Soft Computing
https://doi.org/10.1007/s00500-022-06761-1

FOCUS

On the performance metrics for cyber-physical attack detection in smart grid

Sayawu Yakubu Diaba[1] · Miadreza Shafie-khah[2] · Mohammed Elmusrati[1]

Accepted: 3 January 2022

Abstract
Supervisory Control and Data Acquisition (SCADA) systems play an important role in Smart Grid. Though the rapid evolution provides numerous advantages it is one of the most desired targets for malicious attackers. So far security measures deployed for SCADA systems detect cyber-attacks, however, the performance metrics are not up to the mark. In this paper, we have deployed an intrusion detection system to detect cyber-physical attacks in the SCADA system concatenating the Convolutional Neural Network and Gated Recurrent Unit as a collective approach. Extensive experiments are conducted using a benchmark dataset to validate the performance of the proposed intrusion detection model in a smart metering environment. Parameters such as accuracy, precision, and false-positive rate are compared with existing deep learning models. The proposed concatenated approach attains 98.84% detection accuracy which is much better than existing techniques.

Keywords Supervisory control and data acquisition (SCADA) systems · Intrusion detection system (IDS) · Industrial control system (ICS) · Cyber-physical security · Smart grid · Convolutional neural network (CNN) · Gated recurrent unit (GRU)

1 Introduction

Most of the Intrusion Detection Systems (IDS) used in Supervisory Control and Data Acquisition (SCADA) in power distribution networks are currently concentrated on the cyber sector by ignoring the process states in the physical field (Rakas et al. 2020). Attacks on protocol traffic are being detected, but attacks on processes like Replay and Man-in-the-Middle (MITM) attacks are complex to detect. The performance criteria, risk management requirements, and coordination requirements vary between Information Technology System (IT System) and Industrial Control System (ICS) networks. In an IT system, a long delay could be appropriate, but coming to ICS, reaction time is crucial (Ghosh and Sampalli 2019). In IT systems, data security and integrity are most significant, whereas fault tolerance is less significant, while in ICS, human safety is most crucial, followed by process security, and fault tolerance is necessary (Paridari et al. 2018). Many specific communication protocols without ID certification, encryption, and timestamps were utilized in ICS, whereas standard communication protocols are utilized in IT systems. Zero-day, Denial of Service (DoS), Replay and MITM attacks to ICS will trigger the above-mentioned delay, fault, and information leakage caused by protocols (Hu et al. 2019).

Cyber-physical systems are widely used to integrate the physical process and computations so that the system can be controlled effectively. The performance of the system relies on proper control of its elements, like sensors and actuators. Efficient and secure communication between the system element is most important as it directly affects the

Communicated by Joy Iong-Zong Chen.

✉ Sayawu Yakubu Diaba
saywu.diaba@student.uwasa.fi

Miadreza Shafie-khah
mshafiek@uwasa.fi

Mohammed Elmusrati
moel@uwasa.fi

[1] School of Technology and Innovation, Department of TelecommunicationsEngineering, University of Vaasa, Vaasa, Finland

[2] School of Technology and Innovations, Department of Electrical Engineering, University of Vaasa, Vaasa, Finland

Published online: 21 January 2022

system performance. Malfunctioning the device characteristics might lead to a serious issue in industrial control systems. The system elements face serious security threats which affect the sensing and data actuation. Attacks on IT system networks cause congestion or data leakage, but attacks on ICS networks may result in both data leakage as well as harm to physical infrastructure. As a result, for the SCADA systems, which are commonly utilized in the power distribution networks to ensure the security of the controlled processes, cyber-security is considered a significant part of SCADA (Zhang et al. July 2019). The protection of communication protocols, asset management, physical infrastructures, and controlled processes will come under the security of the SCADA system that is the most important element of the smart grid, and these cannot be handled the same as IT system contemporaries. Some of the key components are supporting software such as Human Machine Interface (HMI), Distributed Control Systems (DCS), Programmable Logical Controllers (PLC), Remote Terminal Units (RTU), network equipment, servers, and computers (Cómbita et al. 2019, Pang et al. 2020, Sun et al. 2020, Elnour et al. 2020). Hence, it is essential to protect the system against attacks and secure communication.

Intrusion detection systems are used to detect the security threats and attacks in a system where the systems can able to detect but not able to prevent the attacks. However, by training the detection systems properly the attacks can be detected efficiently without any manual intervention which may reduce the huge loss compared to the loss acquired in a system without an intrusion detection system. These systems will work as a second-line defense in any architecture and plays a vital role in cyber-physical systems to detect different types of attacks. Intrusion detection systems classify normal and abnormal behavior which helps the system to detect unknown attacks. This essential feature is adopted for cyber-physical systems. A wide range of devices and dynamic computing resources, different software, and operating systems are generally included in cyber-physical systems. Detecting intrusion in such systems using machine learning algorithms-based models is crucial due to the heterogeneous deployment nature. Obtaining labels for attacks can be very time-consuming, challenging, and sometimes even impossible. Therefore, unsupervised learning techniques, capable of detecting cyber-attacks without a need for labels, are deemed best for this task (Keshk et al. 2021, Gumaei et al. 2020). However, the most existing unsupervised techniques are not able to deal with the nonlinearity and inherent correlations of multivariate time series, which represent a considerable amount of real-world data, including data streams generated by sensors in CPSs (Hu et al. 2019; Rodofile et al. 2019; Homay et al. 2020). Therefore, a new unsupervised technique independent from any prior knowledge of cyberattacks is needed to detect intrusions in CPSs.

The major contributions of this paper are summarized as follows.

