IADF-CPS: Intelligent Anomaly Detection Framework Towards Cyber Physical Systems

Senthil Murugan Nagarajan\textsuperscript{a}, Ganesh Gopal Deverajan\textsuperscript{1,\textdagger}, Ali Kashif Bashir\textsuperscript{c}, Rajendra Prasad Mahapatra\textsuperscript{1}, Mohammed S. Al-Numay\textsuperscript{d}

\textsuperscript{a}School of Computer Science and Engineering, VIT-AP University, Amaravati, Andhra Pradesh - 522237, India
\textsuperscript{b}Department of Computer Science and Engineering, SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, Uttar Pradesh - 201204, India
\textsuperscript{c}Department of Computing and Mathematics, Manchester Metropolitan University, Manchester, United Kingdom
\textsuperscript{d}Department of Electrical Engineering, King Saud University, Saudi Arabia

Abstract

Cyber-Physical Systems (CPSs) becoming one of the most complex, intelligent, and sophisticated system. Ensuring security is an important aspect towards CPSs. However, increase in sophisticated and complexity attacks in CPSs, the conventional anomaly detection methods are facing problems and also growth in volume of data becomes challenging which requires domain specific knowledge that could be applied directly to analyze these challenges. In order to overcome this problem, various deep learning based anomaly detection system is developed. In this research, we propose an anomaly detection approach by integration of intelligent deep learning technique named Convolutional Neural Network (CNN) with Kalman Filter (KF) based Gaussian-Mixture Model (GMM). The proposed model is used for identifying and detecting anomalous behavior in CPSs. This proposed framework consists of two important process. First is to pre-process the data by transforming and filtering original data into new format and achieved privacy preservation of the data. Secondly, we proposed GMM-KF integrated deep CNN model for anomaly detection and accurately estimated the posterior probabilities of anomalous and legitimate events in CPSs.

Keywords: Intelligent System, Cyber Physical System (CPS), Kalman Filtering, Deep Learning, Anomaly Detection

1. Introduction

Cyber Physical Systems (CPSs) has an essential feature of connecting infrastructure and physical devices such as smart power grid and autonomous vehicles with various intelligent applications such as Industry 4.0 technologies, smart factories, energy sector, and intelligent transportation systems. CPSs is the next generation intelligent system that includes integration of communication, modern computing, and control technologies which improves the stability, safety, reliability, efficiency, and other performance analysis on real-operating systems \cite{1,2}. Nowadays, various researchers, government decision makers, technical staffs, industrial experts are working towards CPSs due to growth of numerous applications in international standards and critical infrastructures for smart cities developments. Furthermore, the defense authorities in various countries are highly dependent in the standard development of CPSs as defense systems such as naval vessels, unmanned grouped vehicles, and unmanned aerial vehicles are essential towards CPSs \cite{1,3}.

CPSs are exposed to various disturbances due to its complexity of intentional and unintentional events. Due to increase in cyber attacks, the behavior of CPSs has become more sophisticated. Meanwhile, new challenges and threats have been emerged in intelligent CPSs which varies from the existing problems. When internet connection is executed with the CPSs, there increases the major challenges related to privacy and security \cite{4,5}. The major reason behind this security issue is that the recent evolution of different hacking techniques that helps attackers to expose the data integrity and devices of CPSs. The privacy issue of CPSs is involved in compromising important information by using the active and passive attacks. The reliable information is sniffed from the public data by using the passive attacks while to gain access/modify or infer the private data, the active attacks are used \cite{6}.

Even with the advancement in research on security controls such as firewall, authentication, access control, encryption tools, and intrusion detection for the protection of CPSs, but still there is significant challenges occur in privacy and security due to complexity for monitoring the network and physical elements. Various research analysis have indulged in the development of integrity and confidentiality in CPSs \cite{7,8}. Confidentiality includes the ability to protect network and physical data from unauthorized users, whereas integrity refers to the ability to protect data from unauthorized changes.

Various privacy techniques are proposed recently for the protection of sensitive information, while different existing techniques venture to protect the valuable data but they are withhold to improve the performance optimization \cite{9,10}. CPSs includes the evolution of optimizing state dynamism and dynamic states...
as it is a feedback closed-loop system which essentially con-
sidering these evolution’s when applying for security and pri-
vacy preserving models. Various detection mechanism were
proposed for analyzing the protection from cyber attacks to-
towards CPSs. Data Mining (DM), Machine Learning (ML), and
statistical approaches were used widely for the development of
privacy preservation methods and intrusion detection system in
CPSs. However, the major problem in developing anomaly de-
tection method is that obtaining the relevant data and analyze it.
This work ensures that proposed system must not compromise
with various cyber attacks for the transformed data [11][12].

