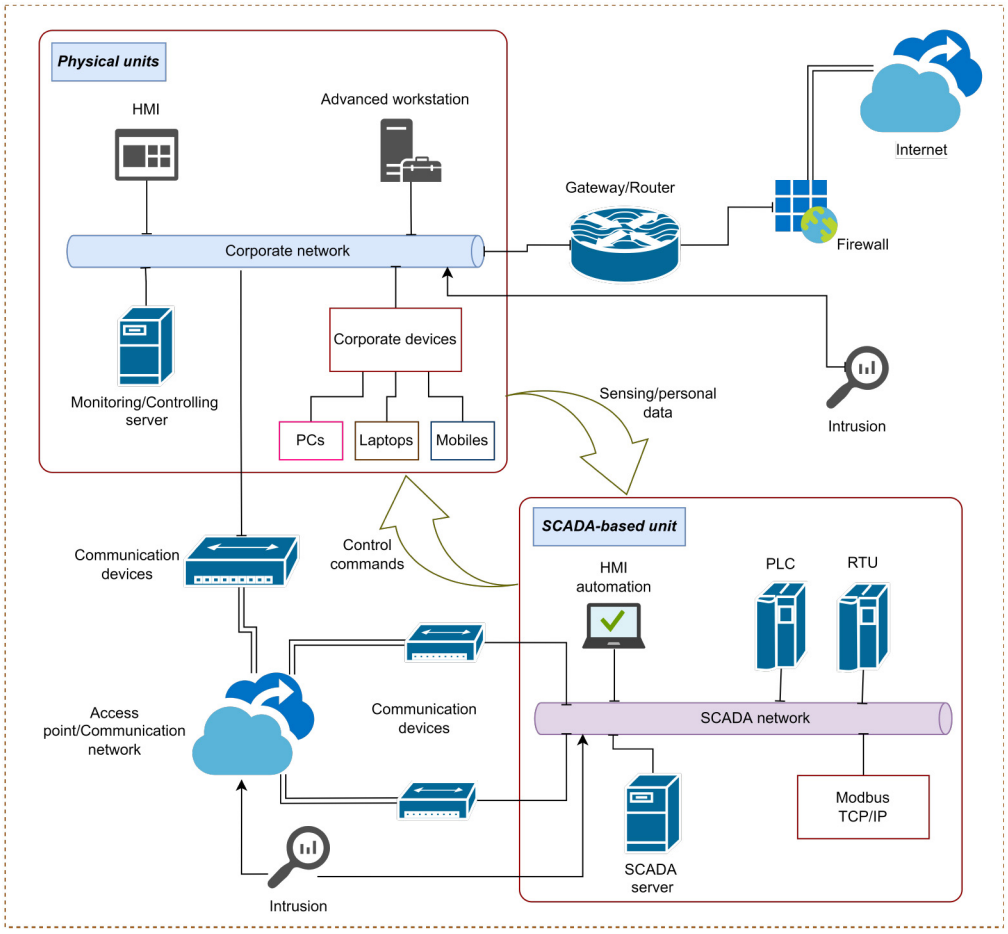


Fig. 2. General architecture of SCADA in ICSs.

TABLE I
SUMMARY AND COMPARISON OF REVIEWED WORKS

Domain	Distinctive characteristics	Limitations	Reference
ICS	Achieved 97% accuracy rate	Detection rate (recall) is less	[10]
ICS	Achieved good accuracy rate	Might not applicable in IoT or IIoT networks	[11]
ICS	Achieved good accuracy rate	False positive rate is high	[12]
IIoT	Better detection of anomalies in smart factory environment	No detection results provided	[13]
IIoT	Achieved good accuracy rate	Only focuses on accuracy	[14]
IIoT	System throughput is high	Only focuses on privacy issues	[15]
IIoT/CPS	Improved data sharing architecture	Low detection rates	[16]
ICS/CPS	Better accuracy rate	Low detection rate	[17]
ICS	Lightweight model	Low accuracy rate	[18]
ICS/IIoT	Can save upto 35% of communication bandwidth	Limited data size restriction	[19]
IIoT	Good accuracy rate	Does not focuses on ICS networks	[20]
IIoT	Offers improved tensor mining approach	Does not cover wide range of cyber-attacks	[21]
IoT	Hybrid system of intrusion detection	Evaluated using non-ICS data	[22]
IoT	Plug and play system for intrusion detection	Not effective against adversarial attacks	[23]
IIoT/IICS	Achieved a good balance between accuracy and false alarms	Centralized learning (Non-federated)	[24]

detect intrusions in ICSs. They employed the SRU architecture to develop their model and report good accuracy rates. However, their false positive rate is high which makes it the least choice for security administrators to utilize it against contemporary cyber-attacks. Khan *et al.* [1] proposed the explainable cyber threat discovery model using the deep learning architecture identify intrusive actions. Specifically, they applied the autoencoder-based method using the sliding window technique to detect intrusions in ICS time-series data. Similarly, the writers of [13] explored the graph neural networks-based technique for the detection of anomalies in diverse IIoT-enabled ICSs. They discussed the potential implementation of their proposed model in several IIoT network scenarios, such as smart factory, smart transportation, and smart energy systems. The writers of [14] proposed the transfer learning method for reducing training time in IIoT systems. They categorize the application of their model in different scenarios, such as centralized and distributed learning, and compared it with the CNN architecture.