(1) CNN and GRU are combined to obtain an intrusion detection system for detecting attacks in smart grid metering infrastructure.
(2) An intense experimental analysis is presented using benchmark datasets to obtain improved accuracy and detection rate performance for the proposed model.

The rest of the paper is arranged as follows: a brief literature analysis is presented in Sect. 2; the proposed intrusion detection model is presented in Sect. 3. Experimental results and observations are discussed in Sect. 4 and finally, the conclusion is presented in Sect. 5.

2 Related works

Recent research works in industrial systems and their evolutions are discussed in this section. Intrusion detection is the major objective and the research directed toward analyzing the features of existing intrusion detection systems. The authors of Khan et al. (2019) have introduced a new method called anomaly detection for ICS. This method utilized a hybrid approach by taking the benefits of the reliable and predictable nature related to communication patterns, which perform in-ground devices in ICS. Initially, few preprocessing approaches were implemented for scaling and standardizing the data. To enhance the performance of anomaly detection, dimensionality reduction algorithms were used. Later, the nearest-neighbor rule algorithm was utilized for balancing the dataset. A signature database was created by noting the system in a time using a bloom filter. Subsequently, a hybrid approach was created for anomaly detection by combining the instance-based learner and package contents-level detection. Here, the developed model has attained the best results when compared to other state-of-the-art models.

The authors of Qian et al. (2020) have suggested a method in a physical way as well as a cyber-way for attack detection. In order to detect malicious behaviors for physical component prevention, process states validation was utilized and being damaged by Zero-day, MITM, and Replay attacks. For branching shaped data sets classification, a nonparallel hyperplane-based fuzzy classifier was developed that was quite complex, complex to classify using two parallel hyperplanes of the Support Vector Machine (SVM) to detect DoS and other cyber-attacks. To test the developed model and validation part, Modbus/Transmission Control Protocol (TCP) traffic data and

simulation process states were used. Thus, it has been proved that the suggested approach was superior to other approaches.

In (Sheng et al. 2021), a cyber-physical technique in the SCADA system for intrusion detection has analyzed the risk levels faced in industries. This was utilized to characterize the structure of the network and SCADA system's process by the extraction and correlation of communication patterns and the ICS device condition. If any violation occurs, then this was considered as an abnormal behavior that was caused due to network attacks. A risk estimation approach was suggested to measure the damage degree of the attack on the infrastructure by associating network intrusions with the state of the SCADA system, providing network teams with more knowledge regarding network attacks. Furthermore, the proposed approach outperformed existing approaches in identifying and evaluating numerous cyber-attacks against the SCADA system.

Privacy-Preserving Anomaly Detection framework named PPAD-CPS is reported in the research work of Keshk et al. (Keshk et al. 2021). The aim is to secure confidential data and discover malicious attacks in power systems and their network traffic. This framework included two modules namely data preprocessing and anomaly detection modules. To filter and transform the real data into a new kind of data, a data preprocessing module was recommended that has attained privacy preservation target. By using Kalman Filter (KF) and Gaussian Mixture Model (GMM), an anomaly detection model was employed for analyzing the posterior probabilities of anomalous and legitimate events. Two public datasets such as UNSW-NB15 and Power System were used for analyzing the proposed framework. This analysis has been proved that the developed PPAD-CPS outperformed the existing methodologies.

In (Kalech 2019) cyber-attack detection models based on temporal pattern recognition (TPR), which searched for anomalies in the data sent by the SCADA elements in the network and found anomalies that were occurred when legal commands like incorrect and unauthorized time intervals were misused. Artificial Neural Networks (ANN) and Hidden Markov Model (HMM) were suggested for evaluating the performance. The evaluation was done on both simulated and real SCADA data using five various feature extraction approaches. The outcomes have shown that TPR models were performed well in detecting cyber-attacks.

Gumaei et al. (Gumaei et al. 2020) have proffered a new security control method for cyber-attack detection in smart grid, which merged feature detection and reduction models for decreasing the features count and attained a better detection rate. For eliminating irrelevant features, the Correlation-based Feature Selection (CFS) technique was utilized that enhanced the detection rate. With the help of optimal features that were selected, cyber and normal attack events were classified using the Instance-Based Learning (IBL) algorithm. By using public datasets of SCADA power system, the experiments were performed relied on tenfold cross-validation approach. This has been revealed that the suggested model consisted of huge detection rate.

Rodofile et al. (Rodofile et al. 2019) have presented a cyber-attack structure, which detects attacks in SCADA systems. The developed model recognized "traditional IT-based attacks, protocol specific attacks, configuration-based attacks and control process attacks" for describing the practical attacks. The recognition of attacks in the whole system has an advantage of allowing us to protect over them with more effectiveness and awareness. A case study was presented by illustrating the sequence of attacks on Distributed Network Protocol 3 (DNP3), which facilitated to affirm the reliability of the developed model.

Reference (Homay et al. 2020) has implemented a robust security control solution as a logic level security on DCS and SCADA systems. In order to establish trust among DCS device components, the developed model ensured message integrity, but this was not considered as the protection layer on industrial automation systems. Malicious attacks like Stuxnet were avoided by the developed solution called low-level security process. From the analysis, the following points are observed as research gaps. The security of SCADA systems has been disrupted using earlier IC system attacks. The significance of defending and securing ICS networks has increased attacks on critical infrastructures. The features and challenges of various methodologies of detecting and blocking cyber-security attack on SCADA systems that has existed earlier are given subsequently. HML-IDS can detect anomalies very fast and it can detect unseen attacks also it can deal with data samples that seem to be hybrid which is complex. But it has some demerits like enhancement of detection rate (Feng et al. 2017).