In this work, a anomaly detection method is proposed for
privacy preservation of CPSs named HDSCNN-KF that used
for the protection of original information and identifies the cy-
ber attacks efficiently in the CPSs. The proposed framework
is evaluated using UNSW-NB15 and cyber power data which is
available publicly. The performance of proposed work is com-
pared with other techniques in order to reveal the superiority
of HDSCNN-KF framework for detecting suspicious events and
preserving valuable data. The main contributions of this re-
search finding are as follows:

1. Developed and detection model for detecting anomaly
activities and threat behavior in CPSs.
2. Proposed an hybrid deep learning model based Siamese
Convolutional Neural Network with Kalman Filtering (HDSCNN-
KF) for improving the issue of over-fitting.
3. We evaluate the systems performance in terms of accu-
rency and computational time using two publicly available
datasets.

The rest of section in this paper is as follows. Section 2 dis-
cussed the recent methodologies and models for security anal-
ysis in CPSs. Section 3 presents the proposed framework for
detecting the cyber attacks. Section 4 discussed the evalua-
tion of the proposed model with comparison results. Section
5 concludes the overall work and suggestions towards future
enhancement.

2. Background and Related Work

This section discussed the anomalies and various methodo-
ologies for the privacy preservation in CPSs.

Authors in [13], analyzed privacy preservation using Supervi-
sory Control and Data Acquisition (SCADA) method for pro-
tecting the original data from revealing or being published by
un-authorized users. Authors in [14], discussed a new research
for alleviating illegal access of important and private data
that collected from the industrial systems. In this privacy pres-
erving technique, authors main goal is to transform, modify, hide,
and distribute information for the prevention of exposing dur-
ing processing of original data when using IDS [15]. Recently,
several methods based on privacy preservation have been de-
developed and these are classified for achieving data transforma-
tion which includes three main categories and these are, aggre-
gation, transformation, and generalization. The confidentiality
of the sensitive data is preserved usign the generalization tech-
niques and it is mapped by sensitive features for getting general
values [15][17].

Authors in [18], proposed a transformation technique for
replacing the original values with alternate values and data di-
imensionality is decreased using the projection methods. Fur-
thermore, authors presented that original data can be divided
into small proportions using the aggregation techniques and pri-
vate values are exchanged in each portion [19]. Authors in [20]
developed a secure SCADA system by using a layered tech-
nique with the various levels of IDS combination and moni-
toring behavior technique. This model has an disadvantage of
scalability issues with the substations of power plant. Authors
in [21] proposed a method to protect the power-scheme which
is used for altering the structure of data. An algebraic criteria is
included to check the power system to protect against various
attacks and making difficult for intruders to reach the system.
In this method, one type of attack only detected easily which
is major disadvantage. Authors in [22] proposed AE-based
solution for detecting different cyber attacks for the industrial
control networks. When internet is connected with control net-
works then there will be existence of various cyber attacks. The
attack sophistication is one of the CPS characteristics that re-
lected by this research problem. This problem is translated into
classification task for machine learning model. Authors in [23]
analyzed the recent advancements of deep learning techniques
that are applied for enhancing accuracy and security problems
in Android-based CPS. The sensitive applications behavior is
detected by implementing deep learning techniques. Typical
devices of Internet of Things (IoT) and Cyber Physical System
(CPS) is operated by using ResNet for resource limitation pur-
pose. Authors in [24] presented the security related issues in
IoT-based CPS. Overflow of failures can occur if even a failure
or small fault inside the CPS’s interdependent networks. Au-
thors focused on to decrease this overflow failures and reduced
the losses. The size of entire network function for the compo-
nent is calculated during this time of attack. Table [1] shows the
research works on cyber physical system.