C. Federated-Based IDSs in ICSs and IIoT Environments

The implementation of federated learning-based models in IIoT has been an emerging area and has been of great interest to both industries and researchers. For example, Wang *et al.* [15] proposed a federated learning-based model for the detection of anomalies in IIoT networks. They used the deep reinforcement learning technique to train the local models. Similarly, the writers of [16] proposed a data sharing architecture using the federated learning model to prevent data leakage in IIoT networks and cyber physical systems (CPSs). To ensure the constraints regarding privacy, they adopted the paillier and partially homomorphic encryption technique. In a similar effort to secure CPSs, the writers of [17] proposed a method named DeepFed to identify cyber-threats. They developed their model using the convolutional RNN architecture and test it on ICS network data. Their experiments achieved good results but the detection rate is low.

To apply the phenomenon of federated learning in securing the ICS networks Jahromi *et al.* [18] proposed a deep learning-based cyber-attack detection model. Although the study is good but they managed to achieved only 90.83% rate of accuracy. Similarly, Huong *et al.* [19] proposed an anomaly detection-based federated learning architecture for securing ICS networks. They implement their model in edge computing environment and claimed that their model can save 35% of communication bandwidth during the transmission between cloud and the edge. In an effort to secure the communication networks of IoT-based supply chain 4.0 systems, Khan *et al.* [20] proposed a federated IDS model that is capable of detecting several kinds of cyber-attacks. They protect the IoT-based supply chain 4.0 systems using the RNN architecture and compared it with centralized training models.

III. Proposed IDS Model

This section discusses the essential steps for the development of the proposed IDS model by presenting the data preparation, model development, and implementing the federated architecture for detecting intrusions.

A. Overview of System Model

This article proposed a deep federated learning architecture to solve the privacy and security issues of IoT-augmented ICS networks. The main steps of the proposed model are presented in Fig. 3, which mainly contains two kinds of components, that are., a global server, and a local server for each ICS network.

1) *Global Server:* The responsibility of the global server (G-server) to build a comprehensive and wide-ranging model for detecting intrusive events by federating the parameters of the model transmit by the local servers at the premise of each ICS network. To obtain an optimum "perfect" IDS model, this step entails multiple communication rounds between the G-server and every participating ICS network.
2) *Local ICS Server:* Every local server (L-server) at each ICS network, on behalf of the stakeholder of the ICS network, is responsible to build a local IDS model based on the data collected from its own IoT-based ICS network and assisting the process of parameter updation by repeatedly communicating with the G-server.

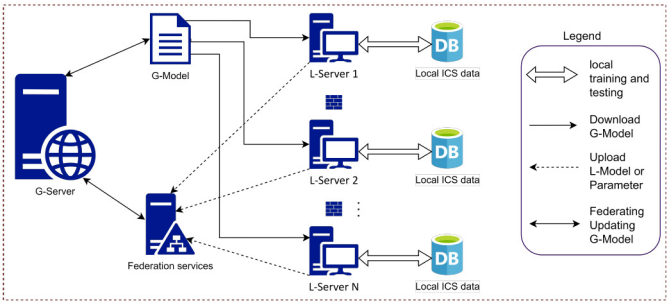


Fig. 3. Proposed federated system model.

B. Data Cleaning

Since ICS and IIoT networks generates enormous quantity of data, it becomes challenging to handle raw data because of power and resource constraints IoT devices. Therefore, extracted raw data from network traffic needs to be preprocessed and cleaned first to make it useful for the model to learn from it in a better and faster manner. The proposed IDS model performs certain preprocessing steps, such as normalization, and feature reduction at this stage with the focus to optimize the features which are briefly defined as follows.

1) *Normalization:* As the proposed IDS model is based on the deep learning architecture which rely on weights, the diverse values of features could be influence by the bias data into specific layers which might cause some weights to upgraded quicker than others. Therefore, it is essential to resolve this problem through the usage of a statistical normalization technique, i.e., the *Z-score* method, which is given as

$$ZS^{(i)} = \frac{fv^{(i)} - \mu}{\sigma} \tag{1}$$

where the mean of n values for a specific feature $(fv^{(i)}(i \in 1, 2, \ldots, n))$ is represented by μ, and the standard deviation is represented by the σ.

2) *Feature Reduction:* Since IIoT and ICS networks generates high-dimensional data, the process of feature reduction is applied with the aim to remove irrelevant and noisy features from the data to improve the processing process by reducing the computational cost and designing a scalable and the lightweight IDS model. Although reducing the features impacts accuracy, but it makes the trained model runs faster and computationally less expensive. As we are aiming to develop an IDS for IoT-based networks (which are resource restricted devices in nature) it is imperative to reduce the number of features because of several benefits, such as: a) fewer computational resources required; b) less memory requirements; and c) less execution time. Therefore, the proposed model uses an ICA method for the feature reduction process, and is given as

$$fs = Av + n \tag{2}$$

where the m-dimensional feature sample is represented by fs, the independent components in n-dimension are represented by the vector v, the constant for the mixing matrix $m \times n$ having $m \geq n$ and n as noise, is represented by A. Assuming that the network data is labeled and captured in noise-free scenario, then the ICA technique can be given as

$$fs = Av \tag{3}$$

and

$$v = Ufs \tag{4}$$

where the mapping function or the unmixing matrix is denoted by U, for projecting fs to v. These independent components are deliberated as the best representation of the data since the ICS technique can offer the higher approximations of A and v given fs having the non-Gaussian data (which is the norm of ICS network data) maximizing constraint. Then, the non-Gaussian data maximizing constraint to diminish the CI between n variables (v_i, where $i = 1, 2, \ldots, n$), is expressed in the following way to solve (4)

$$\mathrm{CI}(v_1, v_2, v_3, \ldots, v_m) = \sum_i D(v_i) - D(v_d) \tag{5}$$

where the differential entropy is represented by D.

