The Scientific Temper (2024) Vol. 15 (4): 3196-3204

E-ISSN: 2231-6396, ISSN: 0976-8653

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.29

https://scientifictemper.com/

RESEARCH ARTICLE

# Machine learning-based ERA model for detecting Sybil attacks on mobile ad hoc networks

**R. Kalaiselvi, P. Meenakshi Sundaram***

## Abstract

Mobile ad hoc networks provide a substantial security threat because they lack central management and sufficient resources. These networks function autonomously without any central authority regulating the inclusion or removal of nodes. Nodes have the autonomy to choose when to join or quit. Dynamic multi-hop networks, either stationary or mobile, provide quick and simple access to data. Predicting the evolution of MANET can be challenging due to the network's dispersion and self-organization, as well as its unpredictable and constantly changing topology. The independent organization of nodes in MANETs, coupled with their dispersion, may complicate the prediction of the network's future growth due to its unstable and constantly changing structure. A Sybil attack, a deceptive tactic, involves a small number of individuals creating multiple counterfeit identities to gain dominance over a substantial portion of the system. To deceive legitimate users into believing that their system is utilizing their identities, the malicious attacker node adopts numerous identities. An attacker aims to gather a substantial number of node IDs, potentially generated at random, to appear and function as distinct nodes. Within the peer-to-peer overlay, the enemy can approach a single object or a group of objects by adopting different identities. Mobile ad hoc networks are intrinsically less secure than wired networks due to inherent security vulnerabilities and limited energy resources. To enhance detection accuracy, it is recommended to employ an ensemble regression arboretum model, which is a type of machine learning prediction model. This study proposes a machine learning-based method for detecting Sybil attacks in MANETs by collecting network metrics such as traffic characteristics, communication patterns, and node behaviors.

**Keywords:** MANET, Sybil attack, Ensemble regressive arboretum model, Mobile ad hoc network, Machine learning.

## Introduction

Mobile area networks (MANETs) are a type of wireless network that do not use a central server to manage packet routing. Variations in communication range, pause durations, and speeds characterize these mobile nodes. Within the realm of inventive technology, there exists a capability to connect physical items through various communication networks in the digital world (Lain Baird *et al*., 2024). A mobile ad hoc network (MANET) is a group of mobile nodes that can communicate with each other either directly or through other nodes in the network. Nodes in MANETs can roam freely and run entirely on battery power. Because of this, a node can disappear or run out of power without alerting its neighbors. When constructing a network system, it is common practice to use a combination of networking protocols, such as MANET, which requires a specific order for the construction of routes by several links. The wireless connections between nearby nodes and the average lifespan of each node determine the system's lifetime. The decentralization and participation of all nodes in route discovery using ad hoc routing protocols improves the data dependability of routing (S. Harihara Gopalan *et al*., 2024). A MANET is formed by multiple compact and portable nodes that have the ability to communicate with one other or other nodes within the network. Increasing number of intermediate nodes along the edges of elements simplifies the process of generating elements of higher order. By linearly estimating the positions of the nodes at both ends of an internal edge, it may determine the position of an intermediate node. Mobile ad hoc networks allow nodes

Department of Computer Science, Maruthupandiyar College, (Affiliated to Bharathidasan University), Thanjavur, Tamil Nadu, India.

***Corresponding Author:** P. Meenakshi Sundaram, Department of Computer Science, Maruthupandiyar College, (Affiliated to Bharathidasan University), Thanjavur, Tamil Nadu, India, E-Mail: sundaramp994@gmail.com

**How to cite this article:** Kalaiselvi, R., Sundaram, P.M. (2024). Machine learning-based ERA model for detecting Sybil attacks on mobile ad hoc networks. The Scientific Temper, **15**(4):3196-3204.

Doi: 10.58414/SCIENTIFICTEMPER.2024.15.4.29

**Source of support:** Nil

**Conflict of interest:** None.



**Received:** 08/10/2024 **Accepted:** 27/11/2024 **Published** : 20/12/2024

to roam freely and operate on batteries. This means that a node can disappear or run out of power without telling the other nodes it's working with. Widespread deployment of MANETs necessitates solving problems including energy resource depletion, slow data transmission rates, and long transmission delays (Sabir Ali Changazi *et al*., 2024)