Authors in [31] proposed a beta distribution for responding and
detecting various malicious activities which occur in CPSs. This
model obtains high response and detection rate. Some re-
searchers used KF-method for protecting the nodes of CPS with
the help of predefined model and the parameters are estimated
based on features linear relationship. Authors in [32] proposed
a framework for safeguarding the privacy of SCADA data based
on a clustering concept that modifies sensitive values and par-
titions the original data based on traffic types. Authors in [33]
described how the VMD method can be used to detect false data
injection attacks (FDI) in estimation problem; nevertheless, the
suggested technique does not detect other types of attacks, such
as communication failure-attacks and coordinated-attacks. In
this work, the VMD technique is used by the authors for devel-
oping the CPADS and its characteristics includes various events
such as physical disturbances normal operation, and cyber at-
tacks based on WAPS. This model has the capability of extract-
ing optimal features to train the model. Furthermore, no prior
work in detecting coordinated cyber attacks in CRAS cyber se-
Table 1: Related Works based on Cyber Physical System

<table>
<thead>
<tr>
<th>Reference</th>
<th>Cyber Attacks</th>
<th>Attack-Targets</th>
<th>Dataset Used</th>
<th>Scenario</th>
<th>Performance Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>[25]</td>
<td>Inference Attacks</td>
<td>Communication links and Sensor Nodes</td>
<td>UNSW NB15 and power system dataset</td>
<td>Smart Power Networks</td>
<td>92% of accuracy is obtained</td>
</tr>
<tr>
<td>[22]</td>
<td>Probe attack, DoS attack, and unauthorized access attacks</td>
<td>Communication links and controllers</td>
<td>NSL-KDD dataset</td>
<td>Industrial Control Networks</td>
<td>97.8% of accuracy is obtained</td>
</tr>
<tr>
<td>[26]</td>
<td>Phasor measurement</td>
<td>Sensor nodes and Controllers</td>
<td>Simulated IEEE 9 bus</td>
<td>Smart grids</td>
<td>94.1% is obtained</td>
</tr>
<tr>
<td>[27]</td>
<td>Unspecified cyber attacks</td>
<td>Sensor nodes and Actuator nodes</td>
<td>Simulated data from gas turbines</td>
<td>Smart grids</td>
<td>FPR rate of 0.000006 is obtained</td>
</tr>
<tr>
<td>[28]</td>
<td>Fuzzy attack, data spoofing, and exploits attacks</td>
<td>Actuator nodes, Communication links, and Sensor nodes</td>
<td>Car hacking dataset and UNSW-NB15</td>
<td>Internet of Vehicles</td>
<td>99% of accuracy is obtained</td>
</tr>
<tr>
<td>[29]</td>
<td>Replay attacks</td>
<td>Controllers</td>
<td>118 bus systems</td>
<td>Smart grids</td>
<td>MAPE obtained as 3.51%</td>
</tr>
<tr>
<td>[30]</td>
<td>False data injection attacks</td>
<td>Actuator nodes and Sensor nodes</td>
<td>SWaT dataset</td>
<td>Water Treatment Plant</td>
<td>89% of accuracy is obtained</td>
</tr>
</tbody>
</table>

Authors in [34] have proposed IDS based on statistical and rule mining methods. Authors defined some set of rules for particular features of network such as flow layer, inter-flow layer, and packet layer. Furthermore, systems historical behavior is used for the statistical analysis where it is conducted for long-term metrics extraction. Detection of trends and anomalies which do not fit with the observed were processed using these metrics. Authors in [35] analyzed the scenarios of false-data injection attacks by studying Linear Parameter Varying (LPV) CPS. Based on state-space model of LPV, a system model is obtained. The expected measurement of future sensor can be predicted using this model. Based on specific threshold, anomalies are detected. Authors in [36] used Long-Short Term Memory (LSTM) with RNN for detection of cyber attacks in CPS by normal behavior modeling of system. Then, calculation of values by RNN prediction is made with the differences between ideal and actual sensors. Furthermore, authors detected the small deviation using cumulative sum method correspond to the anomalies. Authors in [37] proposed sequence-to-sequence encoding-decoding method with the help of RNN. Next values are predicted by encoding the time series data while the future operational data is predicted using decoding with the help of attention method. The anomalies are determined using the difference between the actual and predicted data.

3. Problem Definition

Ensuring security is an important aspect towards CPSs. However, increase in sophisticated and complexity attacks in CPSs, the conventional anomaly detection methods are facing problems and also growth in volume of data becomes challenging which requires domain specific knowledge that could be applied directly to analyze these challenges. In order to overcome this problem, various deep learning based anomaly detection system is developed.