NHFC is capable of modeling a given problem into any degree of accuracy and has high detection accuracy in detecting zero-day, Replay, and MITM attacks. It needs to improve in detecting the attacks for securing the manufacturing process (Wang et al. August 2020). Cyber-physical model is used to detect the network intrusions. It has high accuracy. Though, it is not considered as the secure appliance, because of the lack of multi-factor authentication models. PPAD-CPS has huge privacy levels. It attains the best accuracy, detection rate, and processing times. It needs to perform a principal and independent component analysis for transforming the high-dimensional space into low-dimensional space in order to enhance the performance. TPR can detect cyber-attacks including

legitimate functions. It has high detection accuracy. It needs to reduce the count of false alarms by considering PLC identity (Dhaya 2021; Jacob and Ebby Darney 2021; Haoxiang and Smys 2021; Smys et al. 2021).

CFS-KNN resorts to various correlation measures for removing the irrelevant features and retaining those using predictive power. It is robust to exploit inter-feature relationships. It is sensitive to noise and has an over-fitting problem, which leads to reduce the system performance. DNP3 consists of efficient Internet Provider Security (IPS) and IDS technologies. It can combine four categories for describing the practical attacks. However, troubleshooting the system is quite complex because of distribution over many servers.

To overcome these limitations concatenated deep learning approach is presented in this paper. Deep learning approach has the ability to process high-dimensional data and produce better results than machine learning approaches. (Fig. 1)

3 Proposed work

The proposed concatenated deep learning model is presented in this section to detect intrusions in Cyber-Physical Systems. The simple environment considered for this paper is depicted in Fig. 2. The framework consists of two entities such as industrial agents and cloud servers. The industrial agents are the industrial Cyber-Physical System owners and they oversee the local intrusion detection model. Collecting industrial cyber-physical system data and updating the essential parameters for intrusion detection are some of the regular activities of industrial agents. Industrial agents interact with cloud servers to update the data. Whereas cloud servers take the responsibility of building a comprehensive intrusion detection system using the model parameters obtained from the locally trained model in the industrial agents. The final intrusion detection model could be obtained through multiple interactions between each agent and cloud server.

The threat model we have considered for analysis includes four different threats such as reconnaissance attacks, command injection attacks, response injection attacks, DoS attacks. Cyber threats targeting industrial cyber-physical security are considered for the proposed concatenated deep learning-based intrusion detection system. The reconnaissance attacks are performed to collect industrial cyber-physical security system valuable information. Network architecture details, device features, network protocols are the major aim of the intruder to perform reconnaissance attacks. To mislead or deviate the industrial physical security system behavior, command injection attacks are performed. In this attack, the intruder injects some false information to control a system or provides wrong configuration commands to collapse the system behavior. Unauthorized access, invalid communications, wrong set points are the outcomes of a command injection attack.

To monitor and observe the remote process state in the industrial cyber-physical security system, response injection attacks are performed. These attacks interfere with the system process and provide false responses to the service queries which affect the system state. DoS attacks are quite common and familiar in the network, in the case of industrial cyber-physical security systems the attacker flooded the targets with redundant requests to deplete the server resources in the industrial cyber-physical security system. Due to these boundless requests, the system prevents legitimate requests which affect the system services. Figure 3 depicts the proposed concatenated deep learning