A hierarchical structure is applied to a MANET by means of an appropriate clustering technique in order to resolve these issues. Nodes that are part of a cluster are called member nodes. Factors like mobility, node degree, identity, and residual energy are used to select the cluster head for every cluster (Amol Vasudeva *et al*., 2022). Because it handles all cluster management functions, the representative node is often the first target of attackers. The structure of a dynamic network is always changing. When designing a network, it is crucial to fully account for the fact that network topology might vary greatly. Dynamic topology can adapt to network changes through the use of modulation and coding schemes, transmit power, wireless channel reassignment, wireless backhaul node manipulation, and central planning and assignment of appropriate frequencies, channel bandwidths, and device interfaces (Annu Govind *et al*., 2024). Researchers have recently centered their attention on improving mobile ad hoc networks, in which nodes communicate with one another to provide entertainment services in real-time. However, due to the nature of wireless connections and decentralized architecture, the development of secure routing in MANETs continues to be a significant obstacle. Topology and location-based protocols are two examples of classic MANET routing techniques that work in this direction. In addition to proactive and reactive/on-demand protocols, this collection of routing alternatives also included hybrid protocols (Shaik Shafi *et al*., 2023).

### *Literature Review*

Mobile ad hoc networks are crucial in the modern digital age as they provide wireless communication in rapidly changing and infrastructure-free contexts. Because of their distinct characteristics and decentralized design, MANETs were susceptible to a variety of attacks, including the Sybil attack. Sybil attacks present a significant threat to mobile ad hoc networks (Bhupender Kumar *et al*., 2020). A Sybil attack originates when a malevolent node creates numerous suspicious nodes that seem to be separate entities by impersonating different identities inside the network. A comprehensive literature review on Sybil attack detection on MANETs is presented in this work. Additionally, current surveys are examined to identify any gaps in the research. By utilizing a multitude of false identities, it is possible to circumvent a malicious node's reputation algorithms and artificially inflate their trust scores (Brennan Huber *et al*., 2023). Nodes with a high trust value are able to acquire more access to network resources because they are classified as legitimate, even though they are malicious. Consequently, network security is undermined, rendering networks vulnerable to Sybil attacks.

Currently, there is a discussion regarding mobile ad hoc networks, as well as any potential security vulnerabilities, risks, and solutions that may arise. Energy consumption is higher by nodes in wireless mobile ad hoc networks due to their dynamic nature and reliance on topology (S. Sarika *et al*., 2016). Because of their portability, wireless mobile ad hoc networks are more vulnerable to attacks that try to disable the network totally or partially. It is, therefore, crucial to have a firm grasp of the myriad challenges posed by wireless mobile networks.

The potential new risks associated with cyberattacks utilizing drones and strategies for preventing them (Jean-Paul Yaacoub *et al*., 2020). Exploiting weaknesses in communication channels, smart devices, and hardware, namely smartphones and tablets. Following the hacking cycle, the authors illustrate an attack scenario that exemplifies their simulation of an assault on a particular drone. Ethical hackers should examine this to familiarize themselves with the current vulnerabilities in civilian and military unmanned aerial vehicles (UAVs). Furthermore, it provides them the opportunity to experiment with novel strategies and enhance their defenses against UAV assaults. Thus, countermeasures against drones (both detection and prevention) from civilian and military sectors will be explored.

Details regarding these attack types, including poisoning and inference assaults, their classifications, and operational methodologies inside a federated learning framework (Akarsh K. Nair *et al*., 2023), an exact synopsis of the problems with Federated Machine Learning's security and possible remedies. There are a number of security risks that could arise from using this technique, including data loss, communication problems, poisoning, manipulation of backdoor systems, and others. These types of assaults can be classified in several ways according to their operational mechanisms.

### *Mobile Ad Hoc Security Vulnerabilities*

Mobile ad hoc networks are vulnerable to attacks because it lack safe bounds. Any node within the radio range of any other node in the network can launch an all-weather attack on the mobile ad hoc network, and the attacker can choose which node(s) to target (Naveen Kumar *et al*., 2024). These nodes spontaneously arrange themselves in patterns that are both random and unpredictable. Having a wireless system that can transfer data across locations while considering the mobility of the nodes is crucial in this situation. As a result, any data packets sent to a node within its frequency range will be received by it. Consequently, the receiving node is permitted to deviate from the frequency range at its discretion, even as the nodes remain mobile. In areas without such infrastructure, it opens the opportunity

for people and devices to interconnect (Abu Jahid *et al*., 2022).

### *Lack of Secure Boundaries*

To compromise a wired network, an attacker would need to physically access the medium that the network uses. They might even be required to pass through multiple gateway and firewall levels (Christopher Morales Gonzalez *et al*., 2024). However, as long as the node is inside the network's frequency range, accessing the network in MANETs is a breeze. Therefore, MANETs are not a safe way to establish boundaries.

### *Power and Computational Limitations*

The availability of electric power supplies is not an issue for wired networks, but it is a limitation of wireless networks. Therefore, if a node in a network has a limited amount of electricity, it may behave selfishly.