4. Proposed Anomaly Detection Mechanism

In this section, we present a system model for detecting anomaly activities and threat behavior in CPSs with an aim of achieving the two fundamental goals such as security and privacy. To accomplish this goal, we proposed a hybrid deep learning model based Siamese Convolutional Neural Network (SCNN) with Kalman Filtering (HDSCNN-KF) in order to improve the issue of over-fitting and increase anomaly detection accuracy in CPSs. System model for anomaly detection is shown in Fig. 1.
4.2. Data Acquisition

With advancement and development of computing and communication technologies, CPS becomes large and generates huge amounts of data emerging from various CPS components including massive physical sensor devices such as 3D stereoscopic camera and the LiDAR sensors in automotive applications, mobile applications such as mobile intelligent robots and robotics system and networking devices deployed at each CPS layer as shown in Fig. 3. Due to the complexity of a large-scale CPS and the huge quantity of data produced, SCNN-KF approaches are used to identify and detect anomalous behavior in CPS. For this, we collect data from three layers: physical layer, network layer and mobile application layer. From the physical layer, we collect safety-critical data from sensor devices deployed in the physical world for accessing personal and environmental characteristics. From the network layer, data information related to data packets and network surroundings prone to threats are collected. This information includes header information, SNR, state information of network channel, packet drop rate, and mean of round trip time physical system to cyber system. At the mobile application layer, data related to the computing system responsible for damaging hardware and software systems are collected including CPU utilization, memory utilization, file storage credentials and command execution.

4.3. Data Preprocessing using GMM

Since, CPSs contain a variety of characteristics features from physical and cyberspace including sensor data feature, network data feature and mobile app feature so it is necessary to pre-process and fuse this data before constructing the proposed framework. The physical space is the collection of internal data from the various sources where as cyberspace is the world of information gathered from the source through the internet. The GMM is probability based approach that deals with the representation of data from several distributions.

Cyber power systems have multivariate characteristic features since they produce multiple variables from physical, cyber and network space. The GMM consists of various essential elements that perform data fusion of features acquired from physical, network and mobile application systems and each feature is considered as one-dimensional. The Probability Density Function (PDF) of each feature is computed using given Eqn. 1:

\[ p(\mathbf{u}) = \frac{2}{\sqrt{2\pi}v} e^{-\frac{(\mathbf{u}-\mathbf{x})^2}{2v^2}} \]  

Where, expectation and variance of \( \mathbf{u} \) is represented by \( \mathbf{x} \) and \( v^2 \) respectively whereas the mathematical notation of \( \mathbf{u} \) is \( \mathbf{u} \sim F(\mathbf{u} | \mathbf{x}, v) \). since the data feature has A attributes therefore their PDF can be computed using Eqn. 2 and 3.

\[ p(\mathbf{u} | \mathbf{\omega}) = \sum_{a=1}^{A} \omega_a F(\mathbf{u} | \mathbf{\omega}_a) \]  

\[ F(\mathbf{u} | \mathbf{\omega}_a) = F(\mathbf{u} | \mathbf{\omega}_a, v_a) = \frac{1}{\sqrt{2\pi}v_a} \frac{1}{|v_a|^{\frac{a-a_0}{2}}} e^{-\frac{(\mathbf{u}-\mathbf{x}_a)^2}{2v_a^2}} \]

Where \( \omega_1, \ldots, \omega_A \) represent attribute weight that identify proportion of mixing and each \( \omega \) denotes Gaussian mixing (GMM)
parameter that describe attributes. The GMM weight summation should be constrained with following probability condition as given in Eqn. 3:

\[ 0 < \omega_a < 1, a = 1 \ldots A \quad \text{and} \quad \sum_{a=1}^{A} \omega_a = 1 \]  

(4)

To estimate the GMM parameters \( (\omega) \) as given in Eqn. 3 the Expectation Maximization (EM) method is used. Given the \( U \) data, the Maximum Likelihood Estimation (MLE) \( p(u|\omega) \) is calculated using Eqn. 5 of the GMM parameter of model:

\[ \omega^* = \frac{\text{argmax}_\omega p(u|\omega)}{\omega} \]  

(5)

