LISF: A Security Framework for Internet of Things (IoT) Integrated Distributed Applications

D. Shravani1*, Imtiyaz Khan2, Amogh Deshmukh3, Veeramalla Anitha4, Masrath Saba5, Syed Shabbeer Ahmad6

1Associate Professor, Department Of ADCE, SCETW, OU, Hyderabad, TS, India, drdasarishravani@stanley.edu.in
2Professor, Department Of CSE, Shandan College of Engineering and Technology JNTUH Hyderabad, TS India, imtiyaz.khan.7@gmail.com
3Assistant Professor CSE, School of Technology, Woxsen University Hyderabad TS India, amogh.deshmukh@woxsen.edu.in
4Assistant Professor, Department Of CSE, SCETW, OU, Hyderabad, TS, India, vanitha@stanley.edu.in
5Assistant Professor, Dept of CSE, KMIT, JNTUH Hyderabad, TS India, masrathsaba.be@gmail.com
6Professor, Department Of CSE, MJCET, OU, Hyderabad, TS, India, shabbeer.ahmad@mjccollege.ac.in

*Corresponding author’s E-mail: drdasarishravani@stanley.edu.in

Abstract

Distributed applications where Internet of Things (IoT) technology integrated are vulnerable to different kinds of attacks. Machine learning algorithms are widely used to detect intrusions in such applications. However, there is need for an effective unsupervised learning approach which can detect known and also unknown attacks. Towards this end, in this paper, we proposed a framework to protect security of IoT integrated architectures that are distributed in nature. Our framework is named Learning based IoT Security Framework (LISF). The framework is designed to have machine learning based security to IoT integrated use cases. Since IoT networks cause network traffic that is to be monitored and protected from external attacks, the proposed system uses deep learning technique for automatic detection of cyber-attacks. Particularly, the system exploits deep autoencoder which comprises of encoder and decoder for automatic detection of different kinds of intrusions. It is based on unsupervised learning which is crucial for distributed environments where network flows cannot have sophisticated training samples. We proposed an algorithm named Deep Autoencoder based Cyber Attack Detection (DAE-CAD). Experiments are made using IoT use case dataset known as UNSW-NB15. Our empirical results revealed that DAE-CAD outperforms existing methods with highest accuracy 91.36%.

Keywords: Internet of Things, Distributed Architectures, Machine Learning, Deep Learning, Security

1. Introduction

Security plays an important role in different kinds of networks and applications. With the emergence of Internet of Things (IoT) technology, many distributed use cases came into existence. For instance, smart city is one of the use cases that has distributed architecture and very complex network containing number of smaller networks. In such scenarios, the applications are more vulnerable to attacks. Therefore, it is important to have security mechanisms in place for secure end to end communications. There are many existing solutions that contributed towards security of such applications using machine learning techniques. In fact, many researchers contributed towards security of IoT integrated distributed applications.

directions. BIM and IoT integration enhance construction efficiency. Darabkh et al. [11] explored global IoT implementation, enabling technologies, challenges, and future directions. Ubiquitous sensing through wireless networks forms the backbone of IoT technologies. Kaur et al. [15] focused on IoT, a pinnacle in communication, transforms real-world objects into smarter devices, notably in precision agriculture. Our contributions in this paper are as follows.

1. We proposed a framework named Learning based IoT Security Framework (LISF) to have machine learning based security to IoT integrated use cases.
2. We proposed an algorithm named Deep Autoencoder based Cyber Attack Detection (DAE-CAD).
3. We built an application to evaluated LISF and our algorithm DAE-CAD using a benchmark dataset.

The remainder of the paper is structured as follows. Section 2 reviews literature on existing security models used for distributed architectures. Section 3 presents our methodology including system model and proposed framework. Section 4 presents experimental results while section 5 concludes our work.

2. Literature Review


Ahanger et al. [14] stated that IoT transforms global interactions, posing significant security challenges. Solutions must address privacy, trust, and security across all architectural levels. Kaur et al. [15] focused on IoT, a pinnacle in communication, transforms real-world objects into smarter devices, notably in precision agriculture. Comprehensive research reviews contributions and future directions. Elijah et al. [16] opined that the global population surge and resource challenges drive smart agriculture, employing IoT and data analytics for efficiency and productivity enhancement. Wu et al. [17] stated that IoT integrates with smart city, water, transportation, and manufacturing. Cloud-edge orchestration powered by AI optimizes data processing, but challenges persist. Xu et al. [18] found that IoT involves vast networks of physical devices; centralized security has limitations. Blockchain (BCT) offers security solutions, but challenges persist in integration and application scalability.