### *Failure of Centralized Management Facility*

Ad hoc networks have specific security flaws because they don't have a centralized management structure. It is not possible to monitor and regulate the traffic in a large-scale, extremely dynamic ad hoc network, and the absence of centralized management equipment makes attack detection a particularly challenging task.

### *Cooperativeness*

Routing algorithms in MANETs typically operate under the premise that nodes in the network are cooperative and not intentionally malevolent. So, by deliberately breaking protocol specifications, an attacker can simply become a crucial routing agent and disrupt network operations.

### *Description of the Proposed Process*

For learning issues involving a numerical target variable, ensemble regression integrates many models to enhance prediction accuracy. The three steps of ensemble learning for regression are generation, pruning, and integration. Induction involves generating a collection of candidate models, pruning involves selecting a subset of those models, and integration involves integrating the outputs of the models to obtain a prediction. Classification problems have dominated ensemble learning studies (Hong Li *et al*., 2024). Nevertheless, methods that work well for classification aren't necessarily going to work for regression. So, although they go hand in hand, ensemble learning approaches have evolved in their own unique ways.

## Methodology

This methodology outlines a systematic approach for developing and evaluating an ensemble regressive arboretum model for detecting Sybil attacks on mobile ad hoc networks based on machine learning techniques.

### *Data Collection*

The attribute specifications for a dataset related to the detection of Sybil attack. The main characteristics and their relevance to Sybil attack detection are explained in Table 1.

The dataset comprises various attributes that capture essential information about network connections and user behavior. Duration, a real-valued attribute, denotes the duration of the connection, offering insights into the length of interactions. Categorical attributes like protocol_type specify the type of protocol used, while service indicates the network service employed, aiding in understanding the nature of connections. The flag attribute provides status information, crucial for identifying abnormal connection states. Real-valued attributes src_bytes and dst_bytes quantify the volume of data transferred, potentially indicating malicious data exfiltration. The binary attribute land flags connections originating and terminating at the same host/port, a characteristic often exploited in land attacks. Another binary attribute, logged_in, denotes whether users are authenticated, essential for detecting unauthorized access attempts. Num_compromised, a real-valued attribute, quantifies the extent of compromised conditions, adding depth to anomaly detection. Finally, the categorical attribute class serves as the target

**Table 1:** Attributes of the Sybil attack detection dataset

| *Attributes* | *Type of attribute* | *Explanation* |
|---|---|---|
| Duration | Real-valued attribute | Represents the duration of the connection |
| protocol_type | Categorical attribute | Indicates the protocol type used (tcp, udp, icmp). |
| service | Categorical attribute | Defines the service being utilized by the network. |
| flag | Categorical attribute | This signifies the connection's state. |
| src_bytes and dst_bytes | Real-valued attributes | Represent the number of source and destination bytes. |
| land | Binary attribute (0 or 1) | Indicates if the connection is from/to the same host/port. |
| logged_in | Binary attribute (0 or 1) | Indicates if the user is logged in. |
| num_compromised | Real-valued attribute | Indicates the number of compromised conditions. |
| Class | Categorical attribute | Defines whether the connection is 'normal' or an 'anomaly' (Sybil attack). |

variable, distinguishing between 'normal' connections and 'anomalies' such as Sybil attacks, forming the basis for classification tasks in network security analysis.

### *Workflow Architecture*

Figure 1 illustrates the proposed ERA model, designed specifically for detecting Sybil attacks on mobile ad hoc networks (MANETs). The ERA Model begins its process by importing essential libraries such as scikit-learn and networks. It then establishes a function to extract features from the MANET dataset, including node degree, average neighbor degree, clustering coefficient, betweenness centrality, and other pertinent features derived from the network topology. After generating a labeled dataset containing both normal and Sybil nodes, the ERA Model partitions it into training and testing subsets. To train the classifier, it utilizes the features retrieved from the training data and the ERA Model with decision tree weak learners. Evaluation metrics, such as accuracy, precision, recall, and F1-score, are computed on the testing subset. ERA model is being operationalized with real-time Sybil attack detection by continuously monitoring the network, extracting features of new nodes, and predicting their status using the trained ERA model classifier. It concludes by visually representing the network to highlight identified Sybil nodes and emphasizes the necessity for periodic updates to adapt the classifier to evolving network dynamics.

### *Preprocessing and Extract features*

There are a number of stages involved in preprocessing a dataset to make it ready for analysis and modelling by cleaning and standardizing the data. For the attributes provided in the table, the following preprocessing steps can be applied:

*Handling missing values*

Check for missing values in each attribute. If any missing values are found, impute them using appropriate strategies such as mean, median, mode imputation.