C. Enhanced Simple Recurrent Units for Attack Classification

The SRU variant of RNN is intended to simplify the training process of deep learning algorithms in a parallel manner [25]. The basic idea behind the SRU is to ease the shortcomings of other RNN variants, such as LSTM and GRU. The core improvements of SRUs as compared to other two variants are that time dependency phenomenon is extinguished; and computations are performed in parallelized manner to accelerate the model training. The core structure of SRU contains two gates: 1) the forget gate (forg) and 2) the memory unit. Where the *forg* is responsible to specifies the status of the earlier step to the present state, by adjusting the memory unit. Whereas, these units of memory are responsible to compute the final output state. This computation process of single SRU layer can be expressed as

$$\tilde{\mathrm{in}}_{ts} = \mathrm{wepar}_{\mathrm{in}}\ \mathrm{in}_{ts} \tag{6}$$

$$\mathrm{forg}_{ts} = \sigma\left(\mathrm{wepar}_{\mathrm{forg}}\ \mathrm{in}_{ts} + bsv_{\mathrm{forg}}\right) \tag{7}$$

$$\mathrm{memu}_{ts} = \mathrm{forg}_{ts} \odot \mathrm{memu}_{ts-1} + \left(1 - \mathrm{forg}_{ts}\right) \odot \tilde{\mathrm{in}}_{ts} \tag{8}$$

$$\mathrm{outp}_{ts} = \mathrm{actfun}(\mathrm{memu}_{ts}) \tag{9}$$

where in_{ts} denotes the input, ts denotes the time-stamp, *wepar* denotes the weight parameters, *bsv* denotes the bias vector, and σ and *actfun* denote the activation functions. Whereas, the $\tilde{\mathrm{in}}_{ts}$ in (6) denotes impermanent form, the forg_{ts} in (7) denotes the forget gate (responsible to specify the significance of earlier phase to the present phase), the memu_{ts} in (8) denotes memory unit, and outp_{ts} in (9) denotes the networks' final output.

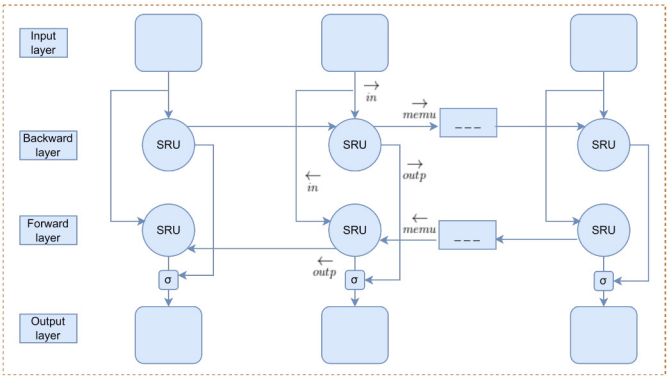


Fig. 4. Schematic architecture of bidirectional SRU.

As realized from the calculations stated above, the transformation between input and gate control unit only rest on the present input step, thereby, causing the matrix operations having vast calculation process to be parallely handled. Although the determining stage of memu_{ts} is still contingent on the former time-step, the measurement of memu_{ts} and outp_{ts} in this SRU method requires less cost of computation since it only encompasses point multiplication.

To alleviate the gradient vanishing problem, this work utilizes skip connections at the final step of the network. First, the res_{ts} (reset gate) is set as

$$\mathrm{res}_{ts} = \sigma\left(\mathrm{wepar}_{\mathrm{res}}\ \mathrm{in}_{ts} + bsv_{\mathrm{res}}\right). \tag{10}$$

Then, applying the skip connections to calculate outp_{ts} as

$$\mathrm{outp}_{ts} = \mathrm{res}_{ts} \odot \mathrm{actfun}(\mathrm{memu}_{ts}) + (1 - \mathrm{res}_{ts}) \odot \mathrm{in}_{ts}. \tag{11}$$

In the above (11), the $(1 - \mathrm{res}_{ts}) \odot \mathrm{in}_{ts}$ keeps the gradient to propagate directly to the ex layer using the skip connection strategy, that is same as adding one to the partial derivative of loss function of cell-state: $([\partial lo]/[\partial \mathrm{outp}]) = ([\partial(\mathrm{forg} + \mathrm{outp})]/[\partial \mathrm{outp}]) = 1 + ([\partial forg]/[\partial \mathrm{outp}])$. Even though the derivative value is small, this method can alleviate the disappearing gradient problem by effectively backpropagating through the error.

The classical models for time-series analysis usually process the input sample sequence from forward to backward position, and then acquire the frontward information. However, this classical approach is inappropriate for data occurrences that contain complex and undefined correlations input data since it may impact the interpretation of the subsequent samples. Therefore, in this work we used an enhanced bidirectional architecture (the structure of which is showed in Fig. 4) to efficiently capture the underlying information of the sequences from the detection samples. The $\overrightarrow{\mathrm{in}}$ in Fig. 4 denotes the forward readings of the input sequence, $\overleftarrow{\mathrm{in}}$ denotes the backward readings, and $\overrightarrow{\mathrm{memu}}$ and $\overleftarrow{\mathrm{memu}}$ denote the forward and backward memory units of the SRUs, respectively. The output of both the forward and backward states of the SRUs are represented by $\overrightarrow{\mathrm{outp}}$ and $\overleftarrow{\mathrm{outp}}$, respectively.

