HOMOMORPHIC ENCRYPTION-ENABLED DEEP LEARNING MODEL FOR INTELLIGENT CYBER THREAT DETECTION

Abstract:

Cybersecurity is one of the emerging fields which address information security, protecting computer systems and networks from intrusions.Cyber-attack detection is crucial for safeguarding computer networks against unauthorized access and malicious activities. Traditional machine learning approaches often suffer from overfitting, leading to reduced accuracy. To overcome this, an optimized deep learning approach with a homomorphic encryption-based authentication protocol is proposed. Cyber threat data is collected from open sources and pre-processed using data cleaning and binning method. Feature extraction is carried out using TF-IDF, followed by dimensionality reduction through Principal Component Analysis. The processed data is then classified using an Extended Physics-Informed Neural Network (EPINN) for cyber-attack detection. A Signature-based Authentication Protocol ensures secure user authentication, while BFV encryption secures data storage in the cloud. Experimental results show high efficiency with 97.36% accuracy, 89.78% precision, and a low false positive rate of 2.17%. This approach enables automatic and robust cyber threat detection, improving proactive defense mechanisms for organizations.

Keywords:*Cyber security; Feature extraction; Principal component analysis; Extended Physics-Informed Neural Network; Classification*.

1. Introduction

The term "cyber security" describes the procedures and policies intended to shield computers, networks, software, and data from harm, illegal access, and cyberattacks. In order to protect the confidentiality, integrity, and availability of information, it entails utilizing technologies, procedures, and policies to stop, identify, and address threats [1].Cybersecurity is “the art of protecting networks, devices, and data from unauthorized access or criminal use and the practice of ensuring confidentiality, integrity, and availability of information [2]. The proliferation of online and mobile banking has led to a steady increase in cybercrime, which includes a variety of frauds like identity theft, credit card scams, spam, and ATM theft. The financial industry's data is particularly sensitive, despite its substantial monetary value [3]. Hackers can use stolen financial information and banking credentials to generate income in a variety of ways.Due to the significant value of stored data, which gives hackers several opportunities to make money, the banking industry is especially susceptible to these kinds of attacks [4].

Cyber-attack detection in banking platforms can be expensive and hard to operate, particularly for smaller businesses with limited funds and infrastructure, due to the possibility of false positives, the inability to respond to quickly emerging threats, and the need for substantial computational resources [5]. To lessen these issues, automated cyberattack detection based on artificial intelligence (AI) was developed. Large datasets are analyzed by AI in automated cyberattack detection to find suspicious trends, allowing for quicker and more precise detection. In a subfield of artificial intelligence, machine learning algorithms are employed to help robots learn from experiments more effectively [6].

2. Related works

A Deep Belief Network (DBN) was developed for cyber-attack detection in industrial automation using temporal patterns in network data to identify anomalies over time [7]. The method achieved 93.2% accuracy with reduced false positives and an error rate of 6.8%. However, its reliance on labeled data limits effectiveness against zero-day and unseen attacks.Deep Neural Network (DNN) was proposed for real-time cyber-attack detection by analyzing large-scale network traffic patterns [8]. It achieved 94.5% accuracy, 92.7% precision, and 91.2% specificity, supported by image enhancement techniques. Despite strong performance, high computational complexity makes deployment difficult on resource-constrained devices.

A Long Short-Term Memory (LSTM) framework was introduced for cyber-attack detection using feature selection to reduce data dimensionality and improve accuracy [9]. The model achieved 94.9% detection accuracy and 93.4% precision, effectively identifying sophisticated attack patterns. However, the feature selection process may exclude critical features, causing occasional misclassification.Recurrent Neural Network (RNN) was developed for cyber-attack detection using temporal data [10]. It monitors variations in network behavior over time, effectively capturing subtle and continuous attack patterns. By employing a patch-based approach, the system enhanced detection performance, achieving an accuracy of 94.8% and a specificity of 91.7%. However, it relies heavily on large volumes of sequential data, making it less effective in cases where data is sparse or lacks continuity.