Expectation (E) and Maximization (M) are the two major stages in continually estimating the model parameters. The Expectation step computes the posterior probability based on the current state of the parameter, and the Maximization-step finds the optimum parameters in relation to the MLE computed in the previous step. The parameters of the EM are chosen carefully, with the number of attributes (A) set to 3 for grouping normal, physical and cyber attack data points and the number of iterations set to 1000 to end the model’s execution. Its three basic parameters (mean \( x \), variance \( v^2 \), and weight \( \omega \) are determined using equations 6, 7, and 8.

\[ x = \sum_{a=1}^{A} p(a|u_a, \omega) \frac{u_a}{p(a|u_a, \omega)} \]  

(6)

\[ v^2 = \frac{\sum_{a=1}^{A} p(a|u_a, \omega) \frac{u_a}{p(a|u_a, \omega)}}{\sum_{a=1}^{A} p(a|u_a, \omega)} \]  

(7)

\[ \omega = \frac{1}{A \sum_{a=1}^{A} p(a|u_a, \omega)} \]  

(8)

4.4. Intelligent Anomaly and Threat Detection

In this detection model we use a hybrid approach of combining SCNN and Kalman Filter to anomaly and threat detection in CPS. For this first we designed SCNN based few shot learning models that process the sensor, network and mobile app data obtained from physical, network and mobile application and fused together by GMM. SCNN analyzes and processes each data and determines whether the data are anomalous or normal. The detected anomalies or threatening activities in CPS are excluded from CPS and normal data are fed into Kalman Filter that use a failure detector for further analysis and anomaly detection. KF then analyzes and processes the data and identifies the anomalies that are missed by SCNN. Finally, in order to achieve a greater degree of consistency, contaminated or anomalous data are removed, and the remaining data that are recognized as normal are fused together.

4.4.1. Anomaly Detection using Siamese Convolution Neural Network

The proposed SCNN based anomaly detection is intended to address the problem of insufficient labeled anomaly samples. Unlike traditional classification models, our SCNN based anomaly detection calculates the distance between input samples in terms of their optimal feature representations rather than simply predicting the class for each input sample data. In particular, SCNN is built to deal with the few-shot learning issue, allowing new classes to be recognized even when just a few samples are available. Figure 5 depicts the SCNN based anomaly detection architecture for anomaly identification in CPS.

As illustrated in Fig. 4 two identical CNNs are used to train the SCNN based deep learning model in which two input sample data sets including query set and support-set are fed one is taken from each class. Original features are reduced using the relative feature representation technique and it is assisted for neural network to overcome the over-fitting problem. As a result, improving detection performance. To extract feature vector embedding, the Siamese network employs a combination of two identical convolution layers and sampling layer. The distance between these two feature vector embedding will be computed throughout the testing phase to determine if these two input samples belong to the same class.

The two input sample sent to SCNN are \( s_i \) and \( s_j \) and the feature vectors of these two samples using SCNN is computed using Eqs. 9 and 10.

\[ f(s_j) = \text{SCNN}_{\text{emp}}(s_j, \alpha_{\text{encode}}) \]  

(9)

\[ f(s_j) = \text{SCNN}_{\text{emp}}(s_j, \alpha_{\text{encode}}) \]  

(10)

Where, \( \alpha_{\text{encode}} \) represent encoding parameter of SCNN. Here, the encoding-parameter is used for encoding of the incoming data into the SCNN. The pair-wise Euclidean distance which
is represented as $Dis$ is used to compute the distance between two feature vector embedding from two input samples $s_i$ and $s_j$ using Eqn. [11]

$$Dis(f(s_i), f(s_j)) = ||f(s_i) - f(s_j)||^2$$

(11)

Finally, SCNN output is produced based on the last two layers i.e. soft max (SM) and fully connected (FC) layer as given in Eqn. [12]

$$Prob(s_i, s_j) = SM(FC(Dis(f(s_i), f(s_j))))$$

(12)

Where, SM(*) represent SoftMax layer function and $FC(*)$ represents fully connected layer function, and $Prob(s_i, s_j)$ represents the probability function to indicate whether $s_i$ and $s_j$ associated with same or different class.

Three losses are addressed in our cost designed function to guarantee prediction accuracy rate and error rate for anomaly detection from huge volumes of CPS data with few labeled samples. The loss of transformation $L_{rela}$ is provided in the relative-feature format. During the SCNN encoding process, the loss of encoding $L_{enc}$ is calculated to quantify the variance between the converted relative-features and extracted feature vectors. $L_{pred}$ is a triplet loss depending on the distances between positive and negative samples, and the reference sample.