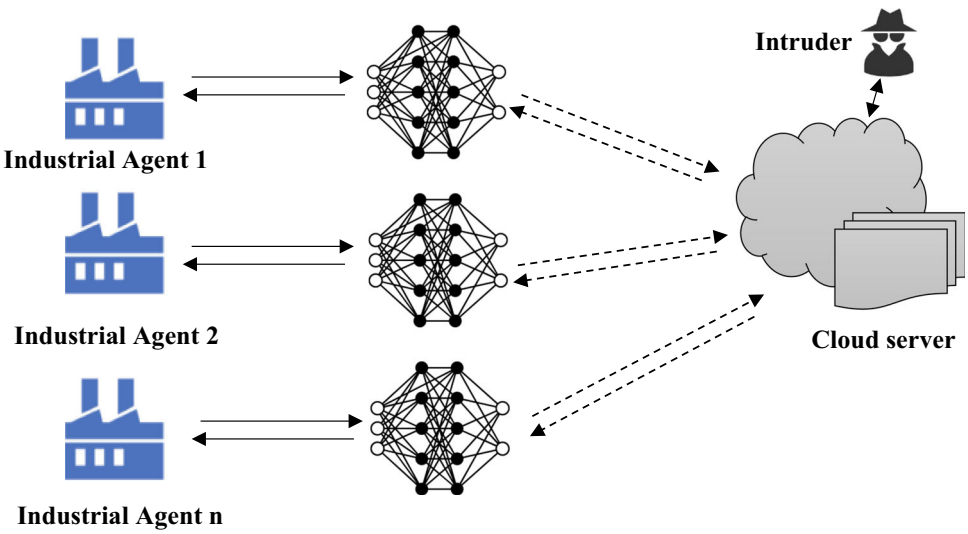

Fig. 1 Proposed system model

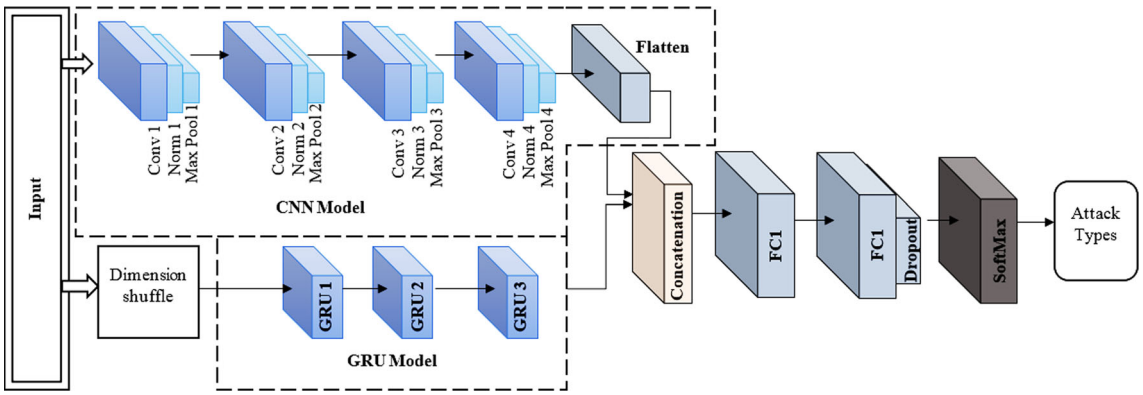

Fig. 2 Proposed concatenated deep learning model

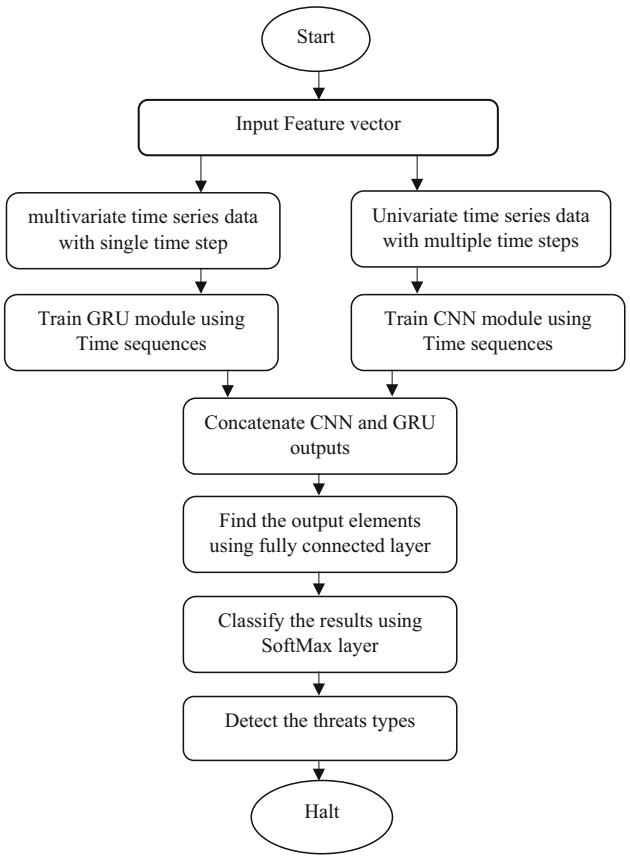

Fig. 3 Process flow of proposed Intrusion detection model

model for intrusion detection in the industrial cyber-physical security system.

The proposed intrusion detection model includes a CNN model and a GRU unit module. The outputs of both modules are combined and processed by a fully connected layer followed by the SoftMax layer. The building block of the convolutional network model includes four convolution blocks. Each block includes a convolutional layer, a max-pooling layer. Batch normalization is performed between the convolution layer and max-pooling layer. The GRU is comprised of three GRU layers. The results of the CNN and GRU model are concatenated and then processed using two fully connected layers. In order to prevent over-fitting, a dropout layer is included after the fully connected layer. Finally, the SoftMax layer is used to map the output of fully connected to probability distribution and predicts the attack types.

3.1 System model

Consider a feature vector v as input for the proposed model which is a one-dimensional vector function representing the numerical features of industrial data. The input is processed by CNN and GRU model. GRU is a modified LSTM model which includes a gated recurrent neural network. LSTM consists of three gates such as forget gate, input gate, and output gate whereas GRU comprises of two gates such as update gate and reset gate. Due to this it requires less parameter for training which provides quick convergence compared to LSTM. Owing to this reason instead of LSTM, GRU is adopted in the proposed design. The long dependency features are captured by the GRU module and learn essential information from the historical data using memory cell. The reset gate is used to forget or remove unnecessary information. Generally, the input for GRU module will be time sequence data and the given input feature vector is a multivariate time series data with a single time step. Therefore, prior to GRU module, a dimension shuffling process is performed that transposes the dimensions of the feature vector v. The dimension shuffling is given as

$$\tilde{v} = \mathrm{shu}f(v) \tag{1}$$

The GRU module process the dimensional shuffled data $\tilde{v}$ and produces output. The output of first layer is given as input to the second layer and it repeats. The output of GRU module is obtained using two activation functions such as

tan$h$ and σ. The essential formulations observed for the GRU module is summarized as follows

$$\Gamma_U = \sigma\left(\omega_U\left[\tilde{v}^{(t-1)}, x^{(t)}\right] + b_U\right) \tag{2}$$

$$\Gamma_R = \sigma\left(\omega_R\left[\tilde{v}^{(t-1)}, x^{(t)}\right] + b_R\right) \tag{3}$$

where Γ_U and Γ_R represents the update gate and reset gate respectively. The range of $\Gamma_U \in \{0, 1\}$ and the range of $\Gamma_R \in \{-1, 1\}$. ω_U and ω_R are the weight functions of update and reset gate. b_U and b_R represents the bias vectors for update and reset gate correspondingly. The candidate activation function for the recurrent unit is given as

$$\tilde{v}^{(t)} = \tanh\left(\omega_v\left[\Gamma_R \times \tilde{v}^{(t-1)}, x^{(t)}\right] + b_v\right) \tag{4}$$