A Convolutional Neural Network (CNN) was designed to detect cyber-attacks in IoT networks using a graph-based method, where devices and their communications are represented as nodes and edges [11]. The model achieved an accuracy of 92.4% and a precision of 90.7%, with computational efficiency that supports real-time applications. Nonetheless, as IoT networks scale, the increased data complexity can reduce performance, making it more challenging to maintain detection accuracy in large environments.Graph Neural Networks (GNN) was proposed to identify cyberattacks on dispersed networks [12]. The distributed model protects data privacy while assisting in the detection of cyberattacks in a variety of contexts. The complexity of the approach is increased by communication overhead between federated systems, which make scaling for real-time detection across several nodes challenging.

A verifiable dynamic access control mechanism with randomized ensemble SVM was developed to address user revocation in IoT architectures [13]. The approach integrates VOMAACS (Verifiable Outsourced Multi-Authority Access Control Scheme) with CPASBE (Ciphertext Attribute Set-Based Encryption) to enhance data security. In CPASBE, sensitive information remains with the data owner, ensuring confidentiality, preventing duplication, and protecting against agreement breaches regardless of provider reliability. When the system receives new or unknown attacks that match pre-existing signatures, its performance suffers due to its dependence on signature-based detection [14]. When presented with noisy or insufficient data, CNN's performance can decrease, decreasing the accuracy of detection in practical settings. Occasionally, sophisticated attacks may be misclassified as a result of the process's unintentional exclusion of crucial features. Scaling for real-time detection over numerous nodes is challenging due to the approach's complexity [15]. To improve the security of Wireless Sensor Networks (WSNs) by detecting and averting cyberattacks, an intelligent hybrid model that combines artificial intelligence and machine learning was proposed. Singular Value Decomposition (SVD) and Principal Component Analysis (PCA) are two feature reduction approaches used in the study. The K-means clustering model enhanced information gain (KMC-IG) is also used for feature extraction. Intrusion detection systems and network traffic classification are presented with the introduction of the Synthetic Minority Excessively Technique for data balancing [16].

3. Proposed Methodology

Cyber security is the practice of defending computers, servers, mobile devices, electronic systems, networks, and data from malicious attacks. Because it guards against theft and destruction to all types of data, cyber security is crucial. False positives from cyberattack detection systems might overload security professionals with warnings that aren't harmful. They might also find it difficult to identify complex, dynamic dangers in real time. In this case, a deep learning-based solution that can identify and safeguard the system in an aberrant state is suggested. Figure 1 depicts the suggested architecture for deep learning-based cyber security threat identification.

To increase the quality of the data, the obtained data are first pre-processed. Feature extraction is carried out using TF-IDF, followed by dimensionality reduction through Multi-Kernel Principal Component Analysis. The processed data is then classified using an Extended Physics-Informed Neural Network (EPINN) for cyber-attack detection. A Verifiable Ring Signature-based Authentication Protocol ensures secure user authentication, while BFV encryption secures data storage in the cloud.

(i) Data Pre-Processing:

In cybersecurity-oriented IoT environments, preprocessing plays a crucial role in preparing raw network traffic, sensor logs, and user access records for reliable intrusion detection and anomaly analysis. Since IoT data is often heterogeneous, large-scale, and prone to noise or missing values, robust preprocessing ensures the integrity and accuracy of subsequent detection models.

Data Cleaning is an essential step which focuses on addressing invalid, inconsistent, or incomplete information. IoT traffic logs may contain missing values due to packet loss, device malfunctions, or communication delays. These can be handled in two ways: (1) ignoring the tuples with missing data, a viable option only in very large datasets, or (2) filling the missing values by imputation techniques such as mean replacement, predictive estimation, or using default values. This preserves data continuity, which is crucial for real-time threat analysis.

Another challenge is noisy data, which may arise from corrupted sensor readings, unauthorized access attempts, or data entry errors in IoT networks. To address this, methods such as the Binning Method are applied, where continuous data (e.g., traffic rates, packet counts) is sorted and divided into equal-sized intervals to smooth out inconsistencies. This reduces false alarms in anomaly detection.