The transformation loss in relative-feature format may be defined and computed for each input sample $s_i$ as given in Eqn. [13]

$$L_{rela} = \frac{1}{N} \sum_{n} -\log\left(\frac{e^{-\text{dis}(f(s_i), m_i)}}{\sum_{k=1}^{K} e^{-\text{dis}(f(s_i), m_k)}}\right)$$

(13)

Where, $\text{dis}(f(s_i), m_i)$ denotes the Euclidean distance. For relative-feature formats, $m_i$ is obtained by averaging the sample instances of class $c$, while $m_c$ is computed for the equivalent representations in each training phase.

The encoding loss is intended to reduce the number of feature vectors while preserving the essential information of features in the original data. For given input sample $s_i$ with a probability distribution, we compute encoding loss $L_{enc}$ using Kullback–Leibler (KL) divergence by using Eqn. [14]

$$L_{enc} = E_{s_i} [\sum_{s_j} p(s_i | f(s_j)) \log \left( \frac{p(s_i | f(s_j))}{q(s_i | f(s_j))} \right)]$$

(14)

Where, $p(s_i | f(s_j))$ represents the true distribution of the input sample data, while the computed distribution is represented as $q(s_i | f(s_j))$ and may be regarded as an approximation to $p(s_i | f(s_j))$.

For prediction loss computation we considered distances between the reference sample and the positive and negative samples. The precise computation may be expressed as Eqn. [15]

$$L_{pred} = \max ((\text{Dis}(f(s'_f), f(s''_f))) - \text{Dis}(f(s'_f), f(s''_f)) + \gamma, 0)$$

(15)

Where, $s'_f$, $s''_f$, and $s''_f$ represent reference, positive, and negative samples. $\gamma \in (0, 1)$ is a variable used to fine-tune error rate in anomaly detection and during the training phase, it is usually set as $\gamma > 0.5$. Based on this adversarial design, the maximum function is utilized to guarantee a minimum loss for $L_{pred}$, allowing the reference sample to be more equivalent to the positive sample instances than the negative one.

4.4.2. Kalman Filter Assisted Anomaly Detection

After processing data in SCNN based anomaly detection, all the data that are anomalous were excluded and normal data are fed into KF for further processing. The initial states are calculated based on prior measurements obtained during the system’s internal operations, and the state-space model is used to describe the CPS dynamics. Algorithm analysis for anomaly detection is given in Algorithm 1 and 2.

Algorithm 1 Training Phase

**Input:** A set of anomalies data samples: $Dt_A = \{(s_A, d_A)| i = 1, 2, 3,...N_A\}$

A set of normal data samples: $Dt_{norm} = \{(s_{norm}, d_{norm})| i = 1, 2, 3,...N_{norm}\}$

A set of query data samples $Dt_Q = \{(s_{Q_i}, d_{Q_i})| j = 1, 2, 3,...N_Q\}$

**Output:** Trained HDSCNN-KF model

1. for each Normal data sample $d$
2. $\omega = \{x, \sigma^2, \alpha\} \leftarrow$ Compute GMM Parameters
3. Compute PDF based on GMM parameter
4. Initialize the parameter $\gamma$ and threshold value for loss ($\tau$)
5. while $OPT_{SCNN} > \tau$
6. for each normal anomalies and query data samples $d$
7. To construct support dataset, select $c$ class with $s$ samples from $Dt_{norm}$ and $Dt_A$
8. To construct query dataset select $c$ class from $Dt_Q$
9. for each $s_i$ in support dataset $d$
10. transforming $s_i$ to
11. Compute transformational loss
12. Transforming $s_i$ into feature vector embedding with SCNN encoder
13. Compute encoding loss
14. Choose reference data sample $s'_f$ and predict $d'_f$
15. Choose positive and negative data sample $s''_f$ and $s''_f$ and predict $d''_f$ and $d''_f$
16. Compute Predictive loss
17. Update SCNN to minimize the optimize cost function $OPT_{SCNN}$
18. for each normal data sample $d_i$ from $i = 1, 2, ..., N_{norm}$
19. Compute posterior probabilities for all data sample using KF
20. Set upper and lower boundaries of normal posterior probability $d_{low}$ and $d_{up}$