D. Federated-SRU-Based IDS Model

The architecture of federated learning is based on the principles of edge computing and distributed machine learning concepts, i.e., the updation of weight parameters is same as that of distributed machine learning. Federated learning is different from distributed machine learning in a way that each participating client/network has complete authority over its local data (which serve as protecting privacy of ICS networks) and can freely decides to join the federation process. This work proposed a federated architecture for the privacy and security of IoT-based ICS networks by taking advantage of local ICS servers. These local servers with reasonable processing capacity are responsible to detect intrusive events on the local ICS network site instead of being carried out the data to the global server as typically observed in traditional architectures. These local servers at the ICS network serves as a central processing and monitoring unit of a local and distinct ICS network to detect intrusions. Further, to accurately serve the process of training and intrusion detection at every local server without collecting data from all the servers, the federated architecture permits the sharing of information among different localities via a centralized federator/aggregator service on the global server.

As shown in Fig. 3, the training and testing procedures are done at every ICS network having the local data of every industrial network, and these local servers only sends the data regarding the parameters of weight matrix of the learned classifier to the global server without sending the entire set of raw data. Even though the global server can have high processing and large storage ability to store high quantity of data generated and gathered in IoT-based industrial systems, but transmitting the entire data to global server could result in both data leakage and privacy issues, and could also result in a delay due to computational intensive tasks. The proposed IDS model overcome these issues using the concept of federated learning by keeping and processing the data at the premise of local industrial site for real-time detection of intrusions. The operational steps of the proposed federated IDS model are as follows.

1) G-server act as a weight aggregator/federator first initialize the model and create a G-model.
2) Then, communication rounds are initiated and G-model is transmitted to the participating L-servers.
3) After publishing the G-model, the G-server waits for the specific weight parameters configuration from its clients/L-servers.
4) Then, the L-servers download the weight parameters of G-model and build their L-model using their own data collected from the IoT-augmented industrial network.
5) For each round of communication, the L-server sends the weight file WP^n_{ts+1} of the trained model to the G-server for aggregation.
6) The G-server after receiving the updated weight file from L-servers, aggregates and calculates the weights of the G-model using

$$WP_{ts+1} = \sum_{j=1}^{N} \frac{D_n}{D} WP^n_{ts+1} \tag{12}$$

where the number of participating industrial networks is denoted by N, the quantity of data possess by the participating n industrial network is denoted by D_n, the total quantity of data possessed by all the participating industrial networks is denoted by D, the weight parameters of the L-model of industrial network n at $ts+1$ time is denoted by WP^n_{ts+1}, and the weight parameters of the aggregated G-model at $ts+1$ time is denoted by WP_{ts+1}.

7) Finally, after the aggregation of updated weights, the G-server transmit the updated G-model for the L-servers to download and update their L-models.

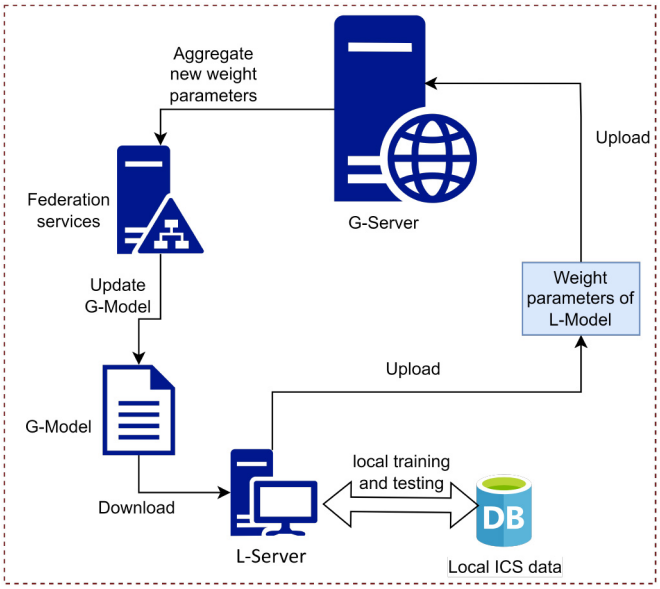


Fig. 5. Proposed federated architecture of communication between G-server and L-server.

To better understand this communication process, these steps are visually shown in Fig. 5. Furthermore, an MD5 check is added to guarantee the integrity of the weight file parameters during the transmission process.

IV. Performance Evaluation of Federated-SRUs

This section presents the experiments conducted to assess the performance of the proposed federated-SRUs IDS model. First, the parameters of experimental environment are outlined which includes the details of experimental setup, the partitioning and description of the data set, the evaluation metrics, and the baseline studies for comparison purposes. Then, experiments are conducted and the comparative results with several state-of-the-art researches and other RNN variants are presented.

A. Experimental Setup

1) *Environment Settings:* The developed IDS is implemented on a workstation having Intel Xeon Silver 4110 CPU 2.10 GHz, NVIDIA RTX2080 GPU, 128 GB RAM, utilizing the Keras API.[1] The parameters of the

[1]Keras: Python deep learning library (https://keras.io).