where ω_v are the weight functions of activation function, b_v is the bias vector and $x^{(t)}$ is the inputs of training data. The output of a single GRU unit is given as

$$v^{(t)} = \left((1 - \Gamma_U) \times \tilde{v}^{(t-1)}\right) + \left(\Gamma_U \times \tilde{v}^{(t)}\right) \tag{5}$$

where $\tilde{v}^{(t-1)}$ is the current unit input which is obtained from the previous unit output. The final output of GRU module is given as I. The same input used for GRU model is parallelly provided to CNN module. The input feature vector v is considered as a univariate time series data with multiple time steps. Since CNN is suitable to process high-dimensional data it does not require any dimensional shuffling module as like GRU. One-dimensional layer is used in the proposed model and the convolution operation is represented as

$$h_1 = \text{convblock1}(v) \tag{6}$$

$$h_2 = \text{convblock2}(h_1) \tag{7}$$

$$h_3 = \text{convblock3}(h_2) \tag{8}$$

$$h_4 = \text{convblock4}(h_3) \tag{9}$$

where h_1, h_2, h_3 and h_4 are the hidden vectors. After each convolution layer, a batch normaliztion and pooling layer is included. The normalization layer normalizes the features which is obtained after the convolution proces and the pooling layer is used to reduce the dimensionality of the data. There are two types of pooling functions such as max pooling an average pooling. In this research work, average pooling is used in the pooling layer. The reduced output from the pooling layer is provided as the input to the next convolution block. The average pooling layer is mathmatically expressed as

$$x^i = \text{avg}P\left(x^{i-1}\right) \tag{10}$$

where x^i is the ouput of pooling layers, and x^{i-1} is the previous values obtained from the convolution layer. i represents the number of pooling layers. The final output of convolution layer is given to flatten layer that converts the data into one-dimensional vector and it is given as

$$J = \text{flatten}(h_4) \tag{11}$$

The output I from the GRU module and J from the CNN module are concatenated which is described as $c_t = \text{concate}(I, J)$. Followed by two fully connected layers are used in the proposed design. To prevent data overfitting a dropout function is used. Then the dropout layer, the final SoftMax layer provides the classification results which provide the attack types. The SoftMax function is described as

$$\widehat{y} = \text{softmax}(\varphi) \tag{12}$$

where φ is the output of dropout layer. In order to evaluate the loss function for the proposed model cross-entropy function is used and it is given as

$$l = -\frac{1}{b}\sum_{i=1}^{n} y_i \log y_i^{'} \tag{13}$$

where the batch size is given as b, n denotes the training sample size, the predicted value is represented as $y_i^{'}$ athe nd actual value is represented as y_i. The process flow of the proposed approach is depicted in Fig. 3.

Figure 3 depicts the process flow of the proposed intrusion detection model. Initially, the process starts with the selection of input feature vectors. The multivariate time series data with a single time step is used to train the GRU model and univariate series data with multiple time steps is used to train the CNN model. The output features of each model are concatenated, and the elements are provided as input to the fully connected layer. The final classified results are obtained from the SoftMax layer and they provide the details of threats and their types.

4 Results and discussion

The proposed intrusion detection model is being experimented on a smart metering environment and the model is comprised of three network configurations, such as Home Area Network (HAN), Wide Area Network (WAN), and Neighborhood Area Network (NAN). The household applications, concentrators, smart electricity meters, data processing centers, and nodes are included in the smart infrastructure. Wireless communication or wired communication is used for communicating the elements. The communication is bidirectional and mainly internal communications are performed through HAN, which is vulnerable to attacks. A DoS attack and a probing attack are the major attacks on HAN. The NAN is used for short-

distance communication, and it is vulnerable to the user to Root (U2R) attack. WAN is used for long-distance communication and it is vulnerable to Remote to Local (R2L) attacks.

To evaluate the performance of the proposed intrusion detection system with a standard benchmark dataset, the NSL-KDD dataset is used. The data include 125,973 attacks and normal data which are provided as input to the proposed intrusion detection. Deep learning models are not able to process character-based features, so to simplify and process the input data, preprocessing steps such as normalization and feature screening are performed before the input is fed into the deep learning model. The attack data present in the dataset is categorized into four types, such as DoS, Probe attack, U2R attack, and R2L attack. The subclasses of attacks are 39 and they cover the attack types of smart metering infrastructure. The experimentation process trains and tests the data in the ratio of 80:20 according to the fivefold cross-validation method. The dataset distribution for the NSL-KDD dataset is listed in Table 1.

In the data preprocessing, the characteristic features are converted into numerical values, specifically as Eigenvectors, using a one-hot encoding process. For this process, flag features, services, and protocol types are considered in the dataset. The protocol type considers attributes such as user datagram protocol (UDP), TCP, and Internet control message protocol (ICMP). The numerical data and one-hot encoding represent the feature vectors in 1×3 dimension as like (0,0,0), (0,1,1), etc. The service feature has 70 attributes and flag features include 11 attributes and these attributes. The features after preprocessing are mapped into numerical features and combined with existing numerical features in the dataset. Also, the labels in the dataset are numerically processed such that, normal behavior is represented as '0', DoS is represented as '1', Probe label is represented as '2', R2L as '3', and U2R as '4'. In order to reduce the feature differences, the dataset is normalized and uniformly mapped. The interval range of uniform mapping is [0, 1].

