Intrusion detection system for cloud environment based on convolutional neural networks and PSO algorithm

Gnanam Jeba Rosline¹, Pushpa Rani²
¹Department of Computer and Software Engineering, Mother Theresa Women’s University, Kodaikanal, India
²Department of Computer Science, Mother Theresa Women’s University, Kodaikanal, India

ABSTRACT

Authentication of clients and their applications to cloud services is a major concern. Network security and the identification of hostile activities are greatly aided by intrusion detection systems (IDS). In general, optimisation strategies can be applied to improve IDS model performance. Convolutional neural networks (CNN) and other deep learning (DL) algorithms is utilised to enhance IDS’s capability to identify and categories intrusions. IDSs can identify prior attacks, adapt to changing threats, and minimise false positives by utilising these strategies. In this work, a lightweight CNN is proposed for intrusion detection in cloud environment. The main contribution of this research is to use particle swarm optimization (PSO), a metaheuristic algorithm to find the CNN’s optimal parameters that comprise the number of convolutional layers, the size of the filter utilized in the convolutional procedure, the number of convolutional filters, and the batch size. Heuristic-based searches are useful for solving these kinds of problems. The experimental outcomes demonstrate that the proposed method reaches 91.70% of accuracy, 91.82% of precision, 91.99% of recall and 91.90% of F1-score. Cloud providers can gain from improved security measures by incorporating the proposed IDS paradigm into cloud settings, thereby minimizing unauthorized access and any data breaches.

Keywords:
Cloud environment
Convolutional neural networks
IoT and authentication
Intrusion detection
Particle swarm optimization

1. INTRODUCTION

Today’s cloud computing (CC) services offer a variety of capabilities, including massive networks and resource pooling [1]. The consumers of the cloud are provided with a platform, services, applications, and infrastructure by cloud providers [2]. On the user’s end merely being able to use the CC system’s user interface is necessary. Network resources like networks, data centres, hardware, software, and utilities are made available on demand through the CC network access model [3]. As a result, CC is a promising technology that provides a number of features such remote data access, storage, and accessibility [4], [5]. Due to its unique qualities, like availability, scalability, and self-services, it considerably lowers expenses. This model is made up, according to the NIST, of three cloud service models [6], [7]. Public clouds are the most popular since they are used by both businesses and private clients. With an exponential rise in clients, cloud service providers and potential hazards that target both users and providers of cloud technology is now expanding quickly. This leads to numerous security concerns involving availability, integrity, and confidentiality when users upload sensitive data to cloud storage. Additionally, the continuous delivery of cloud services without flawless security increases the risk of intrusion [8], [9]. Numerous sectors have

Journal homepage: http://ijee.cs.iaescore.com
adopted CC because of its inherent benefits, including scalability and flexibility. But even with these benefits, security issues continue to be a major obstacle for cloud service providers [10], [11]. As smart farming requires greater protection for the data processing a private cloud is opted.

In general, CC confronts a number of security challenges that slow down the adoption of cloud infrastructure [12], including regulation, data destruction in the cloud, and security concerns [13]. The sensitivity of users in the smart farming systems is one of these challenges. There have been many solutions created and used to protect applications, information, and cloud-based environments against intrusions such firewall and virus attacks, but they still have to be enhanced [14]. Then, intrusion detection is a group of cutting-edge technologies that detect unwanted actions to improve cloud security.

To prevent the release of confidential information or important company data due to unauthorised access, there are many different types of intrusion detection systems (IDS) and solutions available on the market. When applying countermeasures for effective prevention, software and appliances for intrusion detection should take into account the many offensive strategies that might be employed. Illustrations of network intrusion tactics include the use of vulnerability code injection, network flooding or overloading, or switching to attack additional systems on networks once a first host has been compromised. IDS security functions in conjunction with authorization access ability and substantiation, which measured resistance to intrusion using a dual line of resistance. In essence, preparation for intrusion detection is based on a number of key principles. It will be essential to have awareness of potential intrusions, avoid latent incursions, be aware of past intrusions, and take action in reaction to an intrusion. Understanding the many types of intrusions that have occurred or have been performed in the past is, in general, the first step in developing effective detection algorithms that can enhance detection and prevent future attacks. The remainder of the paper is organized as related works are discussed in section 2; proposed convolutional neural network (CNN) architecture and optimization using particle swarm optimization (PSO) are explained in section 3; experimental results are revealed in section 4 with concluding remarks and future perspectives in section 5.

In recent years, academics have looked into ways to improve IDS performance using machine learning (ML) and deep learning (DL) techniques. Large-scale data analysis and precise prediction are capabilities that ML and DL algorithms have shown. In order to detect network intrusions, Faker and Dogdu [15] proposed two methods that combine a deep neural network (DNN) with ensemble methods, random forest (RF) and gradient boosting tree (GBT), using training data from datasets. To carry out flow distribution probability (FDP) outlier detection, Djenouri et al. [16] presented a k-nearest neighbor algorithm for distance-based outlier detection. By using the artificial neural network technique for anomaly classification, the author assessed the dataset’s effectiveness and obtained an accuracy of 76.96% [17]. Recurrent neural network (RNN) and CNN for intrusion detection in cloud-based services are proposed in [18], [19]. The moth-flame optimizer (MFO) algorithm was used by Alazab et al. [20] to create an IDS technique. The evaluation revealed that using the MFO increased the classification accuracy. The bat algorithm was used by Zhou et al. [21] as a feature selection method for creating an IDS. It was assessed using RF. Several metaheuristic approaches are adapted for IDS applications [22]-[24]. Mohan et al. [25], describe five optimisation strategies are applied to the cluster based on K-means and K-nearest neighbor methods, and the outcomes are compared. For the security of the internet of things (IoT), Douiba et al. [26] suggested an optimised IDS combining gradient boosting and decision tree (DT). In order to safeguard industrial IoT (IIoT) edge computing, Mohy-eddine et al. [27] proposed an IDS model employing ensemble learning. Recently, Verma and Ranga [28] examined different ML strategies to find a classification algorithm for securing the IoT. An IDS was recommended by Attou et al. [29] to protect the cloud environment against intrusion. To improve the detection of anomalies, they combine graphic visualisation with RF classifier. The main part of this work is to introduced a lightweight CNN model for IDS in cloud environment and the parameters of CNN are optimized using PSO algorithm to enhance the performance of intrusion detection.

2. PROPOSED ARCHITECTURE

This section talks about the CNN architecture as well as the training dataset. Understanding the network traffic and data properties of the experimental dataset in greater detail is essential prior to model training. Furthermore, this section describes pre-processing, PSO algorithm, proposed CNN architecture, and optimization of CNN architecture using PSO.

2.1. Dataset

CICIDS2018 dataset is created by the Canadian Government’s Communications Security Establishment (CSE) and Canadian Institute for Cybersecurity (CIC) and Amazon Web Services (AWS) [30]. Among the freely accessible intrusion detection datasets, it is the most recent, comprehensive, and significant dataset. Comma-separated values (CSV) files contain both malicious and benign traffic. The collection has a
total of 10 files, which together constitute 6.41 GB [30]. There are 16,233,002 total datasets in the CSE-CIC-IDS2018. A stream of packets and 83 data properties, including duration, packet count and bytes are included in the dataset. In this dataset, a label designating whether network traffic is of the attack or benign categories is the final component of each sample of data. The assault type is separated into an entire of 14 different sorts of attacks by six categories. The number of traffic types in CICIDS2018 dataset are shown in Figure 1.

![Figure 1. Number of network traffic in CICIDS2018 dataset](image)

2.2. Pre-processing

Some datasets may have contained features or outliers which are unusable for training because of the sheer volume of datasets. The trained model might not be able to distinguish between distinct intrusion attacks if there is any inappropriate preprocessing. The primary focus of this effort was data pre-processing, which included data transformation and numerical standardisation. Before the proposed model can use them for training, all of the labels must be converted from their real text format into a numerical value. In this work, a categorization measure is used. Processing was used to categorise the attack data into binary and multi-class categorizations.

2.3. PSO

PSO is a metaheuristic approach instigated from the swarm behaviour of birds flocking and so on. PSO is concerned with shifting the particle's velocity throughout the search space to 'pbest' and 'lbest'. Individual particles in each generation will have their unique 'lbest' and 'gbest' values. Keeping track of the 'gbest' and 'pbest' values, every particle travels towards the best result in the search space. PSO communicates information for example, 'gbest', 'pbest', updated velocity, and location to each particle in the search space. The flowchart of PSO algorithm for watermarking is shown in Figure 2. Where, $u_1$ and $u_2$ are acceleration constants; $\mu$–weighted inertia parameter; $t$–iteration; $\sigma \approx 0.1 \sim 1$; $\varphi \approx 0.1 \sim 0.7$; $p_n$ and $g_n$ are the highest values for nth particle and each particle respectively.

\[
\vartheta_n(t + 1) = \mu \vartheta_n(t) + \sigma u_1(p_n - x_n) + \varphi u_2(g_n - x_n)
\]

\[
\rho_n(k + 1) = \rho_n(k) + \vartheta_n(k + 1)
\]