Feature Type	Feature Name
Network	length
Network	address
Network	time
Network	crc rate
Network	command response
Command Payload	solenoid
Command Payload	setpoint
Command Payload	pump
Command Payload	gain
Command Payload	control scheme
Command Payload	reset rate
Command Payload	system mode
Command Payload	deadband
Command Payload	rate
Command Payload	cycle time
Response Payload	Response Payload
Label	binary attack

Fig. 6. Feature description of gas pipeline data.

Type of Attack	Number of Records
Naïve Malicious Response Injection	7753
Complex Malicious Response Injection	13035
Malicious State Command Injection	7900
Malicious Parameter Command Injection	20412
Malicious Function Code Injection	4898
Denial of Service	2176
Reconnaissance	3874

Fig. 7. Number of records of each attack type in gas pipeline data.

model are selected through the trial and error method having learning rate of 0.001, *amsgrad* set to False, 0.10 as dropout value, batch size of 128, with he_uniform (seed = None) as weight initializer, and Adam as optimizer.

2) *Data Set Partitioning and Description:* Experiments are conducted using a real-scale gas pipeline network (a prominent example of ICS networks) data set[2] [26]. This data set contains seven cyber-attack categories data and one normal category data representing no-attack data or data under normal operations of the ICS network. Every record of this data set has numerous features (as showed in Fig. 6) having their label as normal or attack data. The total number of records in this data set is 274 628 in which 214 580 are normal records (no-attack data) and 60 048 are abnormal (attack data) records. The data is divided into training and testing portions, where 80% of the data is used for training and 20% of the data is used for testing purposes. It is seen from Fig. 7 that the ICS-based gas pipeline data is imbalanced in a way that most of the records in this data set are normal/nonattack records (majority class) covering almost 78% of the total records. Similarly, the attack samples (minority class) are also imbalanced in a way that not all the attack records are equal in numbers (please see Fig. 7 for exact numbers), which indicates that the attack samples are themselves imbalanced in nature. For example, we can see from Fig. 7 that the number of MPCI attack records is 20,412 while the Reconn attack records are only 3874 in numbers.

3) *Evaluation Metrics:* The following evaluation metrics are used to assess the performance of the proposed IDS model.
   a) *Accuracy:* The proportion of model results with correct predictions.
   b) *Recall:* The proportion of model results with records of all attacks correctly predicted as cyber-attack.
   c) *Precision:* The proportion of model results predicted as attack data that are truly attack records.
   d) *F-Measure:* The weighted average of Recall and Precision.

4) *Baseline Researches:* The detection performance of the proposed federated-SRUs IDS model is compared with several state-of-the-art approaches which also used the federated learning architecture. These approaches includes Chen *et al.* [27], Nguyen *et al.* [28], and Schneble and Thamilarasu [29]. These baseline methods are chosen because of two reasons: a) their relevancy to the proposed architecture and b) these approaches are published in reputable venues.

B. Detection Performance and Comparison With Other RNN Variants

As discussed in Section IV-A, experiments are conducted using the Gas pipeline system data to evaluate the performance of our proposed federated IDS model. This data set contains a variety of contemporary attack data, such as denial-of-service attacks, reconnaissance, complex malicious response injections, and command injection attacks. These experiments aims to answer the following research questions.

1) *RQ1:* How accurate is the proposed federated-IDS model in detection complex types of cyber-attack vectors?
2) *RQ2:* How well the proposed federated-IDS model performed as compared to other RNN variants?
3) *RQ3:* To what extent the proposed IDS method alleviate the shortcomings of LSTMs and GRUs?
4) *RQ4:* How robust is the proposed model against imbalanced data?

To answer *RQ1*, the detection accuracy of the proposed federated-IDS model for several types of cyber-attacks is shown in Fig. 8. The evaluation outcomes presented in this figure validate that the proposed IDS model is proficiently capable of detecting complex and contemporary attack vectors in IoT-based ICS networks. As can been seen from Fig. 8, the proposed IDS model achieved higher accuracy rate for all the attack types. These attacks include: 1) DoS attacks: which are capable to shutdown a network or a machine; 2) Reconn attacks: which are capable to acquire information about the targeted network; 3) NMRI attacks: which are capable to inject response packets into the network; 4) MFCI attacks: which are capable to execute commands in a system shell; 5) MSCI attacks: which are capable to exploit a programming flaw;

[2]https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets

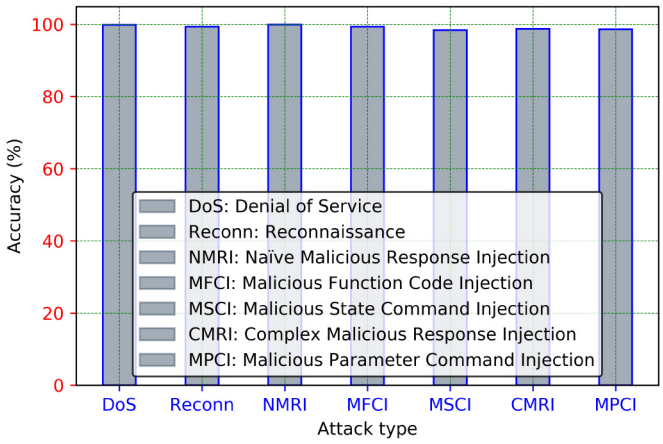


Fig. 8. Performance comparison of the proposed model in detecting different kinds of attacks.