5. Result Analysis

5.1. Dataset Description

In this section, we present performance evaluation of proposed architecture. The performance measurement of proposed work is carried out using two publicly available standard datasets
Algorithm 2 Testing Phase

**Input:** Set of anomalies data samples

**Output:** Normal and anomalous data, then exclude anomalous data

1. for each Normal data sample do
2. Compute GMM parameter
3. Compute PDF based on GMM parameter using Eqn. 3.
4. Classify the testing dataset into normal and anomalous data using SCNN trained model
5. Fed normal data from SCNN into KF
6. Compute posterior probabilities for all data sample using KF
7. for each test data sample do
8. if $d_{i}^{post} < d_{low}^{post} || d_{i}^{post} > d_{upp}^{post}$ then
9. Set label ← cyber-attack
else
10. Set label ← normal

includes Power System dataset and industrial UNSW-NB15 dataset, with various kinds of feature characteristics, namely continual or categorical, are chosen for use in the experimental research. The Power System dataset includes 37 scenarios with multi-class categories, including normal actions (8), meddling actions (28) and no actions (1). The industrial based UNSW-NB15 dataset includes both normal and attack records that are up to date. It has a data volume of around 100 GB and high dimensional observational data instances of about 2,540,046 since it has labeled class 48 feature characteristics. This dataset has a velocity of around 5 to10 MBPS as it travels between different network hops to precisely mimic some genuine network surroundings and it has 10 different classes, one class specifies normal and remaining nine distinct class specifies security events.

5.2. Experimental Setup

The anticipated HDSCNN-KF structure is written in R language and executes on Windows 10 operating system with an Intel 7 CPU processor and primary memory storage of about 16 GB. Sample sizes of 400,000 that includes normal data and anomalous data which are chosen randomly from each datasets for the studies during training and testing stages. The proposed HDSCNN-KF architecture and other methods’ performances are achieved by averaging the five-fold cross-validation outcomes in order to properly evaluate their accuracy without discrimination towards normal or malevolent classes.

5.3. Performance Evaluation

To validate the system, we used stochastic gradient descent (SGD) as the training optimizer model. To examine the training process in the experiment, we set the learning rate to 0.1 and iterated 800 times. The transformation loss, encoding loss, and predictive loss computed in each iterative cycle for UNSW-NB15 dataset as shown in Figure 5. The total performance of the three losses falls quickly and then becomes reasonably steady. It is evident from Fig. 5(a) and 5(c), the error margins of transformation loss and predictive loss vary dramatically throughout the learning process, while in Fig. 5(b) demonstrate that the error margin of encoding loss decreases rapidly and stabilizes after 200 iterative cycle. This training outcome indicates our model’s adaptability and appropriateness for few-shot learning.

![Efficiency Evaluation for Training Phase: (a) Transformation Loss (b) Encoding Loss (c) Predictive Loss](image-url)
sification result our model classification result with five different approach including principal components analysis (PCA), computer vision technique (CVT), naïve base (NB), random forest (RF), filter based support vector machine (F-SVM). The comparative analysis of all the six approaches in graphical format is shown in Fig. 5. In Fig. 5 the significant difference in data distributions shows the dataset imbalance as well as the associated feature characteristics. In other words, the number of normal data samples is far more than the number of malicious or attacked data samples. Form Fig. 5 it is observed that our proposed HDSCNN-KF perform better in clustering. The higher the performance of clustering, greater will be the impact of feature extraction. Furthermore, when compared to the other five techniques, the method produces a clear clustering result with minimal overlaps between characteristics in two distinct groups. This result demonstrates the efficacy of proposed HDSCNN-KF with reduced dimensionality and preserving important feature information throughout the learning phase. Furthermore, the effectiveness of our proposed HDSCNN-KF for detecting anomalies is demonstrated based on real-world attack scenario in CPS. For this comparative experiment was carried out using UNSW-NB15. We compared detected anomalies and true attacks in accordance with network throughput (bytes per second) captured in real-world CPS.