Table 1 Dataset Distribution of NSL-KDD

Type	Total	Training set	Test set
Normal	67,343	53,874	13,469
Dos	45,927	36,742	9185
Probing	11,656	9325	2331
R2L	995	796	199
U2R	52	42	10
Total	1,25,973	100,778	25,195

The proposed intrusion detection model is implemented in Spyder3.0 (Python3.6) operating on Windows 10 OS installed on an i5 Intel processor at 4.20 GHz and 8 GB of memory. The learning rate was set at 0.006 and the number of epochs was 100. The hyperparameter details used in the proposed work are listed in Table 2.

The proposed model performance is further evaluated based on the convergence ability and classification performance in terms of accuracy, detection ratio, and false-positive rate. The training loss and epoch for the proposed concatenated model are depicted in Fig. 4. It is observed from the results. The training loss is gradually decreased and stable after the eighth epoch. This indicates the selection of hyperparameters is reasonable and validates the convergence ability of the proposed model. The confusion matrix for the proposed approach is depicted in Table 3.

The performance of the proposed model in terms of precision, detection rate, f1-score, and false-positive rate is depicted for the four types of attacks in Fig. 5. Based on the values obtained from the confusion matrix, the parameters are evaluated.

From Fig. 5, it can be observed that the detection abilities for DoS and Probe are above 95%, whereas the detection abilities for U2R and R2L are below 90% and 50% due to the limited number of training data. The performance of the proposed model is validated using fivefold cross-validation and the confusion matrix obtained after fivefold validation is depicted in Table 4.

The performance of the proposed model in terms of precision, detection rate, f1-score, and false-positive rate is depicted for the four types of attacks in Fig. 6. Based on the values obtained from the confusion matrix given in Table 4, the parameters are evaluated. It is observed from the results that the detection rate of the proposed model is maximum for DoS and Probe attacks and it has been reduced for R2L and U2R attacks. The reduced performance is due to the minimum number of samples in the

Table 2 Hyperparameter settings

S. no	Parameter	Filter/Neurons
1	Number of filters in CNN	8
	Number of Neurons in CNN and ReLu	16
2	Number of Hidden nodes	60
3	Activation function	ReLU
4	Dense layer	256
5	Cost function	Cross entropy
6	Batch size	128
7	Epoch	100

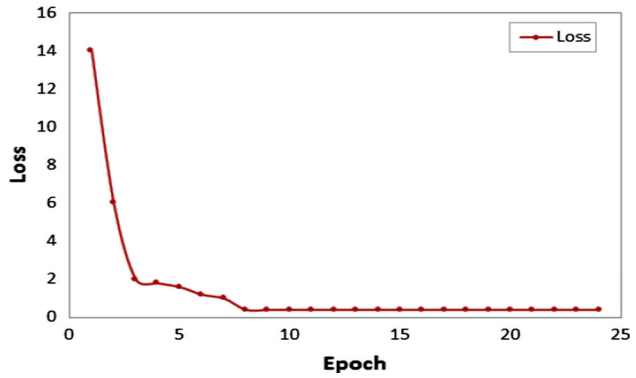

Fig. 4 Training loss vs Epochs

Table 3 Confusion matrix

True Class \ Predicted class	Normal	Dos	Probe	R2L	U2R	
Normal	13358	1	2	1	1	99.78%
Dos	2	9204	0	0	0	99.94%
Probe	5	0	2380	0	0	99.52%
R2L	11	0	0	200	0	90.11%
U2R	2	0	0	0	4	50%
	99.70%	99.73%	98.84%	99.01%	66.78%	
	Normal	DoS	Probe	R2L	U2R	

R2L and U2R attacks whereas for DoS and Probe attacks the number of samples is sufficient to obtain the desired training accuracy which improves the test accuracy.

The performance of the proposed model is compared with existing deep learning techniques like Convolutional Neural Network (CNN), Gated Recurrent Unit (GRU), and Long short-term memory (LSTM) based intrusion detection models. The performance of non-concatenated models is included in the experimental analysis. The results of CNN and GRU are considered as the results of non-concatenated systems as the results are obtained separately by applying the models without any feature fusion. Results clearly depict that the non-concatenated CNN and GRU model exhibits less performance than the proposed model for all the parameters.

The parameters like precision and detection rate (Recall) are considered for analysis, and it is depicted in Figs. 7 and 8. It is observed from the results the performance of the proposed model is better compared to other models. The maximum precision and detection rate attained by the proposed model indicates that all the normal and abnormal activities in the network are detected effectively. The average precision attained by the proposed model considering all the normal and attack categories is 93.6% whereas

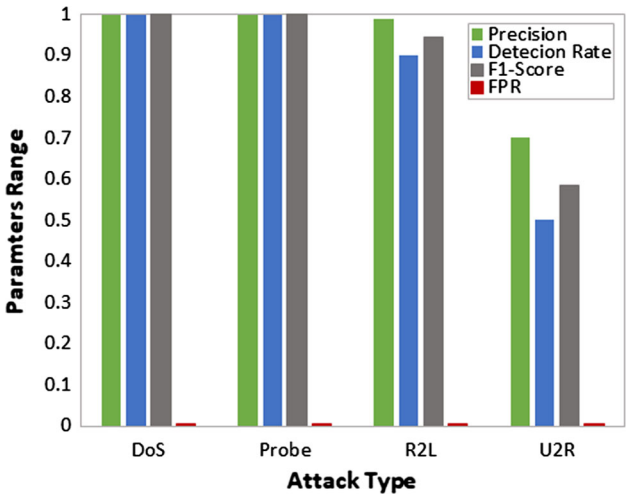

Fig. 5 Performance evaluation of proposed model

Table 4 Confusion matrix

True Class \ Predicted class	Normal	Dos	Probe	R2L	U2R	
Normal	26814	8	9	7	2	99.84%
Dos	0	18404	0	0	0	99.98%
Probe	15	0	4534	0	0	99.32%
R2L	44	0	0	310	0	78.54%
U2R	5	0	0	0	12	49%
	99.54%	99.89%	98.94%	95.84%	66.74%	
	Normal	DoS	Probe	R2L	U2R	

LSTM attains 89.1% and GRU attains 85.8% and 84% attained by the CNN-based detection model.