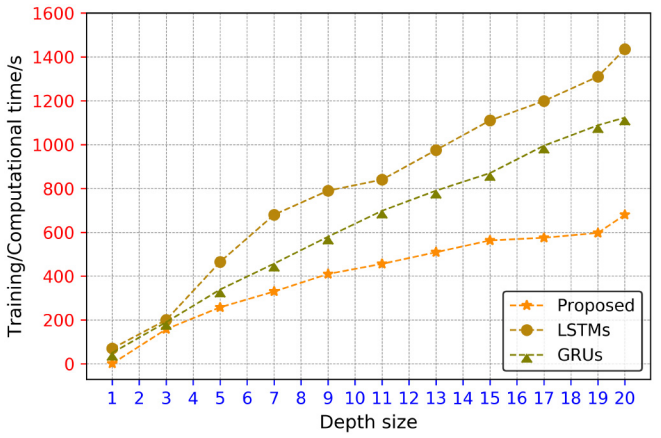


Fig. 9. Computational cost analysis of the proposed model with other RNN variants using different depth size.

TABLE II
NUMERICAL RESULTS OF THE PROPOSED FEDERATED-IDS MODEL AND OTHER RNN VARIANTS

Detection Method	Accuracy	Recall	Precision	F-measure
GRUs	95.705	94.146	95.870	96.269
LSTMs	95.086	92.749	94.649	95.692
Federated-SRUs	99.898	99.834	99.898	99.913

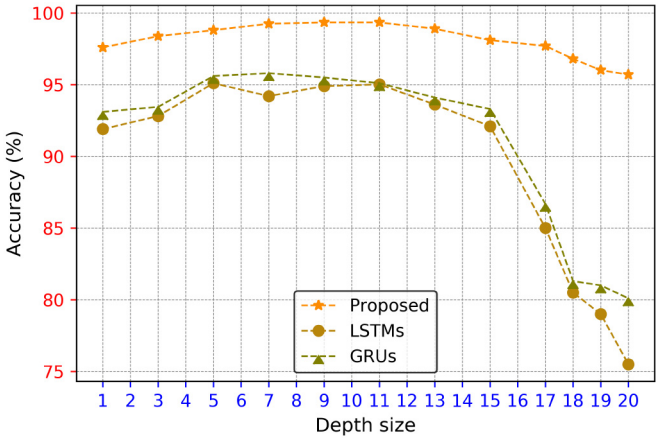


Fig. 10. Accuracy comparison of the proposed model with other RNN variants using different depth size.

and 6) CMRI attacks: which are capable to mimic the normal behavior of the network or system.

To answer *RQ2*, the experimental results are outlined in Table II in terms of four evaluation metrics described in Section IV-A, i.e., accuracy, recall, precision, and f-measure. It can be seen that the proposed federated-IDS design outclasses other RNN variants on all the evaluation metrics by achieving 99.898% accuracy, 99.834% recall, 99.898% precision, and 99.913% f-measure. These results indicates that the proposed SRUs architecture is capable of identifying and learning from underlying data patterns in IoT-based ICS networks as compared to other RNN variants by showing an average improvement of 4.193% in accuracy, 5.688% in recall, 4.028% in precision, and 3.644% in f-measure scores.

To answer *RQ3*, different experiments are conducted which focused on two aspects: 1) the computational cost in terms of training time and 2) the disappearing gradient issue. The numerical results of these experiments are shown in Figs. 9 and 10. It can be seen from Fig. 9 that the proposed IDS model has the shortest training time as compared to other RNN variants (LSTM and GRU). It can also seen that the training time increases with the increase in depth size for all the algorithms. However, the training time of the SRU architecture remains less even for large number of layers. Similarly, Fig. 10 shows the results of the experiment performed to measure the affect of depth size (hidden layer size) in alleviating the vanishing gradient problem. Overall, the three algorithms achieved good accuracy rates having low depth size. However, the vanishing gradient issue starts to appear sooner in LSTMs and GRUs causing a significant drop in detection rate as compared to SRUs. The higher rate of performance of SRUs model with high depth size indicates the efficiency and robustness of the proposed federated-IDS model validating that the proposed model can effectively alleviate the vanishing gradient problem by performing better at higher depth size.

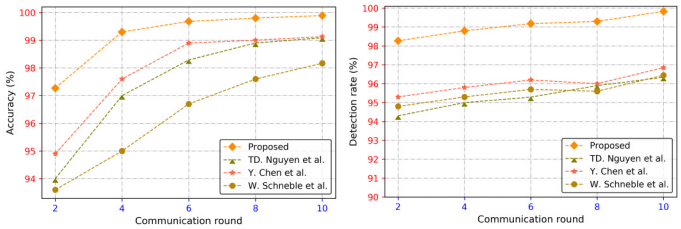
TABLE III
ROBUSTNESS OF THE PROPOSED FEDERATED-IDS MODEL AGAINST IMBALANCED DATA

Attack Type	Accuracy	Detection rate	FAL
DoS	99.86	100	0
Reconn	99.34	99.59	0.66
NMRI	99.90	99.92	0.31
MFCI	99.32	99.30	0.88
MSCI	98.38	99.46	0.45
CMRI	98.69	98.99	0.90
MPCI	98.57	99.01	0.76

To answer *RQ4*, we conducted experiments to validate the robustness of the proposed IDS design against several contemporary cyber-attacks. The numerical outcomes of these experiments are showed in Table III in terms of accuracy, detection rate, and false alarm rate (FAL). We can see from this table, that the proposed model managed to achieved good accuracy and detection rate while having low false alarms rate for all the categories of contemporary cyber-attacks that could be tossed against IoT-enabled ICS networks. These results