Liu et al. [19] tackled integrating field-level manufacturing data with cloud manufacturing, suggesting an IIoT gateway for efficient data management. The approach enhances decision-making and transforms traditional manufacturing into cloud systems. Adhikari et al. [20,23,26] explored fog computing for efficient real-time IoT applications, emphasizing reduced latency, energy consumption,
and challenges with potential solutions. From the literature it is observed that there is need for a security framework that could detect known and unknown cyber-attacks. Chander et al. [22] Detection of Anomalies and Leaf Disease Prediction in Cotton Plant Data IIoT environment. [24], [25] They investigated the concept of security using machine learning and deep learning methods for malware detection, as well as android malware detection with classification based on hybrid analysis and N-gram feature extraction. Chander et al. [27] data, identification and detection of outliers/anomalies is a challenging issue and raised as the primary importance of data analysis in IoT applications. Bilahari et al. [28] computing applications in cyber security, and analyzes the scenario of enhancing the cyber security potentials by suggests that of accelerating the intelligence of the security systems.

3. Materials And Methods

We proposed a novel methodology that is based on the system model presented in Figure 1. It is an IoT integrated distributed application scenario where the application is vulnerable to different kinds of attacks unless security is implemented. To overcome this issue, we proposed separate layer in the system model which is elaborated in Figure 4 to have a learning-based security framework. Ours is an AI based solution towards intrusion detection. Our deep learning model takes care of monitoring application for different kinds of attacks and ensures that the system is able to detect such attacks.
As presented in Figure 1, an overview of IoT integrated system model is provided reflecting a distributed application scenario. It has provision for security framework which takes care of learning-based protection to the system. It can detect cyber-attacks by employing machine learning techniques. It is an IoT integrated system model which is distributed in nature. It has interface layer where actual IoT or distributed applications run. The service layer provides required business logic and other related services. Network layer provides desired network infrastructure. Sensing layer has sensor network for realizing different smart activities.
As presented in Figure 2, it shows an encoding process which can protect applications from attacks that intend to steal data. It is based on many transformations to protect data from different attacks. It helps data to be protected when it is at rest and in transit. It focuses on stronger encryption model. The given file is encrypted using a modified AES algorithm. Then the resultant data is subjected slicing using IDA method. This makes the data more robust to ensure data integrity. Few slices can help in re-establishing the whole data. Finally, the data is subjected to hashing in order to achieve data integrity verification as and when needed.

**Figure 2:** Overview of the proposed encoding process
Figure 3: Overview of the proposed decoding process

As presented in Figure 3, there is a decoding process which is opposite to the encoding process. It considers encoded data as input along with secret key and converts data into slices. Afterwards, there is integrity verification with the help of hashing. The IDA converts the data into an encrypted file. Then the data is decrypted using the modified AES algorithm in order to obtain original content.
We proposed a framework known as Learning based IoT Security Framework (LISF) for protecting the system from cyber-attacks. It is based on deep autoencoder model which is based on unsupervised learning. Thus, the proposed framework can detect known and unknown attack scenarios. Encoder converts input data into some reduced representation while the decoder reconstructs it to detect different kinds of cyber-attacks. We proposed an algorithm named Deep Autoencoder based Cyber Attack Detection (DAE-CAD).

**Algorithm:** Deep Autoencoder based Cyber Attack Detection (DAE-CAD)

**Input:** UNSW-NB15 dataset D

**Output:** Attack detection results R

1. Begin
2. \((T_1, T_2) \leftarrow \text{DataSplit}(D)\)
3. \(F \leftarrow \text{ExtractFeatures}(T_1)\)
4. \(F' \leftarrow \text{OptimizeFeatures}(F)\)
5. Construct encoder
6. Construct decoder
7. Train encoder using \(F'\)
8. Train decoder using \(F'\)
9. \(R \leftarrow \text{AutoEncoder}(\text{encoder}, \text{decoder})\)
10. Return R

As presented in Algorithm 1, the given dataset is divided into training (T1) and test (T2) datasets. Features are extracted and optimized from T1. Then the optimized features are used to train encoder and decoder implicitly. The encoder and decoder perform miniature representation of data and reconstruction of data respectively. When it comes to the attack detection using test data, the autoencoder is employed to detect different kinds of attacks based on the encoding and decoding process.
3. Results and Discussion
We evaluated our framework with deep learning-based implementation to protect IoT use cases from cyber-attacks. Our algorithm named Deep Autoencoder based Cyber Attack Detection (DAE-CAD) is evaluated using an IoT use case dataset known as UNSW-NB15 [21]. This section presents results of experiments.