Based on the precision and detection rate values obtained in the previous analysis, the performance is measured in terms of F1-score and depicted in Fig. 9. The maximum F1-score attained by the proposed model indicates the maximum detection performance compared to existing deep learning methods. The overall accuracy based on the above parameters is calculated for the proposed model and existing models and depicted in Fig. 10. It is observed that the maximum accuracy is attained by the proposed model. 98.84% is the acquired detection accuracy of the proposed model whereas the existing techniques like LSTM, GRU, and CNN attain accuracy of 94.11%, 96.65%, and 97.07%, respectively. Due to efficient feature selection and concatenated process, the proposed model exhibits maximum accuracy compared to other models.

From the results, it can be observed that the proposed model efficiently detects attacks on the network and provides better detection rate and accuracy. The results were

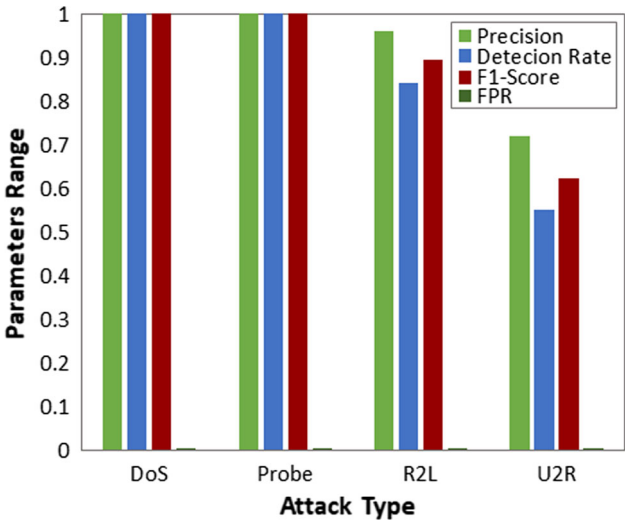

Fig. 6 Performance evaluation of proposed model

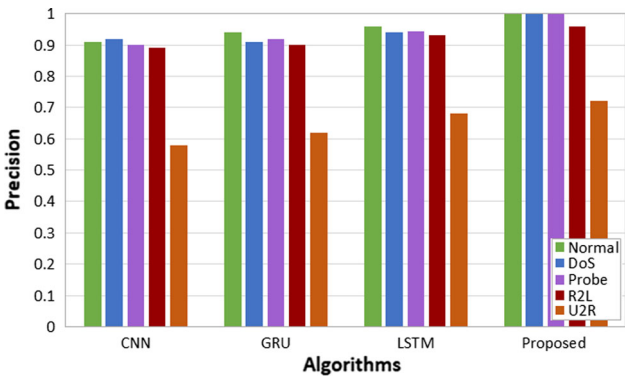

Fig. 7 Precision analysis

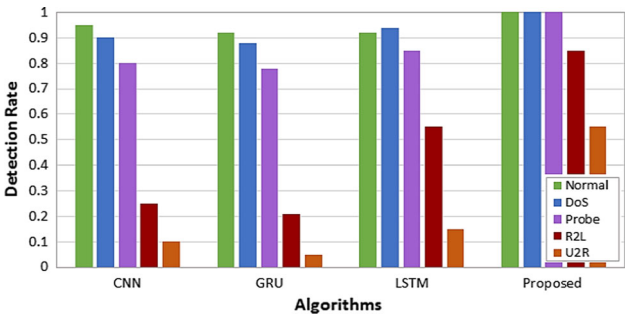

Fig. 8 Detection rate analysis

obtained for the standard dataset and the same performance can be expected in real-time smart grid data. The computational complexity of the proposed model is slightly above the mark than the existing techniques due to the initial parameters for two models used in the setup. However, for an industrial system intrusion detection model the proposed model is sufficient to detect the attacks efficiently. There may be a slight change in the detection performance due to

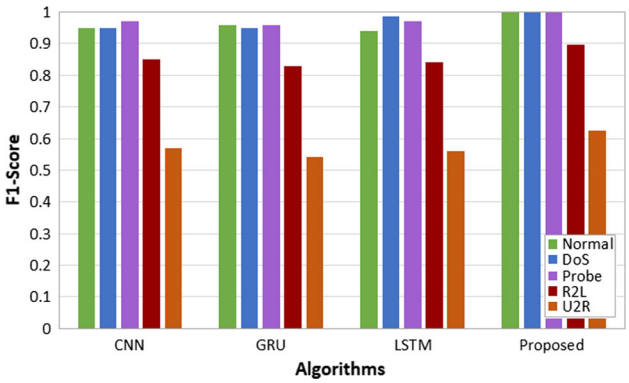

Fig. 9 F1-Score analysis

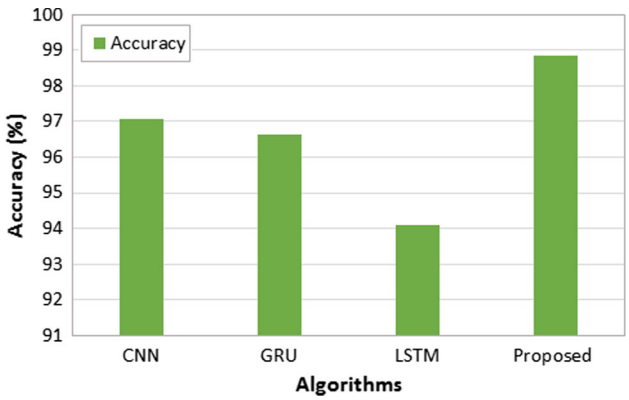

Fig. 10 Comparative analysis of Accuracy

environmental and system changes, which is the minor limitation of this paper.