TABLE IV
NUMERICAL RESULTS OF THE PROPOSED FEDERATED-IDS MODEL WITH BASELINE APPROACHES

Detection Method	Accuracy	Recall	Precision	F-measure
Y. Chen *et al.* [27]	99.13	96.85	99.03	97.90
TD. Nguyen *et al.* [28]	99.09	96.34	99.91	97.77
W. Schneble *et al.* [29]	98.17	96.45	98.79	97.57
Federated-SRUs	99.89	99.83	99.89	99.91



(a) (b)

Fig. 11. Performance Comparison of proposed and baseline methods where (a) shows accuracy and (b) shows detection rate (recall).

demonstrate that the proposed federated-SRUs design has performed very well and is robust against imbalanced data of ICS networks.

C. Performance Comparison With Baseline Methods

We also conduct experiments and compared the numerical results of our proposed IDS model with other compelling state-of-the-art approaches mentioned in Section IV-A. These experiments are conducted with different rounds of communication, and the best results are outlined in Table IV using the metrics of accuracy, recall, precision, and f-measure rates for all the compared methods. It can be seen that the proposed federated-IDS model performed better than other state-of-the-art benchmark approaches. Also, it is worth mentioning that the performance of all IDS models gradually stabilizes and improves with increasing number of communication rounds indicating the stability and adaptiveness of the proposed IDS model. These results are also visualized in Fig. 11(b) using the accuracy and detection rate (recall) metrics, which indicates the superiority of the proposed federated-IDS model as compared to baseline researches.

D. Discussion and Implications

Conclusively, it is evident from the above-mentioned evaluation results that the federated IDS approach adequately achieved better performance than other RNN variants and existing benchmark approaches. It is, therefore, worth mentioning that it would be an applicable security solution for all industrial networks owing to its ability in preserving privacy of the data at their own premises and high efficiency in detecting intrusions and cyber-attacks of multiple types.

Through this research work, the proposed IDS model delivers several valuable practical implications in the process of adapting the federated learning architecture for the security of IoT-based ICS networks. The proposed federated-SRUs IDS model could potentially be serve as a decent guideline in developing robust and scalable federated architecture and transforming from nonfederated to federated-based learning. The numerical results of experiments offer insight into the effectiveness of SRUs using different depth size and computational cost, which can be further studied to avoid cold start problems in federated architectures and assist the security professionals to employ the federated architecture in real-word IoT and industrial networks. Additionally, the proposed federated-IDS model both lessens the shortcomings of LSTM and GRU variants of RNN, as well as the distributed learning architecture improves privacy fears by empowering sharing of computational resources and permitting the security of data that can result in an operative and robust method for the discovery intrusive activities in IoT-enabled ICSs.

V. CONCLUSION AND FUTURE WORK

In this research work, the authors put forwards an IDS model named federated-SRUs for the accurate detection of cyber-attacks against IoT-augmented industrial networks (specifically ICSs). The proposed security model has proved to have achieved the high performance in detecting and identifying multiple kinds of attacks, whilst having the advantage of fast training time as an improved RNN architecture is applied. The proposed federated-SRUs model allows the data to be processed at the local premises of industrial networks, which empowers the ICS owners to preserve the privacy and security of data during the transmission process. The federated architecture of the proposed IDS model distributes the detection and monitoring tasks to local ICS servers, so that they can deal with the generated data inside their industrial network, and aggregates the model parameters through communication rounds with the local servers. The proposed IDS model proved to be accurate in detecting contemporary attack vectors in IoT-based ICS environments.

For the future work, we plan to optimize the weights communicating between the local and global servers in the IDS model using asynchronous FL techniques (as mentioned in [30]), as well as testing the proposed model with more suitable training data from different ICS networks. Furthermore, we plan to test the proposed model with incremental learning approach since it has the capability to update the security model in real time with a gradual change in devices, environment, and time without much computation when device is added or new network data arrive.

REFERENCES

[1] I. A. Khan, N. Moustafa, D. Pi, K. M. Sallam, A. Y. Zomaya, and B. Li, "A new explainable deep learning framework for cyber threat discovery in industrial IoT networks," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11604–11613, Jul. 2022.
[2] I. A. Khan, D. Pi, Z. U. Khan, Y. Hussain, and A. Nawaz, "HML-IDS: A hybrid-multilevel anomaly prediction approach for intrusion detection in SCADA systems," *IEEE Access*, vol. 7, pp. 89507–89521, 2019.
[3] "Securing industrial control systems: A unified initiative FY 2019–2023." CISA. Accessed: Feb. 2022. [Online]. Available: https://www.cisa.gov/publication/securing-industrial-control-systems
[4] G. Falco, C. Caldera, and H. Shrobe, "IIoT cybersecurity risk modeling for SCADA systems," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4486–4495, Dec. 2018.
[5] B. Bencsáth, G. Pék, L. Buttyán, and M. Félegyházi, "Duqu: A Stuxnet-like malware found in the wild," CrySyS Lab, Budapest, Hungary, Rep. v0.93, vol. 14, pp. 1–60, Oct. 2011.

