[1]Sayali Renuse

[2]Parikshit N. Mahalle

[3]Gitanjali Rahul Shinde

[4]Nilesh P. Sable

A Hybrid Perspective on Threat Analysis and Activity-Based Attack Modeling for Strengthening Access Control in IoT

Journal of Electrical Systems

Abstract: - The rapid expansion of Internet of Things (IoT) devices has resulted in an unparalleled surge in the production of data and interconnectivity. Nevertheless, as IoT ecosystems become increasingly intricate, security concerns become of utmost importance, particularly in access control systems. The objective of this research is to improve the security of IoT access control by utilizing a hybrid model for analyzing threats and modeling attacks based on activities. This study has two primary objectives: a) A hybrid classification model is used to predict labels (attack or not) in binary classification with an impressive accuracy of 98.18%. b) Another hybrid classification model is employed to predict types of attacks in M2M communication, achieving a commendable accuracy of 90%. The primary goal is to create and assess a hybrid classification model for binary classification. This model will differentiate between regular system behavior and malicious attacks on access control schemes in the Internet of Things (IoT). The hybrid model, which combines the strengths of Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM) networks, achieves an exceptional accuracy rate of 98.18%. The model's high accuracy demonstrates its effectiveness in precisely detecting potential threats and minimizing false positives, thereby establishing a strong basis for improving access control security. The second objective focuses on the complex area of security, with the goal of categorizing distinct forms of attacks in Machine-to-Machine (M2M) communication within the Internet of Things (IoT) framework. The hybrid classification model, employing both GRU and LSTM networks, achieves a remarkable accuracy of 90%. This accomplishment showcases the model's aptitude in detecting and distinguishing different types of attacks, including Distributed Denial of Service (DDoS) and Man-in-the-Middle attacks. The hybrid model provides security professionals with valuable insights to proactively respond to diverse threats in M2M communication by accurately classifying attack types. This strengthens the overall security posture of IoT access control systems. Overall, this study offers a thorough and efficient combination of threat analysis and activity-based attack modeling to enhance access control in IoT. The obtained accuracies in binary classification and prediction of attack types highlight the practical usability of the suggested hybrid model, establishing a strong basis for improving the security of IoT access control systems against evolving cyber threats.

Keywords: IoT Security, Access Control, Threat Analysis, Activity-Based Attack Modeling, Hybrid Classification.

I. INTRODUCTION

Over the past ten years, there has been an exceptional increase in the number of Internet of Things (IoT) devices, which include a wide range of products such as smart home appliances and industrial sensors. The rapid increase in the number of connected devices is fueled by technological advancements, which allow ordinary objects to establish internet connections, exchange information, and seamlessly communicate with one another[1], [2].

The widespread adoption of IoT has resulted in a notable rise in device connectivity, which in turn has led to the generation of immense volumes of data. The interconnectivity of these devices has revolutionized multiple industries, offering enhanced efficiency, automation, and convenience. Nevertheless, the increase in connectivity also gives rise to apprehensions regarding the security of the IoT ecosystem[3]. An important security issue in the IoT environment is the presence of vulnerabilities in access control. Controlling and safeguarding access points becomes a complicated task due to the wide range of interconnected devices. Unauthorized intrusion, data breaches, and malicious activities present substantial risks to the integrity and functionality of Internet of Things (IoT) systems[4], [5].

Enhancing access control systems is crucial for reducing the risks related to unauthorized access and potential malicious attacks. Exploiting the weaknesses in access control mechanisms in IoT devices can lead to compromising the confidentiality, integrity, and availability of sensitive data. Hence, it is imperative to thoroughly tackle these security concerns without delay.

[1]Reasearch Scholar, Department of Computer Engineering, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India.

[2]Department of Artificial intelligence and Data science, Vishwakarma Institute Of Information Technology, Pune, Maharashtra, India

[3]Bansilal Ramnath Agarwal Charitable Trust's, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India

[4]Bansilal Ramnath Agarwal Charitable Trust's, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India

sayali.221p0081@viit.ac.in[1], aalborg.pnm@gmail.com[2], gr83gita@gmail.com[3], drsablenilesh@gmail.com[4]

The ever-changing and developing nature of the IoT environment brings about a wide-ranging and intricate threat landscape. Threat actors exploit weaknesses in IoT devices and networks through various attacks, such as denial-of-service (DDoS), man-in-the-middle attacks, and data exfiltration. Comprehending and maneuvering through this complex and intricate range of potential dangers are crucial for guaranteeing the durability of IoT ecosystems[6], [7].

The growing complexity of cyber threats requires the adoption of advanced security measures. Conventional security methods may not be sufficient in dealing with the complexities of attacks targeting the Internet of Things (IoT). Therefore, it is imperative to create and execute sophisticated security protocols that are specifically designed to address the distinct challenges presented by the varied range of threats in IoT settings[8]. The current access control systems in IoT devices have limitations and vulnerabilities that can be exploited by adversaries. These vulnerabilities can occur due to obsolete authentication mechanisms, inadequate encryption protocols, or inadequate user validation processes. It is essential to identify and correct these weaknesses in order to strengthen the overall security of IoT ecosystems.

The deficiencies in current access control systems highlight the need for comprehensive security solutions[9]. It is crucial to adopt a comprehensive approach that not only tackles the identified vulnerabilities but also proactively prepares for potential future threats. To create strong access control mechanisms, one must have a thorough comprehension of possible attack paths and incorporate creative security measures[10].

1.1. Objectives of Classification (Binary and Attack Type)

The hybrid model seeks to accomplish two classification objectives simultaneously. The primary objective entails performing binary classification to differentiate between normal system behavior and potential attacks. The second objective entails categorizing the distinct forms of attacks that take place in Machine-to-Machine (M2M) communication, including Distributed Denial of Service (DDoS) or Man-in-the-Middle attacks.