5 Conclusion

This paper presents an efficient intrusion detection system for cyber-physical attack detection in the smart grid metering infrastructure. Cyber physical attacks on the SCADA systems are considered for the smart grid metering infrastructure and various types of attacks are identified. The attack types are related to standard benchmark dataset types and evaluation is performed to avoid real-time computational complexities. The NSL-KDD dataset is used for experimentation and the performance is evaluated in terms of accuracy, detection rate, precision, and false positive rate. Existing methods such as CNN, GRU, LSTM are compared with the proposed model and the results clearly demonstrate that the performance of the proposed model is superior to others. Further, this paper can be improved by focusing on the other parameters related to the grid environment.

Acknowledgements Sayawu Yakubu Diaba would like to thank the Evald and Hilda Nissi Foundation for awarding me scholarship.

Funding Open Access funding provided by University of Vaasa (UVA). No funding was received to assist with the preparation of this manuscript.

Declarations

Conflict of interests The authors declare that they have no conflict of interest to declare.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Cómbita LF, Cárdenas ÁA, Quijano N (2019) Mitigating sensor attacks against industrial control systems. IEEE Access 7:92444–92455

Dhaya R (2021) Light weight CNN based robust image watermarking scheme for security. J Inf Technol Digital World 3(2):118–132

Elnour M, Meskin N, Khan K, Jain R (2020) A Dual-isolation-forests-based attack detection framework for industrial control systems. IEEE Access 8:36639–36651

Feng C, Li T, Chana D (2017) Multi-level Anomaly Detection in Industrial Control Systems via Package Signatures and LSTM Networks. In: 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp 261–272

Ghosh S, Sampalli S (2019) A survey of security in SCADA networks: current issues and future challenges. IEEE Access 7:135812–135831

Gumaei A, Hassan MM, Huda S, Hassan MdR, Camacho D, Ser JD, Fortino G (2020) A robust cyberattack detection approach using optimal features of SCADA power systems in smart grids. Appl Soft Comput 96:106658

Haoxiang W, Smys S (2021) A survey on digital fraud risk control management by automatic case management system. J Electr Eng Autom 3(1):1–14

Homay A, Chrysoulas C, El Boudani B, Sousa MD, Wollschlaeger M (2020) A security and authentication layer for SCADA/DCS applications. Microprocess Microsyst 6:103479

Hu Y, Sun Y, Wang Y, Wang Z (2019) An enhanced multi-stage semantic attack against industrial control systems. IEEE Access 7:156871–156882

Jacob IJ, EbbyDarney P (2021) Design of deep learning algorithm for IoT application by image based recognition. J ISMAC 3(3):276–290

Kalech M (2019) Cyber-attack detection in SCADA systems using temporal pattern recognition techniques. Comput Secur 84:225–238

Keshk M, Sitnikova E, Moustafa N, Hu J, Khalil I (2021) An integrated framework for privacy-preserving based anomaly detection for cyber-physical systems. IEEE Trans Sustain Comput 6(1):66–79

Khan IA, Pi D, Khan ZU, Hussain Y, Nawaz A (2019) HML-IDS: a hybrid-multilevel anomaly prediction approach for intrusion detection in SCADA systems. IEEE Access 7:89507–89521

Pang Y, Xia H, Grimble MJ (2020) Resilient nonlinear control for attacked cyber-physical systems. IEEE Trans Syst, Man, Cybern: Syst 50(6):2129–2138

Paridari K, O'Mahony N, El-Din Mady A, Chabukswar R, Boubekeur M, Sandberg H (2018) A framework for attack-resilient industrial control systems: attack detection and controller reconfiguration. Proceedings of the IEEE 106(1):113–128

Qian J, Du X, Chen B, Qu B, Zeng K, Liu J (2020) Cyber-physical integrated intrusion detection scheme in SCADA system of process manufacturing industry. IEEE Access 8:147471–147481

Rakas SVB, Stojanović MD, Marković-Petrović JD (2020) A Review of research work on network-based SCADA intrusion detection systems. IEEE Access 8:93083–93108

Rodofile NR, Radke K, Foo E (2019) Extending the cyber-attack landscape for SCADA-based critical infrastructure. Int J Crit Infrastruct Prot 25:14–35

Sheng C, Yao Y, Fu Q, Yang W (2021) A cyber-physical model for SCADA system and its intrusion detection. Computer Netw 185:107677

Smys S, Vijesh Joe C (2021) Metric routing protocol for detecting untrustworthy nodes for packet transmission. J Inf Technol 3(2):67–76

Sun Q, Zhang K, Shi Y (2020) Resilient model predictive control of cyber-physical systems under DoS attacks. IEEE Trans Industr Inf 16(7):4920–4927

Wang C, Wang B, Liu H, Qu H (2020) Anomaly detection for industrial control system based on autoencoder neural network. Wireless Commun Mobile Comput 2020:3

Zhang F, Kodituwakku HADE, Hines JW, Coble J (2019) Multilayer data-driven cyber-attack detection system for industrial control systems based on network, system, and process data. IEEE Trans Industr Inf 15(7):4362–4369