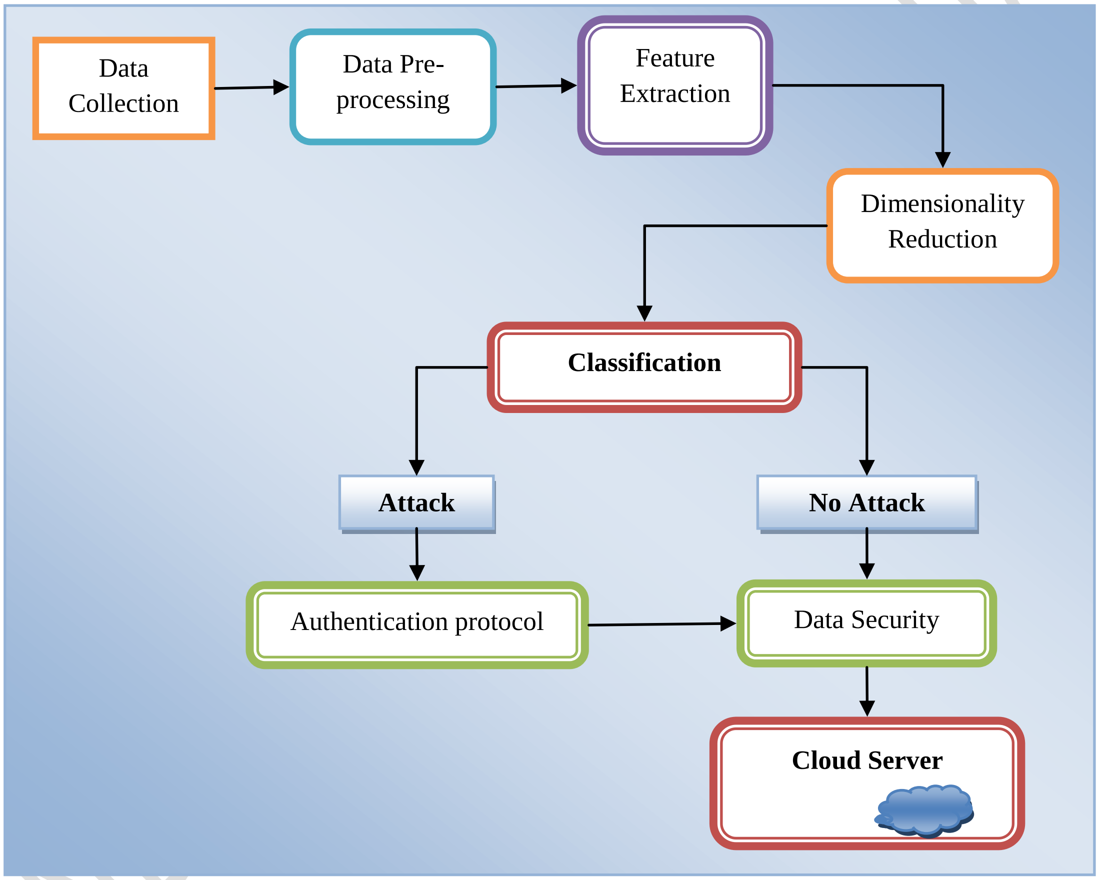

Figure 1: Block diagram for Proposed Model

(ii) Feature Extraction:

The process of turning unstructured IoT or cybersecurity data into valuable features that serve as crucial information for analysis or modeling is known as feature extraction. By choosing, transforming or producing variables that emphasize the most pertinent patterns, this stage is essential to improving the performance of machine learning algorithms. Effective feature extraction aids in the detection of harmful traffic patterns, illegal access attempts, and unusual device behavior when communicating in cybersecurity-based IoT systems.

Term Frequency–Inverse Document Frequency (TF-IDF) is a popular method for feature extraction in cybersecurity data that is based on text and logs.TF-IDF is a statistical measure that evaluates the importance of a word (or token) in a dataset relative to its frequency across multiple documents. TF-IDF can be used to extract discriminative features for anomaly or attack detection in network traffic logs, system call traces, or access records in the context of IoT cybersecurity. The term frequency measures how frequently a term occurs in a document, which is given in equation 1,

$$TF(t,d) = \frac{f_{t,d}}{\Sigma f_{t',d}} \qquad (1)$$

Here, $f_{t,d}$ is the number of times term 't' appears in document 'd'. Inverse Document Frequency (IDF) reduces the weight of common terms and increases the weight of rare terms across the dataset and it is given in equation 2,

$$IDF(t) = \log \frac{N}{1+n_t} \qquad (2)$$

where 'N' is the total number of documents, and n_t is the number of documents containing term t.