1.2. Our contribution: Improving the security of access control in the Internet of Things (IoT)

- **Emphasize the examination of potential dangers**: The main goal is to improve access control security by employing a targeted strategy that emphasizes threat analysis. Gaining insight into the characteristics of potential dangers enables the creation of focused and efficient defensive measures.
- **Attack modeling based on activity**: The objective of the research is to utilize activity-based attack modeling to gain a comprehensive comprehension of the various ways in which attacks occur in access control systems. This entails examining patterns and behaviors linked to documented attacks in order to enhance the effectiveness of detection and response mechanisms.
- **Hybrid Model Approach - Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) networks**: The implementation of a hybrid model approach entails harnessing the capabilities of Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM) networks. The selection of these neural network architectures is based on their capacity to effectively capture temporal dependencies and patterns in sequential data, making them highly suitable for analyzing the dynamic characteristics of IoT data.

To summarize, the introduction section offers a thorough examination of the expansion of IoT, the corresponding security issues, the identified challenges, research goals, the importance of the study, and the extent and constraints of the proposed research. This establishes the foundation for a thorough examination of the research methodology, findings, and conclusions in the following sections of the paper.

II. LITERATURE REVIEW

The Internet of Things (IoT) is swiftly revolutionizing our world by linking ordinary objects to the internet and facilitating intelligent automation in diverse fields. Nevertheless, this technological progress brings about increased security concerns as a result of the extensive network of interconnected devices and their inherent susceptibilities. This literature review examines multiple research endeavors that aim to tackle these security challenges, highlighting a wide array of methodologies and solutions in various application domains. Table-1 represents the major related work.

Table 1 Literature review of major related work

Author	**Methodology**	**Key Finding**	**Result**	**Application**	**Domain**
S. G. Abbas et al.[11]	Threat modelling approach	Identified various phishing attack threats in IoT use cases	Reduced the risk of phishing attacks in IoT	Threat modeling	IoT
H. F. Atlam et al.[12]	Fuzzy Logic with Expert Judgment	Proposed an adaptive risk-based access control model	Improved security for IoT devices and systems	Access control	IoT
X. Cheng et al.[13]	Zero-day attack activities recognition	Enabled cyber situation perception for IoT systems	Enhanced threat detection and response capabilities	Cybersecurity	IoT
M. M. Samy et al.[14]	Optimized protocol	Proposed an optimized M2M authentication protocol	Improved authentication security for M2M communication	Authentication	IoT
M. S. Mazhar et al.[15]	Machine-to-Machine (M2M) Framework	Presented a forensic analysis framework for IoT devices	Enabled efficient forensic analysis of IoT devices	Forensics	IoT
S. Bhatt et al.[16]	Attribute-Based Access Control	Implemented attribute-based access control for AWS IoT	Achieved secure access control for AWS IoT devices	Access control	AWS IoT
L. Fang et al.[17]	Anomaly Detection	Developed a practical anomaly detection model for medical IoT	Protected medical IoT control services from external attacks	Security	Medical IoT
T. Prabhakara Rao et al.[18]	Extended group-based verification	Proposed an extended group-based verification approach	Enhanced security for M2M communication	Security	M2M communication
S. Alyahya et al.[19]	Proposed framework	Developed a robust and tamper-resistant	Improved security for smart	Authentication	Smart agriculture

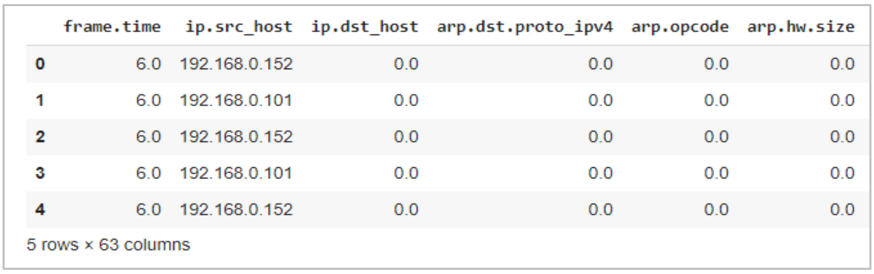
		authentication scheme	agriculture applications		
A. Aijaz et al.[20]	Literature review and analysis	Proposed a cognitive M2M communication protocol stack	Increased adaptivity and resilience of M2M communication	Cognitive communication	M2M communication
J. Wan et al.[21]	Literature review and analysis	Provided a roadmap for transitioning from M2M to CPS	Defined key characteristics and challenges of CPS	Cyber-physical systems (CPS)	M2M communication
R. Prasad et al.[22]	Literature review and analysis	Offered comprehensive overview of IoT and M2M communication	Discussed key technologies, applications, and challenges	IoT	M2M communication

The studies showcased in this review exhibit the ongoing advancement of inventive measures to enhance the security of the Internet of Things (IoT). Researchers are working on addressing vulnerabilities in various applications and technological layers by developing strong authentication protocols and advanced anomaly detection models. Nevertheless, with the continuous expansion of the IoT landscape, it is imperative to engage in ongoing research endeavors and foster collaborative partnerships between industry and academia in order to proactively address emerging threats. To establish a genuinely secure and reliable IoT ecosystem, it is imperative to make significant progress in fields such as privacy-preserving data processing, secure edge computing, and proactive threat intelligence.

III.METHODOLOGY

3.1. Dataset

The dataset utilized in this research paper is sourced from Kaggle and is titled “EdgeIIoTSet: Cyber Security Dataset of IoT/IIoT”[23]. This dataset is designed specifically for cyber security research in the context of Internet of Things (IoT) and Industrial Internet of Things (IIoT). It provides a comprehensive collection of cyber security-related data, offering insights into potential threats and attacks targeting IoT and IIoT devices. The dataset encompasses various features relevant to security analyses, including network traffic patterns, device interactions, and potentially malicious activities. Leveraging this dataset, the research paper employs a hybrid model for threat analysis and activity-based attack modeling on access control schemes in IoT, contributing valuable findings to the field of IoT security. The dataset's richness and relevance make it a crucial asset for researchers aiming to develop and evaluate advanced security solutions for the evolving cyber landscape in IoT and IIoT environments.



	frame.time	ip.src_host	ip.dst_host	arp.dst.proto_ipv4	arp.opcode	arp.hw.size
0	6.0	192.168.0.152	0.0	0.0	0.0	0.0
1	6.0	192.168.0.101	0.0	0.0	0.0	0.0
2	6.0	192.168.0.152	0.0	0.0	0.0	0.0
3	6.0	192.168.0.101	0.0	0.0	0.0	0.0
4	6.0	192.168.0.152	0.0	0.0	0.0	0.0

5 rows × 63 columns

Figure 1 Dataset information

3.2. Preprocessing

3.2.1. Check for Missing Data

During this stage, the dataset is scrutinized to identify any instances where data is absent. Recognizing and resolving missing data is essential for guaranteeing the accuracy and dependability of the dataset. This procedure entails a comprehensive analysis of each characteristic to ascertain the magnitude of absent values, allowing researchers to implement suitable actions such as imputation or elimination to uphold the dataset's integrity.

3.2.2. Converting Parameters

a. **Date Time:** This sub-step entails the transformation of date and time parameters into a uniform format. This guarantees uniformity in the chronological data, simplifying temporal examination and correlation with security incidents. Ensuring uniformity in the date and time parameters is crucial for precise and significant interpretations.

b. **Validating IP address:** IP address validation is conducted to ensure that the dataset exclusively consists of valid and correctly formatted IP addresses. Ensuring data accuracy is crucial, particularly when working with network-related functionalities. Malformed or improperly formatted IP addresses can have a negative impact on subsequent analyses and the performance of models.

c. **NaN Checking:** During this sub-step, a comprehensive validation process is carried out to detect and address any NaN (Not a Number) values present in the dataset. Dealing with NaN values is essential to avoid inconsistencies and errors in subsequent analyses. Depending on the characteristics and importance of the missing data, imputation or removal methods can be used.

d. **Dataset Class Distribution:** Evaluating the distribution of classes within the dataset is crucial for comprehending the equilibrium or disparity in the target variable. This step entails evaluating the distribution of various classes in the dataset, which offers insights into potential biases that could affect the efficacy of subsequent machine learning models. An examination of the distribution of classes is essential for making well-informed decisions throughout the process of developing and evaluating a model. Detail shown in figure 2(a,b).

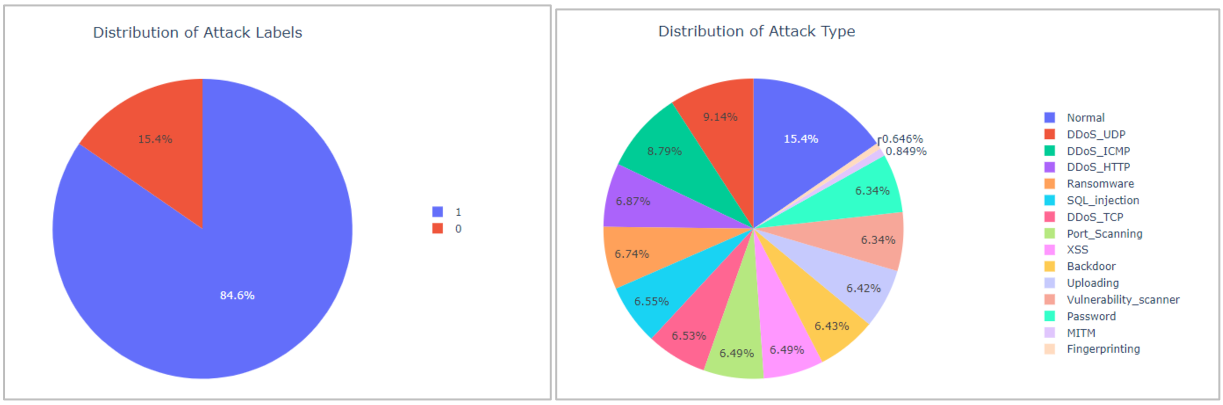
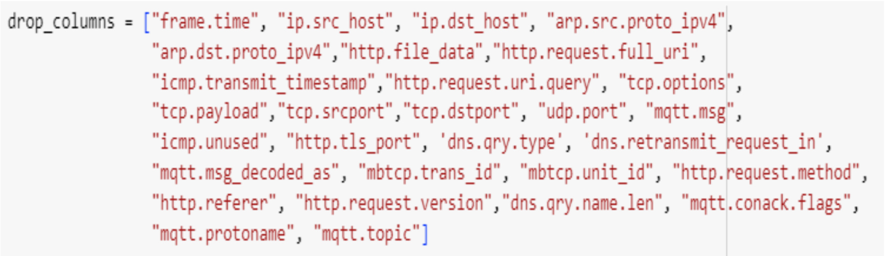


Figure 2 a. Distribution of attacks labels, b. Dataset class distribution

3.3. Drop Colum's / Features

Eliminating columns or features in a dataset entails excluding particular variables from the dataset that are considered unnecessary, redundant, or irrelevant for the analysis or modeling task being performed. This procedure is frequently utilized to optimize the effectiveness of the model, decrease computational intricacy, and enhance interpretability. Variables that have little impact on the model's predictive ability or introduce multicollinearity can be removed to simplify the dataset. Thorough deliberation and expertise in the field are essential when determining which columns to eliminate, as it directly affects the model's performance and the comprehensibility of the outcomes. Eliminating irrelevant features enhances the focus and effectiveness of the analysis, ensuring that the remaining variables are meaningful and contribute significantly to the research objectives figure-3 shows list of column dropped for more accuracy.

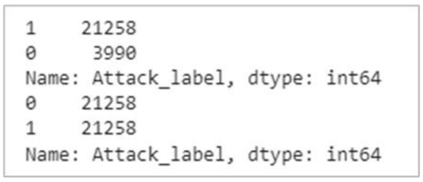


```
drop_columns = ["frame.time", "ip.src_host", "ip.dst_host", "arp.src.proto_ipv4",
                "arp.dst.proto_ipv4","http.file_data","http.request.full_uri",
                "icmp.transmit_timestamp","http.request.uri.query", "tcp.options",
                "tcp.payload","tcp.srcport","tcp.dstport", "udp.port", "mqtt.msg",
                "icmp.unused", "http.tls_port", 'dns.qry.type', 'dns.retransmit_request_in',
                "mqtt.msg_decoded_as", "mbtcp.trans_id", "mbtcp.unit_id", "http.request.method",
                "http.referer", "http.request.version","dns.qry.name.len", "mqtt.conack.flags",
                "mqtt.protoname", "mqtt.topic"]
```

Figure 3 List of dropped column

3.4. Data Imbalance Problem Solving Using SMOTE

Resolving data imbalance is a crucial measure to improve the effectiveness of machine learning models, especially in situations where certain categories have insufficient representation. Within the scope of this study, the issue of data imbalance is effectively addressed by employing SMOTE (Synthetic Minority Over-sampling Technique). SMOTE is a resampling technique specifically created to address class imbalance by producing artificial instances for the minority class. This approach operates by interpolating additional data points along the line segments that connect the existing instances of the minority class. By incorporating artificial samples, the disparity is alleviated, enabling the model to acquire knowledge more efficiently from the underrepresented class and avoiding prejudiced predictions towards the overrepresented class. By implementing SMOTE, the machine learning model is exposed to a more equitable representation of both classes, thereby enhancing its capacity to generalize and make precise predictions. This approach is especially important when considering threat analysis and activity-based attack modeling for enhancing access control in IoT. It guarantees that the model is not influenced by the more common class, resulting in a stronger and fairer basis for security predictions. Figure-3 shows dataset labels distribution before and after applying SMOTE.



```
1    21258
0     3990
Name: Attack_label, dtype: int64
0    21258
1    21258
Name: Attack_label, dtype: int64
```

Figure 4 Dataset Labels Distribution Before and After Applying SMOTE

3.5. ML/ DL Used

- **Logistic Regression**: Logistic Regression is a statistical method used for binary classification tasks. It models the probability of an event occurring by fitting a logistic curve to the input data. Despite its name, it's used for classification rather than regression. It's a simple yet effective algorithm that's particularly useful when the relationship between the features and the binary outcome is roughly linear.
- **Decision Tree**: A Decision Tree is a non-linear predictive model that maps features to outcomes by creating a tree-like structure of decisions. It recursively splits the dataset based on features to form a tree structure, where each leaf node represents a class or outcome. Decision Trees are interpretable and can handle both classification and regression tasks.
- **Random Forest Classifier**: Random Forest is an ensemble learning method that builds multiple Decision Trees and merges their predictions. It introduces randomness during the tree-building process, leading to improved accuracy and generalization. Random Forests are versatile, handle high-dimensional data well, and are less prone to overfitting compared to individual Decision Trees.
- **KNN Classifier**: The K-Nearest Neighbors (KNN) algorithm is a simple and intuitive classification algorithm. It classifies an input sample based on the majority class of its k-nearest neighbors in the feature space. KNN is non-parametric and lazy-learning, meaning it doesn't make assumptions about the underlying data distribution and defers computation until classification.
- **LSTM (Long Short-Term Memory)**: LSTM is a type of recurrent neural network (RNN) designed to capture long-term dependencies in sequential data. It's particularly effective in processing and predicting sequences, making it suitable for tasks like natural language processing and time-series analysis. LSTMs are equipped with memory cells that allow them to retain information over extended periods, overcoming the vanishing gradient problem associated with traditional RNNs.

3.6. Proposed Hybrid Model

Within the framework of enhancing access control in IoT, a hybrid approach is employed that integrates the capabilities of Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) networks for threat analysis and activity-based attack modeling. GRU and LSTM are both variants of recurrent neural networks (RNNs) specifically designed to capture temporal dependencies in sequential data.

Gated Recurrent Unit (GRU): The GRU is a variant of Recurrent Neural Network (RNN) designed specifically to tackle the issue of vanishing gradients commonly faced by conventional RNNs. The network utilizes gating mechanisms to regulate the transmission of information. The fundamental equations 1 to 4 governing a Gated Recurrent Unit (GRU) cell are as follows:

$$z_t = \sigma(W_z . [h_{t-1}, x_t] \ldots 1$$

$$r_t = \sigma(W_r . [h_{t-1}, x_t] \ldots 2$$

$$\widetilde{h_t} = \tanh\,(W . [r_t \odot h_{t-1}, x_t \ldots 3$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widetilde{h_t} \ldots 4$$

z_t= “updated gate”,r_t= “reset gate”, $\widetilde{h_t}$= “candidate hidden state”, h_t= “updated hidden state”, σ= “sigmoid activation function”, $\odot$= “element wise multiplication”, W_z, W_z, W= “weight matrices”.