2.4. Proposed CNN architecture

The suggested CNN design, as displayed in Figure 3, contains five convolutional layers. The convolutional layer operates through calculating the input data in line with stride movement using the filters and kernel. 32 filters in the convolutional layer are shown to exist. Kernel size, also known as the convolution kernel’s window size, is found to be $2 \times 1$. Since max-pooling layers of 1 as well as 2 can successfully prevent overfitting, batch normalisation (BN) and dropout layers are added before output layers in CNN designs.

\[
PRelu(m) = \max(0,m) + \beta \min(0,m)
\]
Following convolution, recovered feature maps will concentrate on important information by removing insignificant noise. Each maximum pooling layer’s output dimension is raised. The output dimension of the convolutional layer was cut in half, and the number of parameters was decreased while still keeping important characteristics. To enhance classification performance in this CNN model, parametric rectified linear unit (PReLU) is utilised as an activation function. PReLU computation is specified below. Here $\beta$ represents the parameter of distribution with mean 0 and standard deviation is 1 that assures the negative axis slope. For categorization, sigmoid is applied for binary class and Softmax for multi-class categorization.

2.5. Optimization of CNN architecture using PSO

In order to optimise the parameters of CNN architectures, two optimisation methods are presented in this section. The PSO approach is used to decide the parameters which should have the greatest priority in order to achieve good CNN performance. The parameters to be improved were determined after evaluating a CNN’s performance in experimental study where the parameters were manually changed. Because, as was already indicated, different CNN parameter settings create a range of results for the same task and the objective is to determine the optimum architectures. The flowchart of the CNN optimization using PSO algorithm is exposed in Figure 4, where the “training and optimisation” block—where the CNN is initially set up to incorporate the parameter optimisation via using the PSO is the most crucial component of the entire process.
The PSO is initialised in this procedure in accordance with the execution parameter (the parameters are detailed below), and this produces the particles. Every solution indicates a finished CNN training since every particle is a potential solution and every position has a parameter which can be optimized. When all of the particles produced via the PSO are assessed for each generation, the training process, which is an iterative loop, comes to an end. The cost of computing is higher and is dependent on database and particles size and number, PSO iterations. In other words, 100 iterations of the CNN training process would be performed if the PSO were run with 10 particles and 10 iterations.

3. RESULTS AND DISCUSSION

In this research, the data set contains 80% training and 20% validation data correspondingly. Furthermore, training set and validation set are selected as ratio of 8:2 from training-validation dataset [30]. The introduced mechanism is executed on Tensorflow 2.2.0 structure with Windows personal computer configuration of 8 GB RAM, 512 GB solid-state drive. The whole work is executed by applying Python programming language as well as its suitable libraries. To shorten as well as calculate the recognition affect of malware, merged valuation metrics, for example, recall (RE), F1-measure, precision (PR) and accuracy (ACU), are useful to evaluate the function of proposed CNN for detecting the intrusion [31]. These four metrics calculations are specified below. Where $F_{\text{Negative}}$denotes the false negatives, $T_{\text{Negative}}$indicates the true negatives, $F_{\text{Positive}}$represents the false positives, and $T_{\text{Positive}}$depicts the true positives.

\[
ACC = \frac{T_{\text{Positive}} + T_{\text{Negative}}}{T_{\text{Positive}} + T_{\text{Negative}} + F_{\text{Positive}} + F_{\text{Negative}}} 
\]

\[
PR = \frac{T_{\text{Positive}}}{T_{\text{Positive}} + F_{\text{Positive}}} 
\]

\[
RE = \frac{T_{\text{Positive}}}{T_{\text{Positive}} + F_{\text{Negative}}} 
\]

\[
F1 - \text{measure} = \frac{2 \times PR \times RE}{PR + RE} 
\]

3.1. Evaluation

Table 1 provides the function of introduced CNN structure for detecting intrusion cultivated on CSE-CIC-IDS2018 dataset. This mechanism reached 91.12% ACU of benign and 91.37% and other attacks recognition. The model reached 92.54%, 93.11%, and 91.87% of ACU for denial of service (DoS), Bruteforce as well as botnet attacks severally. This experimentation illustrated that against DoS, botnet, and distributed (DoS), that are regularly applied by hackers today, the entire offered a better result. As infiltration and Webattacks have a smaller number of samples, the introduced model reached vaguely lower results when...
equated with other attacks. Figure 5 shows the training and legalization accuracy for binary and multi-class categorization using proposed CNN+PSO model.

It is obvious from the preceding findings that the presented method has a strong potential to enhance attack prediction in cloud environment as listed in Table 2. It shows the support vector machine (SVM) CNN+RNN, long short-term memory (LSTM), CNN+LSTM with proposed CNN+PSO model. The proposed lightweight CNN and PSO achieved 91.70% of accuracy which outperformed other ML and DL models.