The terms that appear frequently in a particular document (or log entry) but infrequently in other documents are thus emphasized by TF-IDF, which makes them effective markers of anomalous activity or possible cyber threats.

Features like abnormal command sequences, infrequent access attempts, or abnormal protocol usages that can indicate cyberattacks can be extracted using TF-IDF from cybersecurity-based IoT. In order to increase detection accuracy and reduce false positives, these collected features are subsequently input into classifiers (such as deep learning or ensemble models).

High-dimensional data is common in cybersecurity-based IoT systems due to the massive volume of logs, network packets, and sensor readings. Therefore Principal Component Analysis (PCA) is widely used as a dimensionality reduction technique.PCA transforms the original correlated features into a smaller set of uncorrelated variables known as principal components (PCs), while retaining most of the variance (information) in the data. This helps in reducing noise, improving model efficiency, and focusing on the most discriminative patterns in IoT cybersecurity data.

(iii) Classification:

Classification is the process of assigning data points to predefined categories based on patterns learned from labelled data. Extended Physics-Informed Neural Networks (EPINN) integrates physical laws into neural networks for solving complex scientific problems, improving accuracy in simulations.

Extended Physics-Informed Neural Networks (EPINN) incorporates physical laws and domain knowledge into neural network architectures, improving accuracy and generalization in solving complex, physics-based problems. The Structure Diagram of the Physics Informed Neural Network is shown in Figure 2 is shown below.

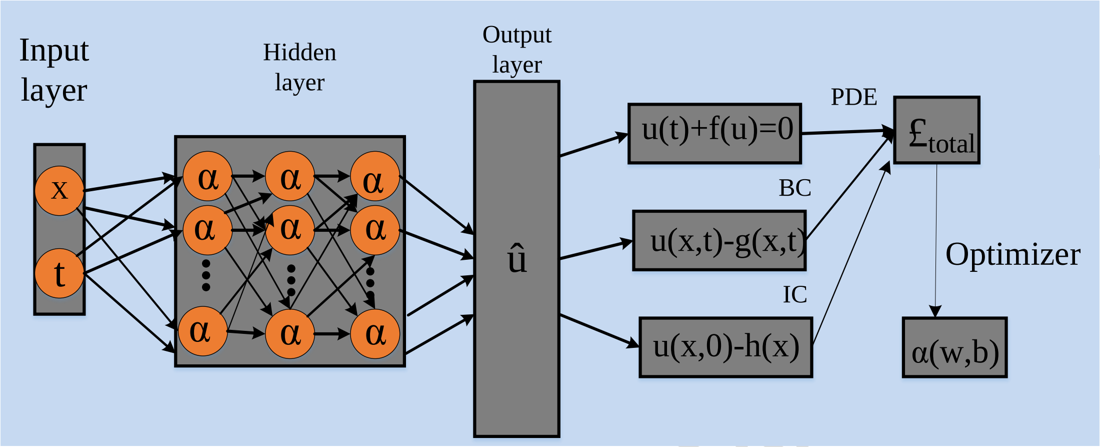

Figure 2: Structure Diagram of the EPINN

A Structure Diagram of a EPINN visually represents how neural networks integrate physical laws into their architecture. It typically includes components such as input layers for initial conditions, hidden layers where the physics-based constraints are enforced, and output layers that produce predictions while satisfying governing equations. This framework enables the model to learn both data-driven patterns and the underlying physics, enhancing accuracy and generalization in scientific applications.

To approximate the solution of the partial differential equation using spatial-temporal coordinates $u(x,t)$ as the input, Fully linked feed forward neural networks with multiple hidden layers are used by EPINN are represented as shown below in equation 3,

$$(na_d^i(L_d(X^{d-1}))_i), d = 1,2 \dots., D-1, i = 1,2 \dots \dots . N_d \quad (3)$$

Once classification process is done then the attacked packets are re-registered with the cloud server using Signature based Authentication Protocol (SAP) and the normal packets are directly send for the encryption processfor secure storage. VRSAP is usedto verify a user's identity before allowing them access to systems or networks and it is important for protecting systems, data, and applications from attacks.

Following user verification, the data is safely stored on a cloud server using Brakerski/Fan-Vercauteren (BFV) cryptography. The BFV scheme, a lattice-based holomorphic encryption technique, is perfect for data security and privacy-preserving computations since it allows safe addition and multiplication operations on encrypted data.

4. Results and Discussion

To improve network security, the proposed homomorphic encryption based authentication protocol based on BFV and EPINN have been created for the detection of cybersecurity attacks.With an Intel Core i5 CPU, an Nvidia GeForce GTX 1650 GPU, and 16GB of RAM, it was designed using Python 3.8.

The Kaggle website is the original source of the data. This dataset has been used extensively to assess anomaly detection and is based on data collected during the DARPA'98 IDS evaluation program.

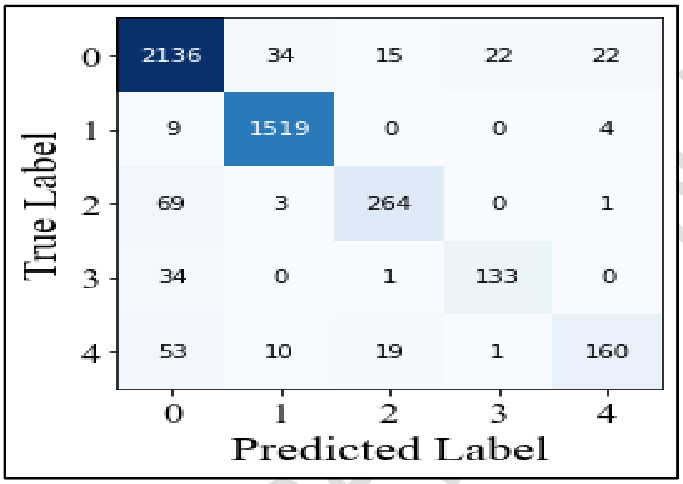

Figure 3: Confusion Matrix for Proposed Model

The confusion matrix in Figure 6 provides a comprehensive evaluation of the proposed classification method for cyber-attack detection. It is used to validate the accuracy of the model by comparing predicted outputs with the actual class labels. These results demonstrate the effectiveness of the proposed method in distinguishing between normal and malicious traffic. By accurately classifying different attack categories, the model proves its robustness and applicability in real-world cybersecurity-based IoT environments where timely and precise detection of various attack types is critical for system security.

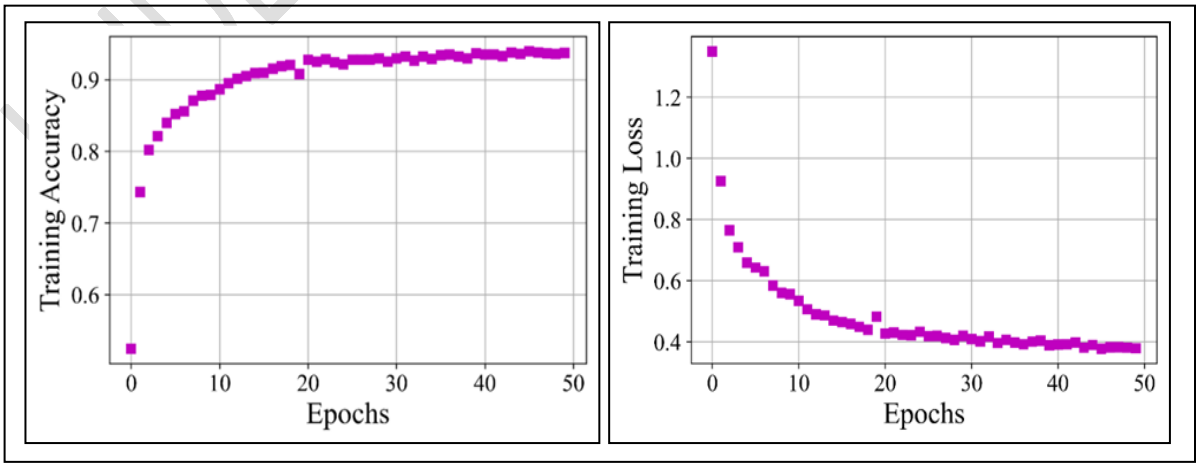

Figure 4: Evaluation of Training Accuracy and Loss with varying Epochs

The figure 4 shows the evaluation of training accuracy and training loss by different numbers of epochs. The model gains more knowledge from the training data as the number of epochs rises and both the training and validation datasets show an improvement in prediction accuracy. This is because the model has more chances to minimize the loss function by modifying its weights and biases.

5. Conclusion

The proposed optimized deep learning system with authentication based on homomorphic encryption demonstrates significant potential for improving cyberattack detection in networks enabled by the Internet of Things. The model achieves excellent accuracy with low false positive rates by combining efficient preprocessing, TF-IDF-based feature extraction, PCA-driven dimensionality reduction, and EPINN classification. Strong data security and user privacy are also guaranteed in cloud environments via BFV encryption and a signature-based authentication scheme. Therefore, this strategy provides a proactive and dependable protection mechanism, enhancing cybersecurity resistance to changing threats.

References:

1. Elsisi, M., & Tran, M. Q. (2021). Development of an IoT architecture based on a deep neural network against cyber attacks for automated guided vehicles. *Sensors*, *21*(24), 8467.
2. Chang, K., & Huang, H. (2023). Exploring the management of multi-sectoral cybersecurity information-sharing networks. *Government Information Quarterly*, *40*(4), 101870.
3. Kumar, V., & Sinha, D. (2021). A robust intelligent zero-day cyber-attack detection technique, Complex Intell. In *Syst* (Vol. 7, No. 5, pp. 2211-2234).
4. Sridhar, S., &Govindarasu, M. (2017). Model-based attack detection and mitigation for automatic generation control. IEEE Transactions on Smart Grid, 5(2), 580-591.
5. Balta, E. C., Pease, M., Moyne, J., Barton, K., & Tilbury, D. M. (2023). Digital twin-based cyber-attack detection framework for cyber-physical manufacturing systems. IEEE Transactions on Automation Science and Engineering, 21(2), 1695-1712.
6. Zhang, F., Kodituwakku, H. A. D. E., Hines, J. W., & Coble, J. (2019). Multilayer data-driven cyber-attack detection system for industrial control systems based on network, system, and process data. *IEEE Transactions on Industrial Informatics*, *15*(7), 4362-4369.
7. Jatothu, M. S., & Devi, D. S. L. (2024). Evolving Deep Belief Network for Cyber-Attack Detection in Industrial Automation and Control Systems. Journal of Cybersecurity and Privacy, 5(1), 1-15.

8. Vijayakumar, K. P., Pradeep, K., Balasundaram, A., &Prusty, M. R. (2023). Enhanced cyber attack detection process for internet of health things (IoHT) devices using deep neural network. *Processes*, *11*(4), 1072.
9. Zarzycki, K., Chaber, P., Cabaj, K., Ławryńczuk, M., Marusak, P., Nebeluk, R., ... & Wojtulewicz, A. (2023). Forgery cyber-attack supported by LSTM neural network: an experimental case study. Sensors, 23(15), 6778.
10. Udas, P. B., Roy, K. S., Karim, M. E., & Ullah, S. M. A. (2023, February). Attention-based RNN architecture for detecting multi-step cyber-attack using PSO metaheuristic. In 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE) (pp. 1-6). IEEE.
11. Nedeljkovic, D., & Jakovljevic, Z. (2022). CNN based method for the development of cyber-attacks detection algorithms in industrial control systems. Computers & Security, 114, 102585.
12. Friji, H., Olivereau, A., &Sarkiss, M. (2023, May). Efficient network representation for GNN-based intrusion detection. In International conference on applied cryptography and network security (pp. 532-554). Cham: Springer Nature Switzerland.
13. Kumar, R. A., & Vinuthna, K. (2021). Randomized Ensemble SVM based Deep learning with Verifiable dynamic access control using user revocation in IoT architecture. Sādhanā, 46(4), 229.
14. Stojanović, B., Hofer-Schmitz, K., & Kleb, U. (2020). APT datasets and attack modeling for automated detection methods: A review. Computers & Security, 92, 101734.
15. Burnap, P., Anthi, E., Reineckea, P., Williams, L., Cao, F., Aldmoura, R., & Jones, K. (2024). Mapping automated cyber attack intelligence to context-based impact on system-level goals. Journal of Cybersecurity and Privacy, 4(2), 340-356.
16. Behiry, M. H., & Aly, M. (2024). Cyberattack detection in wireless sensor networks using a hybrid feature reduction technique with AI and machine learning methods. Journal of Big Data, 11(1), 16.