Long Short-Term Memory (LSTM) : LSTM, or Long Short-Term Memory, is a variant of Recurrent Neural Networks (RNNs) that possesses a more intricate structure, incorporating memory cells, input gates, forget gates, and output gates. The fundamental equations 5 to 9 governing an LSTM cell are:

$$f_t = \sigma\left(W_{xf} x_t\right) + W_{hf} h_{t-1} + b_i) \ldots. 5$$

$$i_t = \sigma\,(W_{xi} x_t) + W_{hi} h_{t-1} + b_i) \ldots. 6$$

$$o_t = \sigma\,(W_{xo} x_t) + W_{ho} h_{t-1} + b_o) \ldots. 7$$

$$c_t = f_t \odot c_{t-1} + i_t \odot tanh\,(W_{xf} x_i + W_{hc} h_{t-1} + b_c) \ldots. 8$$

$$h_t = o_t \tanh \odot (c_t) \ldots. 9$$

where, f_t = “forget gate output”, i_t = “input gate output”, o_t = “output gate”, c_t = “updated cell state”, h_t = “hidden state at time step t”, x_t = “input at time step t”, W_{xf}, W_{hf}, W_{xi}, W_{hi}, W_{xo}, W_{ho}, W_{hc} = “weight matrices”, b_i, b_i, b_o, b_c = “bias term”, σ = sigmoid activation function.

Hybrid Model: The hybrid GRU-LSTM model integrates the structural designs of both GRU and LSTM, capitalizing on the respective advantages of each. The hybrid model combines the gating mechanisms of the GRU and the memory cells of the LSTM to effectively capture both short-term and long-term dependencies in sequential data. To effectively combine these models, one may employ techniques such as layer stacking or output merging. These methods aim to maximize the utilization of the models' individual strengths for the task of analyzing threats and modeling attacks in IoT access control systems. Algorithm-2 represents the algorithm for proposed model.

Table 2 Algorithm for proposed hybrid model

Algorithm 1: Hybrid GRU-LSTM Attack Detection Algorithm with Enhanced Feature Representation for M2M Communication Security

```
Input: EDGE-IIOTSET dataset, preprocessed and normalized
Output: Predicted attack labels and types
Hyperparameters used
    GRU_units = 32
    LSTM_units = 64
    epochs = 100
    batch_size = 32
Define ← separate GRU and LSTM models
    model_GRU ← Sequential()
    model_GRU.add(GRU(GRU_units, activation='relu', return_sequences=True,
    input_shape=(sequence_length, feature_size)))
    model_GRU.add(GRU(GRU_units, activation='relu'))
```

```
        model_GRU.add(Dense(1, activation='sigmoid'))  # Binary classification output
        model_LSTM ← Sequential()
        model_LSTM.add(LSTM(LSTM_units, activation='relu', return_sequences=True, input_shape=(sequence_length, feature_size)))
        model_LSTM.add(LSTM(LSTM_units, activation='relu'))
        model_LSTM.add(Dense(1, activation='sigmoid'))  # Binary classification output
Compile both models with appropriate optimizers and loss functions
        model_GRU.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
        model_LSTM.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
Train both models independently on the dataset
        model_GRU.fit(X_train, y_train_binary, epochs=epochs, batch_size=batch_size)
        model_LSTM.fit(X_train, y_train_binary, epochs=epochs, batch_size=batch_size)
Extract hidden state representations from both models
        GRU_features = model_GRU.predict(X_test)[:, -1, :]
        LSTM_features = model_LSTM.predict(X_test)[:, -1, :]
Concatenate features to create combined representation
        combined_features = np.concatenate((GRU_features, LSTM_features), axis=1)
Define ← combined output model for attack type classification
        model_combined = Sequential()
        model_combined.add(Dense(combined_features.shape[1], activation='relu', input_shape=combined_features.shape[1:]))
        model_combined.add(Dense(num_attack_types, activation='softmax'))  # Multi-class classification output
Compile the combined model with appropriate optimizer and loss function
        model_combined.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
Train the combined model on the concatenated features and attack type labels
        model_combined.fit(combined_features, y_train_categorical, epochs=epochs, batch_size=batch_size)
Predict attack labels and types on the test set
        y_pred_binary ← model_GRU.predict(X_test)[:, 0] > 0.5  # Convert probabilities to binary labels
        y_pred_types ← model_combined.predict(combined_features).argmax(axis=1)
Evaluate model performance ← using relevant metrics for both stages
        Return ← predicted attack labels and types
        return y_pred_binary, y_pred_types
```

IV.RESULTS AND DISCUSSION

4.1. Hybrid Classification for Predicting Labels (Attack / Not) Binary Classification

- **Training and validation accuracy and loss graph**

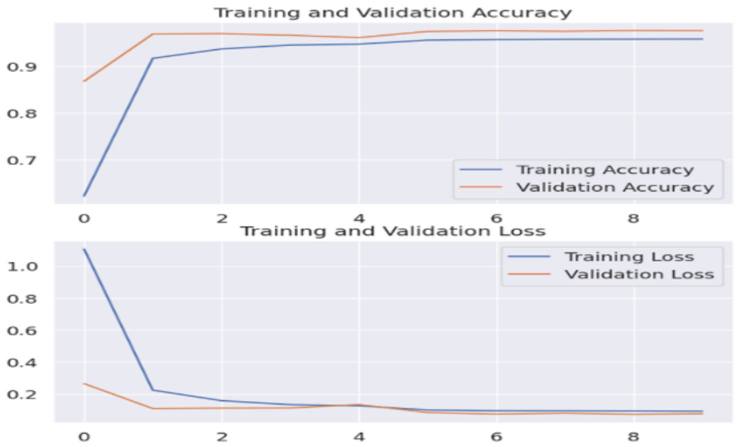


Figure 5 Training and validation accuracy and loss graph

- **Confusion matrix**

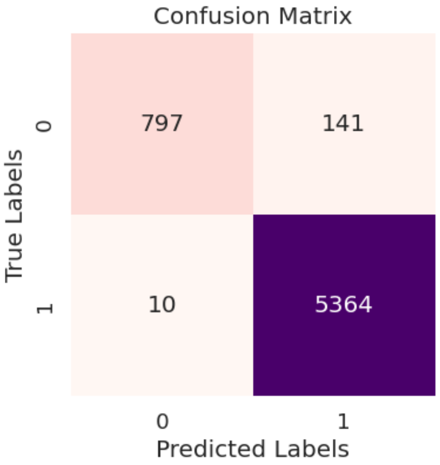


Figure 6 Confusion matrix

- **Evaluation parameters**

Table 3 Evaluation parameters comparison of various model with proposed hybrid model