Table 1. Execution of proposed CNN architecture for IDS

<table>
<thead>
<tr>
<th>Label</th>
<th>ACC</th>
<th>PR</th>
<th>RE</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Binary classification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td>91.12%</td>
<td>91.42%</td>
<td>92.86%</td>
<td>92.13%</td>
</tr>
<tr>
<td>All attacks</td>
<td>91.37%</td>
<td>92.61%</td>
<td>91.41%</td>
<td>92.01%</td>
</tr>
<tr>
<td><strong>Multi-class classification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benign</td>
<td>91.03%</td>
<td>90.12%</td>
<td>90.55%</td>
<td>90.33%</td>
</tr>
<tr>
<td>Bruteforce</td>
<td>93.11%</td>
<td>93.45%</td>
<td>92.82%</td>
<td>93.13%</td>
</tr>
<tr>
<td>DoS</td>
<td>93%</td>
<td>92.36%</td>
<td>92.90%</td>
<td>92.63%</td>
</tr>
<tr>
<td>Webattack</td>
<td>90.67%</td>
<td>90.37%</td>
<td>93.54%</td>
<td>91.93%</td>
</tr>
<tr>
<td>Infiltration</td>
<td>90.51%</td>
<td>92.18%</td>
<td>91.23%</td>
<td>91.70%</td>
</tr>
<tr>
<td>Botnet</td>
<td>91.87%</td>
<td>92.70%</td>
<td>91.20%</td>
<td>91.94%</td>
</tr>
<tr>
<td>DDoS attacks</td>
<td>91.15%</td>
<td>91.57%</td>
<td>91.69%</td>
<td>91.63%</td>
</tr>
</tbody>
</table>

Figure 5. Training and validation accuracy for binary classification using proposed method

Table 2. Comparison of accuracy for intrusion detection

<table>
<thead>
<tr>
<th>Model</th>
<th>Dataset</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN+RNN [32]</td>
<td>KDD Cup 99</td>
<td>85.23</td>
</tr>
<tr>
<td>SVM [33]</td>
<td>BIoT</td>
<td>79</td>
</tr>
<tr>
<td>LSTM [34]</td>
<td>CICIDS2017</td>
<td>85.64</td>
</tr>
<tr>
<td>CNN+LSTM [35]</td>
<td>CICIDS2017</td>
<td>80.91</td>
</tr>
<tr>
<td>Proposed CNN+PSO</td>
<td>CICIDS2018</td>
<td>91.70</td>
</tr>
</tbody>
</table>

4. CONCLUSION

With the addition of ML and DL algorithms, intrusion detection has profited greatly from developments in cyber security. This research introduced a novel CNN-PSO algorithm-based method for noticing intrusions in a cloud infrastructure. The outcomes from our method shows how effective it is at identifying intrusions, with an overall accuracy rate of 91.7%. Additionally, the PSO demands a lot of training time and hardware resources while not producing superior results. Therefore, before adopting an external optimisation technique, we advise creating a DNN with TensorFlow and Keras and assessing the function. The results can then be used to decide if an optimisation algorithm is necessary. Additionally, using feature selection approaches to be helpful in boosting the IDS’s entire effectiveness. Our model’s ability to utilize a small number of variables and still reach a high accuracy rate and shorten forecasting time is one of its significant strengths. By utilising well-chosen elements, our methodology raises operational efficiency and the accuracy rate. The method is time-consuming because of the model learning process, which is one of its drawbacks. In future, transfer learning strategies can be used to overcome these constraints.
REFERENCES


BIOGRAPHIES OF AUTHORS

Gnanam Jeba Rosline completed Master of Computer Applications from University of Madras and Pursuing Ph.D. as part time candidate in Mother Teresa Women’s University, Kodaikanal, Tamil Nadu, India under the guidance of Dr. Pushpa Rani. Currently working as a lecturer in University of Technology and Applied Sciences, MUSCAT. Area of research includes network security and artificial intelligence. She can be contacted at email: jebaroslineies@gmail.com.

Dr. Pushpa Rani completed Master of Computer Applications from Bharathiar University, India and Ph.D. from Madurai Kamaraj University, India. Currently working as a Director, Department of Computer Science, Mother Teresa Women’s University, Kodaikanal, Tamil Nadu, India. A research guide who has published 100 journals, 100 conference proceedings, 4 books, 5 projects and many funded projects. Expertise and research area includes biometrics, adaptive learning system, information retrieval, image processing, cloud computing, network security. She can be contacted at email: drpushpa.mtwu@gmail.com.