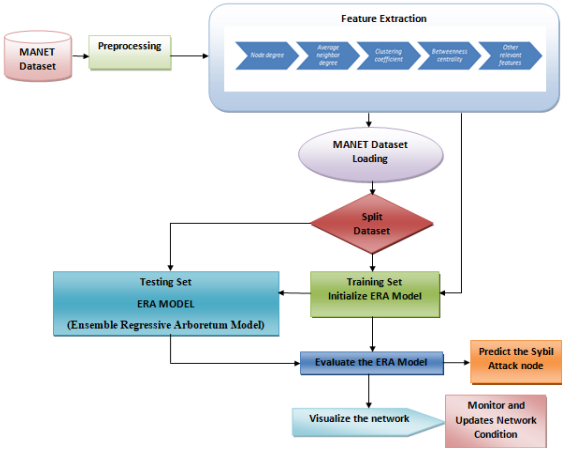


**Figure 1:** Workflow of the proposed ERA model

*Encoding categorical attributes*

For properties like protocol_type, service, and flag, use one-hot encoding to convert category variables to numerical ones. This ensures that categorical attributes can be effectively used in machine learning algorithms.

*Scaling numerical attributes*

Scale numerical attributes like Duration, src_bytes, dst_bytes, and num_compromised to bring them onto a similar scale. This can be done using techniques like Min-Max scaling, which helps in improving the convergence of machine learning algorithms.

*Handling binary attributes*

Binary attributes like land and logged_in can be kept as they are since they are already in a suitable format for modeling. Ensure that their values are consistent and properly interpreted in the context of the analysis.

*Feature selection*

Use feature selection to choose which features should be included in the study. This can be done using techniques like domain knowledge-based selection.Feature extraction involves identifying and extracting meaningful features from the dataset that are relevant to the task at hand, which in this case is the detection of Sybil attacks. For the given dataset, features such as node degree, average neighbor degree, clustering coefficient, and betweenness centrality can be derived from the network topology represented by attributes like protocol_type, service, flag, src_bytes, dst_bytes, land, logged_in, and num_compromised.

*Handling class imbalance*

If you see that one class ('anomaly') is much underrepresented in comparison to the other ('normal') in the target variable (Class), you should look into the possibility of a class imbalance. In cases where class imbalance is detected, methods including oversampling, under sampling, and synthetic data generation can be employed.

*Outlier detection and removal*

It is important to identify and remove any numerical attributes with unusually high or low values that may affect the performance of the models. The Z-score is one statistical method for identifying extreme values. The Z-score denotes the amount of standard deviations that a data point possesses from the mean of the collection. In most cases, data points are deemed outliers if their Z-score is greater than a predetermined threshold. The formula for calculating the Z-score of a data point $x$ in a numerical attribute is:

$$Z = \frac{x-\mu}{\sigma} \qquad (1)$$

Where, $x$ is the data point, μ is the mean of the dataset and σ is the standard deviation of the dataset. If |Z| > threshold, then the data point x is considered an outlier and can be removed or treated accordingly.

*Validation split*

For the purpose to train and test the model, divide the cleaned-up dataset in half. About 70% of the time is devoted to testing, and 30% to training.

By performing these preprocessing steps, the dataset is prepared for further analysis and modeling, ensuring that the machine learning algorithms can effectively learn from the data and make accurate predictions, particularly for detecting anomalies like Sybil attacks in mobile ad hoc networks.

## *Proposed Ensemble Regressive Arboretum Model (ERA Model)*

The proposed ensemble regressive arboretum model aims to detect Sybil attacks within MANET datasets, where the input consists of features from the MANET dataset, and the output is the prediction of whether a node is normal or a Sybil node. The Sybil attack detection model function is initialized by the algorithm. This function sets up a bunch of models and specifies a bunch of parameters, including how many estimators to use for the random forest model and how deep it can go, and how many estimators to use for the gradient boosting model and how fast it learns. This function takes in parameters and returns the initialized models when invoked.

The fit function encapsulates the fitting process, which consists of training three regression models (Linear Regression, Random Forest, and Gradient Boosting) using input data that has been separated into training and validation sets. After being added to the list of models, each model undergoes training using the training data. The predict function aggregates predictions from all models by computing the mean of their individual predictions.

*Random forest regression model*

Let $y^{rf}$ be the predictions from the random forest model. The formula is utilized throughout the training process of the random forest model:

$$y\hat{}rf = \mathrm{RF}(Xtrain, ytrain) \quad (2)$$

Here, RF represents the random forest model.

*Gradient boosting regression model*

Let y^gb be the predictions from the gradient boosting model. The gradient boosting model is trained using the formula:

$$y^{gb} = \mathrm{GB}(Xtrain, ytrain) \quad (3)$$

Here, GB represents the gradient boosting model.

*Linear regression model*

Let y^lr be the predictions from the linear regression model. The linear regression model is trained using the formula:

$$y^{lr} = \mathrm{LR}(Xtrain, ytrain) \quad (4)$$

Here, LR represents the linear regression model.

*Ensemble prediction*

The final prediction is the mean of predictions from all models:

$$\hat{}ensemble = \frac{1}{3}(y\hat{}rf + y\hat{}gb + y\hat{}lr) \quad (5)$$

In actuality, the algorithm receives the ensemble model and its parameters by loading and preprocessing the MANET data, and then it calls the SybilAttackDetectionModel function. After fitting the ensemble model to the data, it evaluates its performance on a test set. Metrics like as recall, accuracy, precision, and F1-score are calculated to assess the ensemble model's capability to accurately differentiate between normal and Sybil nodes in the MANET dataset. Figure 2 explains random forest regression model which is a powerful and versatile algorithm for regression tasks, capable of capturing complex relationships in data while mitigating overfitting. It aggregates predictions from

*Algorithm: Proposed Ensemble Regressive Arboretum Model*

*Input: MANET Dataset*

*Output: Predict Normal or Sybil node*

1. *FUNCTION SybilAttackDetectionModel(n_estimators_rf=100, max_depth_rf=NULL, learning_rate_gb=0.1, n_estimators_gb=100)*
2. *models = []*
3. *RETURN models, n_estimators_rf, max_depth_rf, learning_rate_gb, n_estimators_gb*
4. *FUNCTION fit(X, y)*
5. *X_train, X_val, y_train, y_val = train_test_split(X, y, test_size=0.2, random_state=42)*
6. *rf_regressor = RandomForestRegressor(n_estimators=n_estimators_rf, max_depth=max_depth_rf)*
7. *rf_regressor.fit(X_train, y_train)*
8. *ADD rf_regressor TO models*
9. *gb_regressor = GradientBoostingRegressor(learning_rate=learning_rate_gb, n_estimators=n_estimators_gb)*
10. *gb_regressor.fit(X_train, y_train)*
11. *ADD gb_regressor TO models*
12. *lr_regressor = LinearRegression()*
13. *lr_regressor.fit(X_train, y_train)*
14. *ADD lr_regressor TO models*
15. *RETURN models*
16. *FUNCTION predict(X)*
17. *predictions = []*
18. *FOR EACH model IN models*
19. *predictions.APPEND(model.predict(X))*
20. *RETURN MEAN(predictions)*
21. *X, y = load_and_preprocess_data()*
22. *ensemble_model, n_estimators_rf, max_depth_rf, learning_rate_gb, n_estimators_gb = SybilAttackDetectionModel()*
23. *models = ensemble_model.fit(X, y)*
24. *predictions = ensemble_model.predict(X_test)*
25. *accuracy = accuracy_score(y_test, predictions)*
26. *precision = precision_score(y_test, predictions)*
27. *recall = recall_score(y_test, predictions)*
28. *f1 = f1_score(y_test, predictions)*

**Figure 2:** Pseudo code of the proposed ensemble regressive arboretum model (ERA Model)

multiple decision trees to provide robust and accurate predictions for unseen data points.

### *Experimental Setup*

The setup for experimentally detecting Sybil attacks on MANETs through machine learning employs the network simulator (NS) tool. NS tool serves as a versatile and extensively utilized simulation platform, facilitating the modeling and analysis of network protocols and scenarios. In this particular study, NS tool is configured to replicate the dynamic and decentralized characteristics typical of mobile ad hoc networks. A range of machine learning algorithms are incorporated into the setup to scrutinize network behaviors and detect potential Sybil attackers. Through this experimental framework, the proposed detection methodology undergoes systematic evaluation, offering insights into its efficacy in fortifying the security of mobile ad hoc networks against Sybil attacks.

## Results and Discussion

The process of selecting source-destination pairings is a complicated one, which makes it difficult to get the network architecture properly. In order to analyze the behavior on the network and identify potential Sybil attackers, the system employs a number of different algorithms. It is vital to make use of performance indicators such as accuracy, precision, recall, and F1-score while designing the results analysis section for the purpose of identifying Sybil attacks on mobile ad hoc networks through the application of machine learning.

### *Evaluation Metrics*

In crafting the results analysis section for detecting Sybil attacks on mobile ad hoc networks using machine learning, it is crucial to include performance metrics such as accuracy, precision, recall, and F1-score. Here's a breakdown of each metric along with its formula:

*Accuracy*

By this metric, researchers can see how effectively the system can identify Sybil attacks on the network.It is determined by dividing the number of Sybil nodes that were accurately detected by the total number of nodes in the network.

$$\text{Accuracy} = \frac{\text{Number of correctly detected Sybil nodes}}{\text{Total number of Sybil nodes}} \quad (6)$$

*Precision*

Precision quantifies the accuracy of the system when it flags nodes as potential Sybil attackers. It is computed as the ratio of true positive detections to the total number of nodes identified as Sybil attackers.

$$\text{Precision} = \frac{\text{True Positive Detections}}{\text{Total number of nodes identified as Sybil attackers}} \quad (7)$$

*Recall*

Recall, or sensitivity, quantifies the system's capacity to accurately detect and identify all real Sybil attackers that exist within the network. The calculation involves dividing the number of correctly identified Sybil nodes (true positives) by the total number of Sybil nodes.

$$\text{Recall} = \frac{\text{True Positive Detections}}{\text{Total number of Sybil nodes}} \quad (8)$$

*F1-score*

The F1 score offers a well-balanced evaluation of the overall performance of the system by considering both precision and recall. It is especially beneficial when there is a requirement to achieve a compromise between incorrect positive results and incorrect negative results. The F1-score is computed by taking the harmonic mean of the precision and recall values:

$$F1 - \text{score} = 2 \times \frac{\text{Precision x Recall}}{\text{Precision+ Recal}} \quad (9)$$

The above metrics together offer a thorough assessment of the system's ability to detect Sybil assaults in mobile ad hoc networks, offering insights into its accuracy, balance between false positives and false negatives, and ability to identify actual Sybil attackers.

In Table 2 explains the ERA model showcases remarkable effectiveness in detecting Sybil attacks within mobile ad hoc networks, as demonstrated by its high performance across various metrics throughout both the training and testing phases. Throughout this training phase, ERA model exhibited an accuracy of 98.26%, indicating its proficiency in accurately identifying Sybil nodes among network entities. Moreover, the precision of 98.14% in this phase suggests that the majority of flagged nodes indeed corresponded to actual Sybil attackers. Additionally, with a recall of 98.43%, the model demonstrated a robust ability to capture the majority of genuine Sybil attackers present in the network. As the model transitioned to the testing phase, its performance further improved, with accuracy reaching 98.72% and precision slightly rising to 98.29%. These results underscore the model's consistent and reliable performance in accurately discerning Sybil attacks. However, the relatively low F1-scores of 1.32 and 1.69% during the training and testing phases, respectively, suggest a potential imbalance between precision and recall, indicating areas for further optimization to enhance the model's overall performance. Nonetheless, ERA model presents a promising approach for bolstering the security of mobile ad hoc networks against Sybil attacks, offering a potent tool for network defense and resilience.

Figure 3 presents a graphical representation of performance analysis of the ERA model. The findings collected from the assessment metrics will be presented

**Table 2:** Performance analysis of proposed ERA model

| *Performance metrics* | *Ensemble regressive arboretum model (ERA model)* | |
|---|---|---|
| | *Training phase hold-out validation (70%)* | *Testing phase hold-out validation (30%)* |
| Accuracy | 98.26% | 98.72% |
| Precision | 98.14% | 98.29% |
| Recall | 98.43% | 98.54% |
| F1-score | 1.32% | 1.69% |

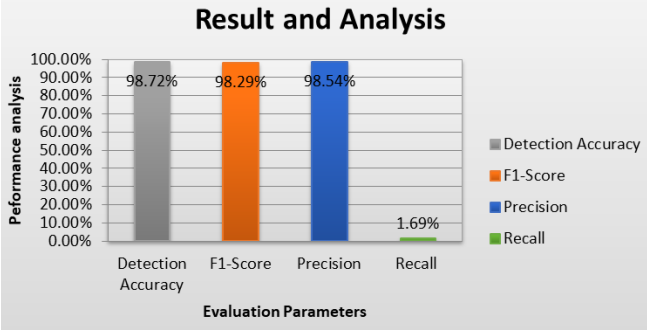


**Figure 3:** Performance analysis of proposed ERA model

and analyzed to evaluate the effectiveness of the proposed Sybil attack detection system on mobile ad hoc networks. These measures together offer vital insights to the system's reliability, efficiency, and its practicality for real-world deployment in protecting mobile ad hoc networks from Sybil assaults.

### *Comparative Analysis*

The performance of the Mulischeme, LogitBoost, RandomCommittee, Stacking, AdaBaggingand ERA Model in detecting Sybil attacks within mobile ad hoc networks was compared with metrics like accuracy, F1-score, precision, and recall.

Table 3 gives a comparative evaluation of various models overall performance in detecting Sybil attacks inside cellular ad hoc networks, as expressed by key metrics including accuracy, precision, recall, and F1 score. Among models evaluated, the ERA Model stands out with the highest accuracy score of 98.72%, indicating its exceptional capability in correctly classifying instances. Furthermore, the ERA model demonstrates incredible precision and recall percentages of 98.29 and 98.54%, respectively, highlighting its potential to correctlypick out Sybil attackers while

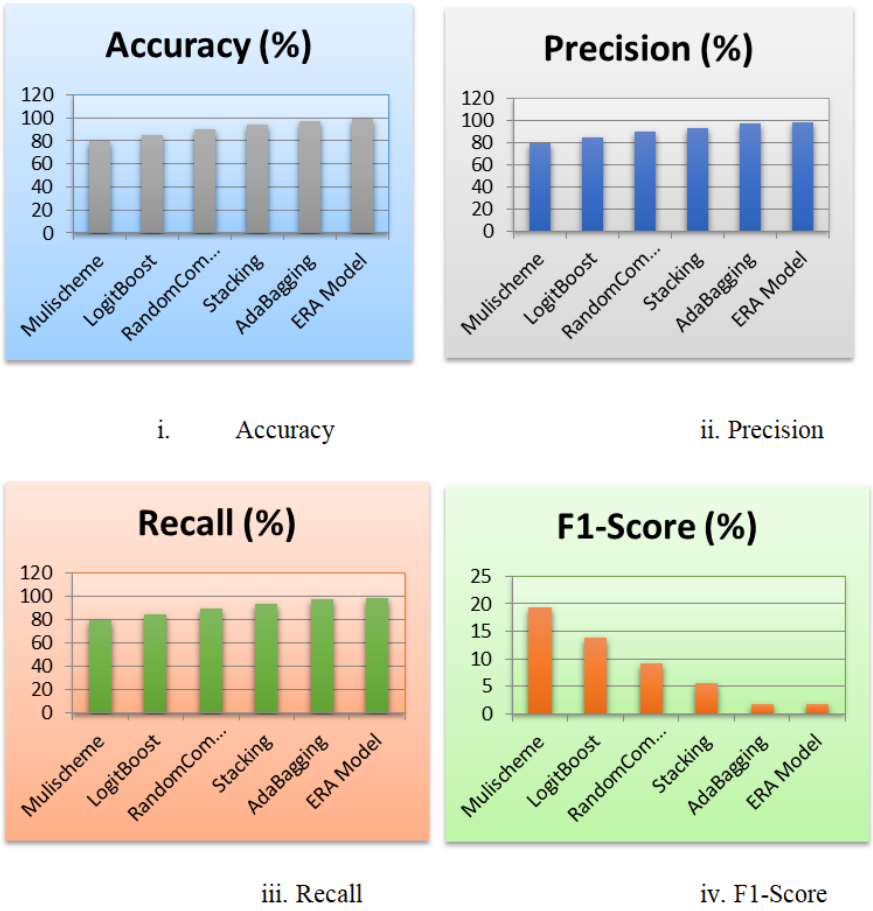


**Figure 4:** Comparative analysis of various machine learning algorithm

**Table 3:** Comparative analysis of various algorithm

| | *Accuracy (%)* | *Precision (%)* | *Recall (%)* | *F1-Score (%)* |
|---|---|---|---|---|
| Mulischeme | 79.68 | 79.59 | 79.32 | 19.30 |
| LogitBoost | 84.53 | 84.46 | 84.40 | 13.93 |
| Random Committee | 89.75 | 89.69 | 89.58 | 9.23 |
| Stacking | 93.87 | 93.40 | 93.47 | 5.67 |
| AdaBagging | 97.07 | 97.03 | 97.00 | 1.80 |
| ERA Model | 98.72 | 98.29 | 98.54 | 1.69 |

minimizingfalse positives and negatives. Despite its high accuracy and precision, the ERA model reveals a highly decreased F1-score of 1.69%, suggesting potential room for development in attaining a higher stability among precision and recall. Compared with Mulischeme, LogitBoost, RandomCommitte and Stacking models display varying levels of performance across the metrics, with differences in their ability to accurately detect Sybil attacks. The findings underscore the effectiveness of the ERA Model in enhancing the security of mobile ad hoc networks against Sybil attacks, while also indicating avenues for further optimization to achieve a more balanced performance across all evaluation metrics.

Figure 4 compares the overall performance of various models in detecting Sybil attacks in mobile ad hoc networks with the use of metrics like accuracy, precision, recall, and F1-score. ERA model outperforms others with the highest accuracy, precision, and recall scores, indicating its effectiveness in identifying Sybil attackers accurately. However, its F1-score is relatively lower, suggesting a trade-off between precision and recall. Other models show varying performance levels, highlighting the importance of selecting the most suitable model for effective Sybil attack detection.

## Conclusion

In ad hoc networks, mobility is a common enemy of security services. Mobility enhances security and helps detect Sybil assaults, as shown in this research. The detection method uses the ERA model incorrectly and incorrectly labels groups of nodes that move in tandem as Sybil attackers. This method is used to collect network attributes such communication patterns, node behaviours, and traffic characteristics. It is able to distinguish between legitimate node behaviour and that of a Sybil attacker by including these traits into a proposed ERA model. Accuracy, F1-score, precision, and recall were some of the metrics which may be utilized to evaluate the ERA model. The outcomes of the ERA model will be contrasted with those of alternative methods including Mulischeme, LogitBoost, RandomCommittee, Stacking, and AdaBagging. on order to identify Sybil attacks on mobile ad hoc networks, ERA model methods will be utilized. The findings provide an in-depth evaluation of how well the ERA Model Algorithm detects Sybil assaults in MANETs using machine learning. Accuracy, F1-score, precision, and recall were some of the metrics used to estimate the algorithm's performance. With regard to Sybil assaults, the ERA model algorithm is capable of achieving a detection accuracy of 98.72%, a precision of 98.29%, and a recall rate of 98.54%. The approach is able to successfully identify Sybil assaults on MANETs, as demonstrated by its F1-Score of 1.69% or higher.

## Acknowledgement

We also wish to thank the Editor, Associate Editor and reviewers for their invaluable comments and suggestions, which significantly enriched the manuscript.

## References

Abu Jahid., Mohammed H., Alsharif.,& Trevor J. Hall. (2022). A contemporary survey on free space optical communication: Potentials, technical challenges, recent advances and research direction. Journal of Network and Computer Applications. 200.

Akarsh K. Nair., Ebin Deni Raj., &Jayakrushna Sahoo. (2023). A robust analysis of adversarial attacks on federated learning environments. Computer Standards & Interfaces. 86.

Amol Vasudeva.,& Manu Sood. (2022). On the vulnerability of the mobile ad hoc network to transmission power-controlled Sybil attack: Adopting the mobility-based clustering. Journal of King Saud University - Computer and Information Sciences. 34 (9), 7025-7044.

Annu Govind., Kuldeep Jayaswal., Vijay Kumar Tayal., &Prakash Kumar. (2024). Simulation and real time implementation of shunt active power filter for power quality enhancement using adaptive neural network topology. Electric Power Systems Research. 228.

Bhupender Kumar.,& Bubu Bhuyan. (2020). Game Theoretical Defense Mechanism Against Reputation Based Sybil Attacks. Procedia Computer Science. 167, 2465-2477.

Brennan Huber., Farah Kandah.,& Anthony Skjellum.BEAST: Behavior as a Service for Trust management in IoT devices. Future Generation Computer Systems.144, 165-178.

Christopher Morales Gonzalez., Matthew Harper., Michael Cash., Lan Luo., Zhen Ling., Qun Z. Sun., &Xinwen Fu. (2024). On Building Automation System Security. High-Confidence Computing. 4 (3).

Hong Li., Lixia Bai., Weifeng Gao., Jin Xie., Lingling Huang. (2024). Many-objective coevolutionary learning algorithm with extreme learning machine auto-encoder for ensemble classifier of feedforward neural networks. Expert Systems with Applications.246.

Iain Baird., Isam Wadhaj., Baraq Ghaleb., & Craig Thomson. (2024). Impact Analysis of Security Attacks on Mobile Ad Hoc Networks (MANETs). Electronics. 13 (16).

Jean-Paul Yaacoub., Hassan Noura., Ola Salman., &Ali Chehab. Security analysis of drones systems: Attacks, limitations, and recommendations. Internet of Things. 11.

Naveen Kumar., &Ankit Chaudhary. (2024). Surveying Cybersecurity Vulnerabilities and Countermeasures for Enhancing UAV Security. Computer Networks. 252.

SabirAli Changazi., Asim Dilawar Bakhshi., Muhammad Yousaf.,

Syed Muhammad Mohsin., Syed Muhammad Abrar Akber., Mohammed Abazeed., &Mohammed Ali. (2024). Optimization of network topology robustness in IoTs: A systematic review. Computer Networks. 250.

S. Sarika., A. Pravin., A. Vijayakumar., &K. Selvamani., (2016). Security Issues in Mobile Ad Hoc Networks.Procedia Computer Science. 92,329-335.

Shaik Shafi., S Mounika., &S Velliangiri. (2023). Machine Learning and Trust Based AODV Routing Protocol to Mitigate Flooding and Blackhole Attacks in MANET. Procedia Computer Science. 218, 2309-2318.

S. Harihara Gopalan., V. Vignesh., D. Udaya Suriya Rajkumar., A.K. Velmurugan., D. Deepa., &R. Dhanapal. (2024). Fuzzified swarm intelligence framework using FPSOR algorithm for high-speed MANET- Internet of Things (IoT). Measurement: Sensors. 31.