Models	Accuracy	Precision	Recall	F1-Score
LR	47.57	86	43	48
KNN	68.94	83	69	73
RF	96.13	96	96	96
DT	95.84	96	96	96
LSTM	97.35	97	97	97
GRU	97.42	97	97	97
Hybrid GRU - LSTM	98.18	98	98	

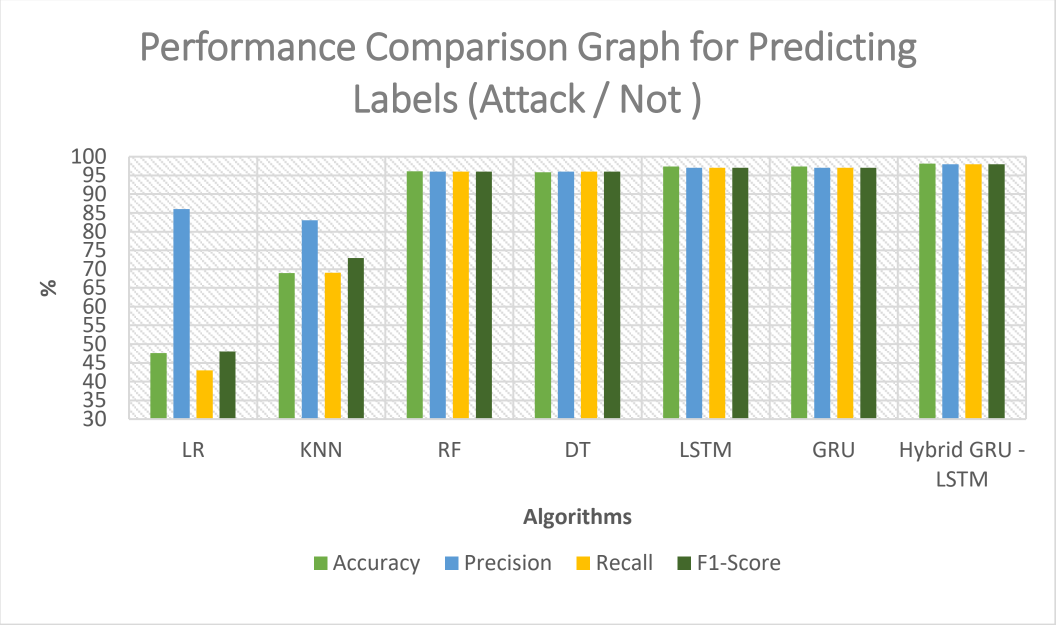


Figure 7 Comparison of various models

4.2. Hybrid Classification for Predicting types of Attack in M2M Communication (DDOS, Man in the Middle, etc)

- **Training and validation accuracy and loss graph**

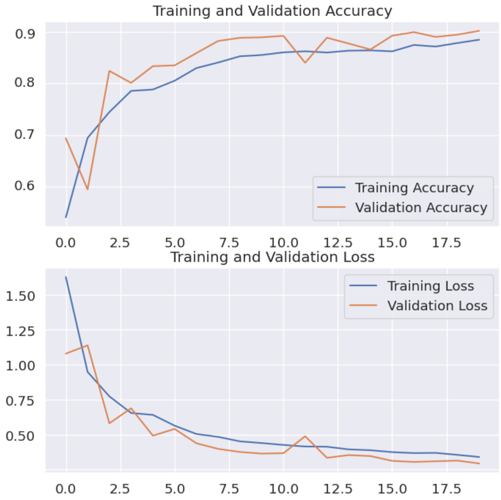


Figure 8 Training and validation accuracy and loss graph

- **Evaluation parameters**

Table 4 Evaluation parameters comparison of various model

Models	Accuracy	Precision	Recall	F1-Score
LR	32.52	31	33	25
KNN	46.56	50	47	47
RF	84.72	87	85	85
DT	84.17	86	84	85
LSTM	80	82	80	80