With our proposed work we efficiently detect different types of cyber attack for different set of feature attribute percentage (A=25%, 50%,75%,100%) represented in Fig. 6. For both industrial based UNSW-NB15 and Power System dataset, detection rate (DR) proposed work is progressively rise for attribute set A=25% and A=50%, while graph plotting are almost same for attribute set A=75% and A=100%. The improved DRs are due to higher fitting parameter properties of the GMM, SCNN and KF, which can distinguish between normal and cyber-attacks posterior bounds. Fig. 6(a) show that when A=75%, the proposed method can detect normal data with a DR of around 98.90% and a false negative rate of 1.10% using the Power System dataset. Furthermore, with same attribute A=75%, the majority of the cyber-attacks in this dataset can be identified with a DR ranging from 91.56% to 99.82%. When using UNSW-NB15 dataset with attribute set of A=75%, the proposed method can identify most of the cyber-attack class type, as shown in Fig. 6(b), with DRs ranging from 79.98% to 99.35% and Q=75%. Furthermore, normal data are recognized with about 94.90% accuracy and just about 5.10% false negatives.

From the result outcome depicted in Fig. 7 it is observed that proposed HDSCNN-KF framework outperforms as it combine the functionality of three different approach including GMM, SCNN and KF. The combination of three approaches properly limits the normal region which can help in identifying different types cyber-attack efficiently. The proposed work use GMM show better result with attribute percentage set of A=75% and A=100% because we pre-process data using GMM that can fit the lower and upper and lower limits of data features with probability distributions and compute the PDF of each feature vector accurately. As a result, suspicious data may be accurately identified based on the normal profile borders, where malicious data are identified as top and bottom outlier. Since we pick 75% of all features, the proposed framework performance is substantially enhanced.

According to the experimental findings as shown in Fig. 7 we demonstrated that our proposed identified various types of cyber attacks in minimum computational time than other peer-
ing methods including principal components analysis (PCA), computer vision technique (CVT), naive base (NB) classifier, random forest (RF) classifier, filter based support vector machine (F-SVM) classifier. From Fig. [7] we observe that our proposed HDSCNNKF framework takes 65 s to train 11,000 observations while for same observational instances other existing approach takes on an average 70 to 80 seconds.

6. Potential Applications and Limitations of Proposed Model

There are various potential applications can be applied using the proposed work such as military applications where securing confidential information, medical applications to transfer the health records securely to the patients or practitioners, and in smart cities to obtain the secure transmission of data through various sources.

The major limitations of this proposed model is that the huge data processing can be less efficient but with some improvements in the model can overcome this limitation. Furthermore, reduction in performance when new anomalies are identified.

7. Future Research Challenges

1. We have to implement the model with various benchmarks in CPS to analyze the performance of the proposed model.
2. For a certain period of time, the sudden change in false data injection attack cannot be identified by the model which should be addressed in future research.
3. To avoid catastrophic events, real-time response from CPS for faults or attacks immediately could be a challenging task where the response time must be improved and can be worked as a future direction.

8. Conclusion

In this paper, we presented the HDSCNN-KF to cope with the limited labeled and unbalanced datasets produced in real-world CPS for intelligent anomaly and threat detection in order to improve CPS security and protection in more intelligent way. The proposed infrastructure is built on three essential components. First, we use GMM with attempts to combine various data feature characteristics from physical, network and mobile apps with different distributions into a single value, which reduces processing time and aids in the preservation of information in the cyber physical system. Second, instead of providing the prediction result directly, we use encoded convolution neural network (CNN) based on Siamese network architecture that quantify the separation between input data samples in relative feature format in optimized form. To enhance training efficiency we designed cost optimization function that include three losses: loss occurs in transforming feature into its relative format, loss during encoding process of SCNN and the predictive loss occurs while computing the distances between three different data sample including reference and positive and negative. Third, we use KF, which can effectively match the dynamics of the CPS by accurately estimating and controlling them in order to identify outliers as anomalies and cyber attacks. The experimental findings of the proposed method can detect normal data with a DR of around 98.90% and a false negative rate of 1.10% using the Power System dataset. The proposed work use GMM show better result with attribute percentage set of A=75% and A=100%. Our proposed framework identifies different types of cyber attack very efficiently and successfully. In future research, we will perform additional assessments in various scenarios to enhance the algorithm’s accuracy and efficiency. In addition, we may use variation methods such as PCA and ICA for dimensionality reduction that can help in improving present work performance.

9. Acknowledgment

Research Grant: Mohammed Al-Numay acknowledge financial support from the Researchers Supporting Project number (RSP-2021/150) King Saud University, Saudi Arabia.

References