[6] D. Lee, *Flame: Massive Cyber-Attack Discovered, Researchers Say*, vol. 5, BBC News, London, U.K., 2012.
[7] N. Ito, "A historical review of the techniques in Japanese buildings for resisting various loads, focusing on seismic attacks," in *Proc. ICOMOS IWC–XVI Int. Symp.*, 2007, pp. 1–5.
[8] I. A. Khan *et al.*, "Efficient behaviour specification and bidirectional gated recurrent units-based intrusion detection method for industrial control systems," *Electron. Lett.*, vol. 56, no. 1, pp. 27–30, 2020.
[9] I. A. Khan *et al.*, "A privacy-conserving framework based intrusion detection method for detecting and recognizing malicious behaviours in cyber-physical power networks," *Appl. Intell.*, vol. 51, no. 10, pp. 7306–7321, 2021.
[10] H. Gu, Y. Lai, Y. Wang, J. Liu, M. Sun, and B. Mao, "DEIDS: A novel intrusion detection system for industrial control systems," *Neural Comput. Appl.*, vol. 34, pp. 9793–9811, Feb. 2022.
[11] W. Wang *et al.*, "Abnormal detection technology of industrial control system based on transfer learning," *Appl. Math. Comput.*, vol. 412, Jan. 2022, Art. no. 126539.
[12] J. Ling, Z. Zhu, Y. Luo, and H. Wang, "An intrusion detection method for industrial control systems based on bidirectional simple recurrent unit," *Comput. Elect. Eng.*, vol. 91, May 2021, Art. no. 107049.
[13] Y. Wu, H.-N. Dai, and H. Tang, "Graph neural networks for anomaly detection in industrial Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9214–9231, Jun. 2022.
[14] X. Liu, W. Yu, F. Liang, D. Griffith, and N. Golmie, "Toward deep transfer learning in industrial Internet of Things," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12163–12175, Aug. 2021.
[15] X. Wang *et al.*, "Toward accurate anomaly detection in industrial Internet of Things using hierarchical federated learning," *IEEE Internet Things J.*, vol. 9, no. 10, pp. 7110–7119, May 2022.
[16] A. Makkar, T. W. Kim, A. K. Singh, J. Kang, and J. H. Park, "SecureIIoT environment: Federated learning empowered approach for securing IIoT from data breach," *IEEE Trans. Ind. Informat.*, vol. 18, no. 9, pp. 6406–6414, Sep. 2022.
[17] B. Li, Y. Wu, J. Song, R. Lu, T. Li, and L. Zhao, "DeepFed: Federated deep learning for intrusion detection in industrial cyber–physical systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5615–5624, Aug. 2021.
[18] A. N. Jahromi, H. K. Schulich, and A. Dehghantanha, "Deep federated learning-based cyber-attack detection in industrial control systems," in *Proc. 18th Int. Conf. Privacy Security Trust (PST)*, 2021, pp. 1–6.
[19] T. T. Huong *et al.*, "Detecting cyberattacks using anomaly detection in industrial control systems: A federated learning approach," *Comput. Ind.*, vol. 132, Nov. 2021, Art. no. 103509.
[20] I. A. Khan, N. Moustafa, D. Pi, Y. Hussain, and N. A. Khan, "DFF-SC4N: A deep federated defence framework for protecting supply chain 4.0 networks," *IEEE Trans. Ind. Informat.*, early access, Sep. 1, 2021, doi: 10.1109/TII.2021.3108811.
[21] L. Kong, X.-Y. Liu, H. Sheng, P. Zeng, and G. Chen, "Federated tensor mining for secure industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 2144–2153, Mar. 2020.
[22] G. Bovenzi, G. Aceto, D. Ciuonzo, V. Persico, and A. Pescapé, "A hierarchical hybrid intrusion detection approach in IoT scenarios," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, 2020, pp. 1–7.
[23] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," 2018, *arXiv:1802.09089*.
[24] I. A. Khan, M. Keshk, D. Pi, N. Khan, Y. Hussain, and H. Soliman, "Enhancing IIoT networks protection: A robust security model for attack detection in Internet industrial control systems," *Ad Hoc Netw.*, vol. 134, Jul. 2022, Art. no. 102930.
[25] I. A. Khan *et al.*, "XSRU-IoMT: Explainable simple recurrent units for threat detection in Internet of medical things networks," *Future Gener. Comput. Syst.*, vol. 127, pp. 181–193, Feb. 2022.
[26] T. Morris and W. Gao, "Industrial control system traffic data sets for intrusion detection research," in *Proc. Int. Conf. Crit. Infrastruct. Prot.*, 2014, pp. 65–78.
[27] Y. Chen, X. Qin, J. Wang, C. Yu, and W. Gao, "FedHealth: A federated transfer learning framework for wearable healthcare," *IEEE Intell. Syst.*, vol. 35, no. 4, pp. 83–93, Jul./Aug. 2020.
[28] T. D. Nguyen, S. Marchal, M. Miettinen, H. Fereidooni, N. Asokan, and A.-R. Sadeghi, "DïoT: A federated self-learning anomaly detection system for IoT," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, 2019, pp. 756–767.
[29] W. Schneble and G. Thamilarasu, "Attack detection using federated learning in medical cyber-physical systems," in *Proc. 28th Int. Conf. Comput. Commun. Netw. (ICCCN)*, vol. 29. Valencia, Spain, 2019, pp. 1–8.
[30] G. Aceto, D. Ciuonzo, A. Montieri, V. Persico, and A. Pescapé, "Know your big data trade-offs when classifying encrypted mobile traffic with deep learning," in *Proc. Netw. Traffic Meas. Anal. Conf. (TMA)*, 2019, pp. 121–128.