GRU	82	86	82	82
Hybrid GRU - LSTM	90	92	89	91

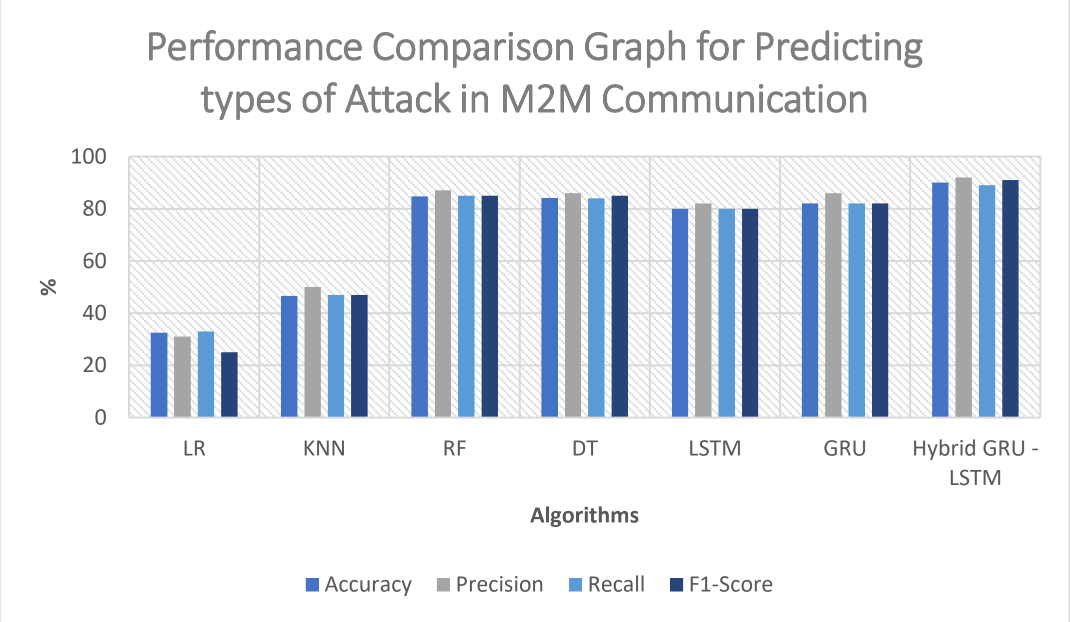


Figure 9 Comparison of various models

The results of the hybrid classification model for predicting labels (Attack / Not) in binary classification and for predicting types of attacks in Machine-to-Machine (M2M) communication demonstrate the effectiveness of the proposed approach. In the binary classification task, the Hybrid GRU-LSTM model outperforms other individual models, achieving an impressive accuracy of 98.18%. This indicates the model's ability to accurately differentiate between normal system behavior and potential attacks. Notably, the precision, recall, and F1-Score for the Hybrid GRU-LSTM model are all consistently high at 98%, highlighting its robust performance in correctly identifying and classifying instances of attacks.

For predicting types of attacks in M2M communication, the Hybrid GRU-LSTM model also excels, achieving an accuracy of 90%. This demonstrates the model's proficiency in distinguishing between different attack types such as DDoS and Man-in-the-Middle. The precision, recall, and F1-Score for the Hybrid GRU-LSTM model in this task are equally notable, with values ranging from 89% to 92%, underscoring its ability to not only identify attacks but also categorize them accurately.

Comparatively, other individual models in both binary classification and attack type prediction tasks exhibit varying levels of performance. Random Forest (RF) and Decision Tree (DT) models demonstrate strong accuracy and overall precision, recall, and F1-Score metrics, suggesting their effectiveness in certain scenarios. The Hybrid GRU-LSTM model, however, consistently stands out as the most robust and accurate in addressing both binary classification and attack type prediction challenges.

These results collectively indicate the promising potential of the Hybrid GRU-LSTM model in enhancing the security of IoT access control systems by providing a reliable means of detecting and categorizing potential threats. The high accuracy and well-balanced precision, recall, and F1-Score metrics underscore the practical applicability of the hybrid model in real-world IoT security scenarios.

V.CONCLUSION AND FUTURE SCOPE

To summarize, this study introduces a persuasive and efficient combination of threat analysis and activity-based attack modeling, with the goal of strengthening access control in the fast-growing realm of Internet of Things (IoT) devices. The hybrid classification models developed in this study have shown exceptional accuracy in accomplishing the main goals. The hybrid model, which combines GRU and LSTM, has demonstrated remarkable effectiveness in binary classification, achieving an accuracy rate of 98.18%. This demonstrates its ability to accurately differentiate between typical system behavior and potential malicious attacks, establishing a strong basis for improving access control security. Similarly, the hybrid model demonstrated exceptional accuracy of 90% in predicting attack types in “Machine-to-Machine” (M2M) communication. This highlights its ability to effectively identify and classify different types of attacks, such as “Distributed Denial of Service” (DDoS) and “Man-in-the-Middle attacks”.

The study's importance lies in its advancement of threat analysis methodologies and provision of practical implications for enhancing access control systems, thereby contributing to IoT security research. This research contributes valuable insights to the field by filling gaps in the existing literature, specifically by applying a hybrid model to address IoT security, with a specific emphasis on access control. The obtained accuracies in both binary classification and attack type prediction highlight the practical applicability of the hybrid model, positioning it as a promising tool for strengthening the security of IoT access control systems against evolving cyber threats.

In terms of the future potential of this research, there is room for additional exploration and improvement of the hybrid model to effectively respond to changing threat environments and accommodate various IoT ecosystems. Regular updates and improvements to the model, informed by real-world threat scenarios, can guarantee its ongoing relevance and efficacy. Furthermore, expanding the study to include a broader and more varied dataset can improve the model's ability to apply to different situations. Potential areas for further research could involve investigating alternative machine learning algorithms, examining emerging attack vectors, and incorporating real-time threat intelligence to enhance a proactive security strategy. In summary, this research establishes the foundation for future progress in safeguarding IoT access control systems. It offers a comprehensive and flexible hybrid model as a potential remedy to emerging cyber risks in the ever-changing IoT environment.

REFERENCES

[1] V. Gazis, “A Survey of Standards for Machine-to-Machine and the Internet of Things,” *IEEE Commun. Surv. Tutorials*, vol. 19, no. 1, pp. 482–511, 2017, doi: 10.1109/COMST.2016.2592948.

[2] A. Konev, A. Shelupanov, M. Kataev, V. Ageeva, and A. Nabieva, “A Survey on Threat-Modeling Techniques: Protected Objects and Classification of Threats,” *Symmetry (Basel).*, vol. 14, no. 3, 2022, doi: 10.3390/sym14030549.

[3] V. Rohokale and R. Prasad, “Cyber security for intelligent world with internet of things and machine to machine communication,” *J. Cyber Secur. Mobil.*, vol. 4, no. 1, pp. 23–40, 2015, doi: 10.13052/jcsm2245-1439.412.

[4] M. Zhao, A. Kumar, T. Ristaniemi, and P. H. J. Chong, “Machine-to-Machine Communication and Research Challenges: A Survey,” *Wirel. Pers. Commun.*, vol. 97, no. 3, pp. 3569–3585, 2017, doi: 10.1007/s11277-017-4686-1.