As presented in Figure 5, the proposed algorithm showed its performance reflected in the form of confusion matrix. Based on this different performance metrics shown in Table 1.

**Table 1:** Performance metrics used in this paper

<table>
<thead>
<tr>
<th>Metric</th>
<th>Formula</th>
<th>Value range</th>
<th>Best Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>( \frac{TP + TN}{TP + TN + FP + FN} )</td>
<td>[0; 1]</td>
<td>1</td>
</tr>
<tr>
<td>Precision (p)</td>
<td>( \frac{TP}{TP + FP} )</td>
<td>[0; 1]</td>
<td>1</td>
</tr>
<tr>
<td>Recall (r)</td>
<td>( \frac{TP}{TP + FN} )</td>
<td>[0; 1]</td>
<td>1</td>
</tr>
<tr>
<td>F1-Score</td>
<td>( \frac{2 \times (p \times r)}{(p + r)} )</td>
<td>[0; 1]</td>
<td>1</td>
</tr>
</tbody>
</table>

The proposed method is evaluated using these metrics. The experiments are made using UNSW-NB15 which has 7 kinds of attack instances. This dataset is used for training the proposed model and help in detection of attacks.

**Table 2:** Experimental results showing performance of different models

<table>
<thead>
<tr>
<th>Attack Detection Model</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logistic Regression</td>
<td>0.78829</td>
<td>0.775965</td>
<td>0.83929</td>
<td>0.806395</td>
</tr>
<tr>
<td>Multilayer Perception</td>
<td>0.80512</td>
<td>0.812175</td>
<td>0.82246</td>
<td>0.81719</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>0.806225</td>
<td>0.818465</td>
<td>0.81719</td>
<td>0.817785</td>
</tr>
<tr>
<td>Random Forest</td>
<td>0.815745</td>
<td>0.81855</td>
<td>0.831555</td>
<td>0.82501</td>
</tr>
<tr>
<td>Proposed Deep Autoencoder</td>
<td>0.913634</td>
<td>0.916776</td>
<td>0.931342</td>
<td>0.924011</td>
</tr>
</tbody>
</table>

As presented in Table 2, the proposed deep learning model is compared against different existing machine learning models.
As presented in Figure 6, performance of different models in intrusion detection for given IoT use case is evaluated. Higher value for any metric used in the evaluation indicates better performance. Logistic Regression (LR) achieved 78.82% accuracy, Multilayer Perceptron (MLP) showed 80.51%, Decision Tree (DT) 80.62%, Random Forest (RF) exhibited 81.57%, while the proposed deep autoencoder based model showed highest accuracy 91.36%. From the experimental results, it is found that the proposed model is capable of improving attack detection accuracy due to its modus operandi and ability to discriminate legitimate and attack traffics.

4. Conclusion
In this paper, we proposed a framework to protect security of IoT integrated architectures that are distributed in nature. Our framework is named Learning based IoT Security Framework (LISF). The framework is designed to have machine learning based security to IoT integrated use cases. Since IoT networks cause network traffic that is to be monitored and protected from external attacks, the proposed system uses deep learning technique for automatic detection of cyber-attacks. Particularly, the system exploits deep autoencoder which comprises of encoder and decoder for automatic detection of different kinds of intrusions. It is based on unsupervised learning which is crucial for distributed environments where network flows cannot have sophisticated training samples. We proposed an algorithm named Deep Autoencoder based Cyber Attack Detection (DAE-CAD). Experiments are made using IoT use case dataset known as UNSW-NB15. Our empirical results revealed that DAE-CAD outperforms existing methods with highest accuracy 91.36%. In future, we intend to improve our framework by using hybrid deep learning model for intrusion detection more efficiently.

References:


