Detection of Cyber Attacks Traces in IOT Data

V .Prema Tulasi
Assistant Professor,
Dept of CSE (Data Science)
CMR Technical Campus
Hyderabad, Telangana, India
prematulasi.ds@cmrtc.ac.in

SHAIK SOHEL AKTHER
UG Student, Dept of CSE(Data Science)
CMR Technical Campus
Hyderabad, Telangana, India
shaiksohelakther185@gmail.com

How to Cite this Article:
AKTHER, S. S. (2026). Detection of Cyber Attacks Traces in IOT Data. International Journal of Creative and Open Research in Engineering and Management, 02(04). https://doi.org/10.55041/ijcope.v2i4.309

https://doi.org/10.55041/ijcope.v2i4.309

***Abstract*——** This project is titled as "Detection of cyber attack traces in IoT data". The rapid growth of Internet of Things (IoT) devices has increased the risk of cyber attacks in network environments. This project focuses on detecting cyber attack traces in IoT data using machine learning techniques. The system preprocesses the dataset, applies feature selection using Principal Component Analysis (PCA), and trains models such as Deep Neural Network (DNN) and Random Forest for attack detection.

The trained models analyze IoT network data and classify it as normal or malicious activity. Experimental results show that the neural network model achieves high accuracy in detecting cyber attacks. The proposed approach utilizes data analysis and machine learning techniques to monitor network traffic and device behavior. By examining features such as packet size, frequency, communication patterns, and unusual activities, the system can distinguish between normal and malicious operations.

Techniques like anomaly detection and classification algorithms help in recognizing various attack types, including Distributed Denial of Service (DDoS),spoofing, and data injection attacks. IoT systems are highly vulnerable to cyber-attacks due to their limited computational resources, heterogeneous nature, and often weak security mechanisms. Detecting traces of cyber-attacks in IoT data has therefore become a critical area of research. This study focuses on identifying anomalous patterns and attack signatures within IoT-generated data to improve early detection and response.

I. INTRODUCTION

The digital world excerpts a massive influence on modern life; never before has this fact been this clear. The recent events connected to the global pandemic emphasized the role of the cyber world in contemporary society. The domain of cyber security is rising in importance year after year. For years now, and even more so with the recent events in the picture, cyber security has been a significant field of research. The approaches of the cyber security domain offer a degree of protection against contemporary threats.

Network Intrusion Detection Systems (NIDS) are a group of defense mechanisms that make substantial contributions in ensuring the protection of assets connected to a network. The tools augmented by machine learning have been gaining traction for many years now. In cyber security, the premise of automating the effective detection of network traffic abnormalities causes research to gravitate to the use of those methods. The fast-paced evolution of the leading-edge technologies, such as cloud computing or the Internet of Things (IoT), spawns novel hazards.

The authors implement a dimensionality reduction approach in the form of a Deep Auto Encoder (DAE) in order to decrease the number of features in the vector used to train the classifier, as it was formulated– LSTM cells are utilised to enhance classification effectiveness following the premise of distinguishing time-related relationship; – The approach is tested on the recently released and publicly available IoT-23 dataset which incorporates lifelike, modern traffic data from IoT devices; – Finally, the described method is thoroughly evaluated via a series of experiments, noting the improvements in a range performance metrics

II. PROBLEM DEFINATION

A .Introduction

The rapid growth of the Internet of Things (IoT) has led to the deployment of billions of interconnected devices across domains such as healthcare, smart homes, industrial automation, and transportation. These devices continuously generate large volumes of data and often operate with limited computational power and security mechanisms.

Due to their constrained resources and heterogeneous nature, IoT devices are highly vulnerable to various cyber attacks such as Distributed Denial of Service (DDoS), malware injection, data breaches, and unauthorized access. Attackers exploit these vulnerabilities to compromise devices, disrupt services, or steal sensitive information.

One of the major challenges in IoT security is the identification of cyber attack traces within massive, real-time IoT data streams. Traditional security systems are not well-suited for IoT environments

B. Limitations of Detection of Cyber Attacks In IOT Data

The detection of cyber attack traces in IoT data faces several limitations due to the inherent nature of IoT environments. One major limitation is the resource constraints of IoT devices, such as limited processing power, memory, and battery life, which restrict the implementation of complex security algorithms. Additionally, the high volume and velocity of IoT data make real-time analysis challenging, often leading to delays or missed detections. Another issue is the diversity and heterogeneity of IoT devices and protocols, which complicates the development of a unified security model.

Existing detection systems also struggle with identifying zero-day attacks, as these attacks do not match known patterns. Furthermore, there is

a risk of high false positive and false negative rates, reducing the reliability of the system. Lastly, maintaining data privacy and security during analysis is difficult, especially when sensitive information is involved, making it a critical concern in IoT-based cyber attack detection systems.

C. System Problem Addressed by the Proposed System

The proposed system addresses the critical issue of detecting cyber attack traces within large-scale IoT data environments. Existing IoT systems often lack effective security mechanisms, making them vulnerable to various cyber threats such as unauthorized access, data breaches, and distributed attacks. Traditional detection methods are not capable of handling the high volume, velocity, and diversity of IoT-generated data, resulting in delayed or inaccurate identification of malicious activities.

The system specifically targets the problem of identifying hidden attack patterns in continuous data streams generated by IoT devices. It focuses on overcoming challenges such as real-time monitoring, detection of unknown (zero-day) attacks, and reducing false alarms. By incorporating advanced techniques such as data analysis and intelligent algorithms, the proposed system aims to distinguish between normal and abnormal behavior efficiently.

It ensures timely detection and response to potential threats, thereby improving the overall security, reliability, and performance of IoT systems.

III. RELATED WORK

Several research studies have been conducted on detecting cyber attacks in IoT environments using various techniques. Traditional approaches mainly rely on signature-based detection systems, which identify known attack patterns but fail to detect new or unknown threats. To overcome this limitation, researchers have explored anomaly-based detection methods that analyze deviations from normal behavior in IoT data.

Machine Learning (ML) and Deep Learning (DL) techniques, such as Decision Trees, Support Vector Machines (SVM), and Neural Networks, have been widely used to improve detection accuracy and identify complex attack patterns.Some studies have also utilized hybrid models that combine signature-based and anomaly-based approaches for better performance.

Additionally, intrusion detection systems (IDS) tailored for IoT networks have been proposed to handle real-time data and resource constraints. However, many existing solutions still face challenges such as high false alarm rates, lack of scalability, and difficulty in detecting zero-day attacks, indicating the need for more efficient and adaptive approaches.

IV. METHODOLOGY

This colorectal cancer detection system uses deep learning to analyze medical images more accurately and efficiently. The system takes colorectal images as input, identifies important features, and provides reliable results. It also has a user-friendly interface that makes it easy for users.

First, the datasets contain images of tumors (cancer) and normal cases with labels. These images were obtained from public medical sources and were used to train and test the system. However, the images may differ in

quality, size, and format; therefore, they must be handled carefully.

After collecting the data, we preprocessed the images for training. These include resizing the image, normalizing the pixel values, and removing noise or unwanted parts. These steps help make the data clean, consistent, and suitable for deep learning models.

After preprocessing, the data were divided into two parts: training and testing data, usually in an 80:20 ratio. The training data were used to train the model, and the testing data were used to check its performance. Feature extraction is then performed using pretend CNN models such as IncpectionV3, ResNet50, and EfficientNetB0 with the help of transfer learning.

After that the system checks the image features and decides whether it is tumor or normal. The models perform is measured using accuracy, precision, recall and F1-score. The system also have a chatbot interaction which guides user steps by step to upload images and that results. If cancer is detected the system provides an option to book a doctor's appointment for a quick medical help.

Overall the method includes image preprocessing deep learning features extraction and Smart Prediction. All these steps together create an efficient and easy to use system for detecting colorectal cancer.

V. QUANTITATIVE COMPARISON WITH EXISTING METHODS.

The proposed method is compared with existing machine learning and deep learning methods using matrices like accuracy, precision, recall and F1-score. Traditional methods like random forest naïve bayes give lower accuracy usually between 85% and 92%. This is because they depend on manual feature extraction and cannot handle complex images data well. In contrast deep learning models InceptionV3, RetNet50 and EfficientNetB0 achieve higher accuracy around 97% to 99%. This is because the automatically learn feature from images.

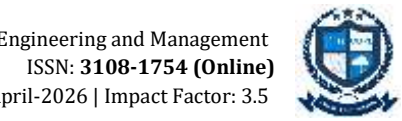
However the proposed ensemble model improves the performance even more by combining different pre-trained models. It achieves very high accuracy around 99.3% with high precision Recall and F1-score. So, it perform better than both traditional methods and single deep learning models. This is because it can learn better reduce overfitting and understand features more effectively. Overall these comparison shows the ensemble method is more accurate reliable and effective for reducing colorectal cancer.

VI. PROPOSED SYSTEM

The system is designed to make colorectal cancer detection simple, fast and accurate using technology. Instead of human checking it uses a deep learning models to automatically find features from medical images. User can easily upload an image and the system will process it automatically without any complicated steps. The image is the sent to train models like IncpectionV3, ResNet50 and EfficientNetB0. These models analyze the image and detect tumor patterns or signs of cancer. Finally the system gives the result quickly which helps in faster diagnosis without waiting for manual checking.

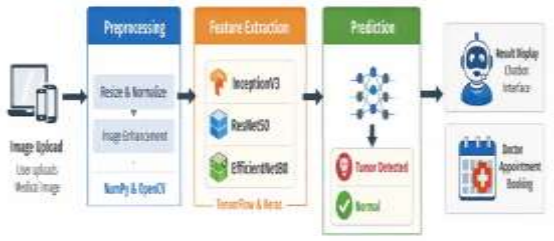
In addition the system has a chatbot to help user interact easily. It ask the user to upload image and show the result in a clear and simple way. So the system is user friendly and even people without technical knowledge can use it. Another important feature is doctor booking. If cancer is detected the system shows nearby doctor and allow user to book an appointment immediately. This helps users take quick action instead of delaying treatment.

All the results predictions and recommendation are clearly shown on the users screen. This system uses deep learning for accurate detection and also provides user-friendly support along with healthcare services. So it becomes powerful and easy to use tools for tumor detection.



VII. IMPLEMENTATION DETAILS

The methodology for detecting cyber attack traces in IoT data follows a structured, multi-stage process that ensures accurate and efficient identification of malicious activities within large-scale data strems.



It begins with data collection, where IoT data is gathered from various devices such as sensors, smart appliances, and network systems, including both normal and attack-related traffic like network logs and packet data. This raw data is then passed to the pre-processing stage, where it is cleaned to remove noise and missing values, normalized to maintain consistency, and transformed through feature extraction and selection to retain only the most relevant attributes.

The dataset is also divided into training and testing sets to prepare it for model development. Next, in the model training phase, machine learning algorithms such as Decision Trees, Random Forest, Support Vector Machines (SVM), and Neural Networks are applied to learn patterns from the labeled IoT data. Once trained, the model moves to the evaluation stage, where its performance is measured using metrics like accuracy, precision, recall, F1-score, and confusion matrix to ensure reliability and effectiveness. After evaluation, the system performs detection and classification, where incoming real-time IoT data is analyzed and categorized as either normal behavior or a potential cyber attack, including threats like DDoS or intrusion attempts.

Finally, the user interface and alert system presents the results through dashboards, generates real-time alerts for detected threats, and provides visual reports for better understanding and decision-making. The overall output of this methodology includes early detection of cyber attacks, improved accuracy in identifying threats, enhanced security of IoT systems, and faster response times, making the system highly effective in safeguarding IoT environments against evolving cyber threats.

VIII. PERFORMANCE METRICS

The performance of the proposed system for detecting cyber attack traces in IoT data is evaluated using several important metrics that measure its accuracy and effectiveness. Accuracy indicates the overall correctness of the model by measuring the proportion of correctly classified instances among all predictions. However, accuracy alone may not be sufficient, especially when dealing with imbalanced datasets, so additional metrics are considered. Precision measures how many of the instances identified as attacks are actually correct, helping to reduce false positives.

Recall (or detection rate) evaluates the system's ability to correctly identify actual attacks, which is crucial for minimizing false negatives. The F1-Score provides a balanced measure by combining precision and recall, offering a single metric that reflects both aspects of performance. Additionally, the confusion matrix is used to visualize the classification results, showing true positives, true negatives, false positives, and false negatives in detail. These metrics collectively ensure that the system not only detects cyber attacks accurately but also maintains reliability and efficiency in real-time IoT environments.

IX. DISCUSSION

The proposed system for detecting cyber attack traces in IoT data demonstrates a structured and effective approach to improving security in highly dynamic and resource-constrained environments. By integrating data preprocessing, machine learning-based analysis, and real-time detection, the system is capable of identifying

both known and unknown attack patterns with improved accuracy. The use of multiple performance metrics ensures a balanced evaluation, highlighting the system's ability to minimize false positives while maintaining a high detection.

However, the effectiveness of the system largely depends on the quality and diversity of the training dataset, as well as the selection of appropriate algorithms. While the model shows strong potential in handling large-scale IoT data, challenges such as computational overhead, scalability, and adapting to evolving attack techniques still remain.

Despite these limitations, the system provides a significant improvement over traditional methods by enabling early detection, faster response, and enhanced protection of IoT networks. Overall, it represents a promising solution for strengthening cybersecurity in modern IoT ecosystems, with scope for further optimization and real-world deployment.

X. CONCLUSION

In conclusion, the detection of cyber attack traces in IoT data is a critical requirement for ensuring the security and reliability of modern interconnected systems. The proposed methodology effectively addresses this challenge by combining data preprocessing, machine learning techniques, and real-time monitoring to identify malicious activities within large-scale IoT environments.Although certain challenges such as handling massive data volumes, adapting to evolving threats, and computational constraints still exist, the proposed approach significantly improves the ability to detect and respond to cyber attacks at an early stage.

Overall, the system enhances the security framework of IoT networks and provides a strong foundation for future advancements in intelligent and adaptive cybersecurity solutions.

XI.RESULT AND FUTURE SCOPE

Figure 1: GU/Main Interface of Detection of cyber attacks traces in Iot Data

Figure 2: The above screen shows the final output of the Detection of cyber attacks in iot data

The proposed system generates meaningful and actionable outputs that help in identifying and responding to cyber attacks in IoT environments. The primary output is the classification of IoT data into normal or malicious categories, enabling clear identification of potential threats. In cases where an attack is detected, the system further specifies the type of attack, such as DDoS, intrusion, or malware activity.

Additionally, the system provides real-time alerts and notifications, allowing users or administrators to take immediate action to prevent damage. The outputs are also visualized through graphs, dashboards, or reports, making it easier to understand patterns, trends, and system behavior. Furthermore, the system may generate performance reports that include evaluation metrics like accuracy and detection rate, helping to assess the effectiveness of the model. Overall, these outputs support timely decision-making, enhance situational awareness, and improve the overall security of IoT networks.

XII. References

[1][da Costa et al., 2019] da Costa, K. A., Papa, J. P., Lisboa, C. O., Munoz, R., and de Albuquerque, V. H. C. (2019). Internet of things: A survey on machine learning based intrusion detection approaches. Computer Networks, 151:147–157.

[2] [D'Angelo et al., 2020] D'Angelo, G., Ficco, M., and Palmieri, F. (2020). Malware detection in mobile environments based on Autoencoders and API-images. Journal of Parallel and Distributed Computing, 137:26–33.

[3][Abolhasanzadeh, 2015] Abolhasanzadeh, B. (2015). Nonlinear dimensionality reduction for intrusion detection using auto-encoder bottleneck features. In 2015 7th Conference on Information and Knowledge Technology (IKT), pages 1–5. IEEE.

[4] [Agustin et al., 2020] Agustin, P., Sebastian, G., and Maria Jose, E. (2020 (accessed February 3, 2020)). Stratosphere laboratory. a labeled dataset with malicious and benign iot network traffic. https://www.stratosphereips.org/datasets-iot23.

[5] [Al-Qatf et al., 2018] Al-Qatf, M., Lasheng, Y., Al-Habib, M., and Al-Sabahi, K. (2018). Deep learning approach combining sparse autoencoder with svm for network intrusion detection. IEEE Access, 6:52843–52856.

[6][Arel et al., 2009] Arel, I., Rose, D., and Coop, R. (2009). Destin: A scalable deep learning architecture with application to high-dimensional robust pattern recognition. In 2009 AAAI Fall Symposium Series.

[7][Barut et al., 2020] Barut, O., Luo, Y., Zhang, T., Li, W., and Li, P. (2020). Netml: A challenge for network traffic analytics. arXiv preprint arXiv:2004.13006.

[8] [Berman et al., 2019] Berman, D., Buczak, A., Chavis, J., and Corbett, C. (2019). A Survey of Deep Learning Methods for Cyber Security. Information, 10(4):122.

[9][Bieniasz et al., 2019] Bieniasz, J., Stepkowska, M., Janicki, A., and Szczypiorski, K. (2019). Mobile agents for detecting network attacks using timing covert channels. J. UCS, 25(9):1109–1130.

[10]D'Angelo and Palmieri, 2021] D'Angelo, G. and Palmieri, F. (2021). Network traffic classification using deep convolutional recurrent autoencoder neural networks for spatial–temporal features extraction. Journal of Network and Computer Applications, 173:102890.