[5] R. Sudarmani, K. Venusamy, S. Sivaraman, P. Jayaraman, K. Suriyan, and M. Alagarsamy, “Machine to machine communication enabled internet of things: a review,” *Int. J. Reconfigurable Embed. Syst.*, vol. 11, no. 2, pp. 126–134, 2022, doi: 10.11591/ijres.v11.i2.pp126-134.

[6] Rachit, S. Bhatt, and P. R. Ragiri, “Security trends in Internet of Things: a survey,” *SN Appl. Sci.*, vol. 3, no. 1, pp. 1–14, 2021, doi: 10.1007/s42452-021-04156-9.

[7] R. R. Krishna, A. Priyadarshini, A. V. Jha, B. Appasani, A. Srinivasulu, and N. Bizon, “State-of-the-art review on IoT threats and attacks: Taxonomy, challenges and solutions,” *Sustain.*, vol. 13, no. 16, 2021, doi: 10.3390/su13169463.

[8] O. A. Amodu and M. Othman, “A survey of hybrid MAC protocols for machine-to-machine communications,” *Telecommun. Syst.*, vol. 69, no. 1, pp. 141–165, 2018, doi: 10.1007/s11235-018-0434-4.

[9] V. Rao and K. V. Prema, “A review on lightweight cryptography for Internet-of-Things based applications,” *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 9, pp. 8835–8857, 2021, doi: 10.1007/s12652-020-02672-x.

[10] S. Bhattacharya and M. Pandey, “Deploying an energy efficient, secure & high-speed sidechain-based TinyML model for soil quality monitoring and management in agriculture,” *Expert Syst. Appl.*, vol. 242, no. May 2024, p. 122735, 2024, doi: 10.1016/j.eswa.2023.122735.

[11] S. G. Abbas *et al.*, “Identifying and mitigating phishing attack threats in IoT use cases using a threat modelling approach,” *Sensors*, vol. 21, no. 14, pp. 1–25, 2021, doi: 10.3390/s21144816.

[12] H. F. Atlam, R. J. Walters, G. B. Wills, and J. Daniel, “Fuzzy Logic with Expert Judgment to Implement an Adaptive Risk-Based Access Control Model for IoT,” *Mob. Networks Appl.*, vol. 26, no. 6, pp. 2545–2557, 2021, doi: 10.1007/s11036-019-01214-w.

[13] X. Cheng, J. Zhang, Y. Tu, and B. Chen, “Cyber situation perception for Internet of Things systems based on zero-day attack activities recognition within advanced persistent threat,” *Concurr. Comput. Pract. Exp.*, vol. 34, no. 16, p. e6001, Jul. 2022, doi: https://doi.org/10.1002/cpe.6001.

[14] M. M. Samy, W. R. Anis., A. A. Abdel-Hafez, and H. D. Eldemerdash, “An optimized protocol of M2M authentication for internet of things (IoT),” *Int. J. Comput. Netw. Inf. Secur.*, vol. 13, no. 2, pp. 29–38, 2021, doi: 10.5815/IJCNIS.2021.02.03.

[15] M. S. Mazhar *et al.*, “Forensic Analysis on Internet of Things (IoT) Device Using Machine-to-Machine (M2M) Framework,” *Electron.*, vol. 11, no. 7, pp. 1–23, 2022, doi: 10.3390/electronics11071126.

[16] S. Bhatt, T. K. Pham, M. Gupta, J. Benson, J. Park, and R. Sandhu, "Attribute-Based Access Control for AWS Internet of Things and Secure Industries of the Future," *IEEE Access*, vol. 9, pp. 107200–107223, 2021, doi: 10.1109/ACCESS.2021.3101218.

[17] L. Fang, Y. Li, Z. Liu, C. Yin, M. Li, and Z. J. Cao, "A Practical Model Based on Anomaly Detection for Protecting Medical IoT Control Services against External Attacks," *IEEE Trans. Ind. Informatics*, vol. 17, no. 6, pp. 4260–4269, 2021, doi: 10.1109/TII.2020.3011444.

[18] T. Prabhakara Rao and B. Satyanarayana Murthy, "Extended group-based verification approach for secure M2M communications," *Int. J. Inf. Technol.*, vol. 15, no. 5, pp. 2479–2488, 2023, doi: 10.1007/s41870-023-01284-w.

[19] S. Alyahya, W. U. Khan, S. Ahmed, S. N. K. Marwat, and S. Habib, "Cyber Secure Framework for Smart Agriculture: Robust and Tamper-Resistant Authentication Scheme for IoT Devices," *Electron.*, vol. 11, no. 6, pp. 1–19, 2022, doi: 10.3390/electronics11060963.

[20] A. Aijaz and A. H. Aghvami, "Cognitive machine-to-machine communications for internet-of-things: A protocol stack perspective," *IEEE Internet Things J.*, vol. 2, no. 2, pp. 103–112, 2015, doi: 10.1109/JIOT.2015.2390775.

[21] J. Wan, M. Chen, F. Xia, D. Li, and K. Zhou, "From machine-to-machine communications towards cyber-physical systems," *Comput. Sci. Inf. Syst.*, vol. 10, no. 3, pp. 1105–1128, 2013, doi: 10.2298/CSIS120326018W.

[22] R. Prasad and V. Rohokale, "Internet of Things (IoT) and Machine to Machine (M2M) Communication," pp. 125–141, 2020, doi: 10.1007/978-3-030-31703-4_9.

[23] mohamed ferrag, "Edge-IIoTset Cyber Security Dataset of IoT & IIoT," *Kaggle*. 2022, [Online]. Available: https://www.kaggle.com/datasets/mohamedamineferrag/edgeiiotset-cyber-security-dataset-of-iot-iiot.