Black hole Attack Detection Using Machine Learning Algorithms in MANET – Performance Comparision

Ms. Katakam Tejaswini1, Mrs. Yannam Adilakshmi2

1Katakam Tejaswini Gudlavalleru Engineering College
2Mrs. Yannam Adilakshmi Associate Professor, Dept. of Computer Science and Engineering, Gudlavalleru Engineering College, Gudlavalleru, Andhra Pradesh, India

Abstract - MANET forms without any stand support devices. Nodes can be established in the network very easily without any prerequisite. Due to this flexibility many security threats can occur in the routing. So, in order to overcome from this, the performance of IDS should be improvised. In the present project a comparison study has been carried out on machine learning algorithms in terms of accuracy and detection rate for IDS improvisation.

Key Words: MANET, Support Vector Machine, Decision tree Classifier, Random Forest classifier, Logistic Regression

1. INTRODUCTION

IDS is to detect the attack before the attacker introduces any harm to the network, it take care of three major functions like monitoring, detecting and generating alarms. The black hole attack is considered as the one of the most affected kind on MANETs. In order to protect the network from blackhole attack, the use of an Intrusion Detection System (IDS) has a major importance in the MANET protection. With the help of the machine learning algorithms here it improve the anomaly-based IDS in MANET’s here the following algorithms are support vector machine, decision tree classifier, random forest classifier, logistic regression algorithm by considering above four algorithms which one is best for detecting anomaly-based IDS in MANETs.

1.1 BLACK HOLE ATTACK USING AODV ROUTING PROTOCOL

Fig(1):Black hole Attack

Black hole attack is one of the major attacks in MANETs. It consists of Source node, Destination node and Neighbouring nodes. The source node sends a RREQ (Route Request Packet) to its neighbouring nodes to search for the route destination. However, black hole node sends a fake route reply to the source node which will degrade the performance of the network. In order to prevent this, the performance of the IDS should be increased with machine learning algorithms.

2. Related Work:

Sujithra L et. al In this paper it improves the conversation of energy in heterogeneous network and also reducing the active time of IDS running in the nodes so in order to achieve this probabilistic approach is implemented, here optimal probabilistic of node to be set. By considering this decreasing the active time of IDS in each node and also conserve the energy of the node that increases the network lifetime significantly.

Sankaranarayanan S et. al Proposed the RSA algorithm in intrusion detection system in MANET it successfully identifies the malicious node Results shows that secure IDS method gives better packet deliver ratio in presence of intruders.

Indira N Proposed Anomaly based intrusion detection technique using the SOM classification method provides higher detection rate than other anomaly detection method. As anomaly-based intrusion detection techniques are based on statistical data they can result in false positive identification of normal pattern as an attack. This false identification of benign behavior as abnormal can result in isolation of non-malicious node as malicious, thus may result in partitioning of the network.

Hangmei Deng proposed two detection models that is distributed hierarchical and completely distributed. IDS is implement- ed in both models and focus on the network layer for detect the attack in the network that hierarchical distribution gets the good detection rate than the complete distributed approach

Houda Moudni proposed-Adaptive neuro fuzzy inference system (ANFIS) algorithm and particle swarm optimization
(PSO) are used for detection and prevention of black hole attack in MANETs. The performance evaluation (ANFIS+PSO) is evaluated by Detection rate and False rate. As a result; less no. of node connections leads to high detection rate and low false rate.

Zalte S.S. et. al proposed- black hole and gray hole attack detection in AODV protocol by IDS. By comparing the normal AODV, attacker AODV and new AODV. As a result; new AODV improves the throughput and packet delivery ratio.

3. PROPOSED METHODOLOGY

Nodes in the MANET share the wireless medium and the topology of the network changes erratically and dynamically. Research in a MANET gets tremendous attention because of its eminent characteristics like instant infrastructure, cheap and easy deployment in hostile terrain where geographical conditions are not suitable for an earthquake, battlefield. MANET can build anytime and anywhere. Since the nodes are roaming, the network topology varies rapidly. The remarkable advantages of MANETs such as multi hop, infrastructure less transmission etc., makes it as a best medium to networks. Though MANETs have surplus things, they have some security issues that will cause severe damages and loss in network. Random linking of mobile nodes leads to add malicious nodes in the network accidentally. To suspect and detect the malicious activity in the network, Intrusion Detection System (IDS) are implemented to analyze the behaviour of the neighbourhood nodes. To improving the anomaly based intrusion detection system in MANETs comparing the four algorithms i.e, support vector machine, decision tree classifier, Random forest classifier and logistic regression.

Fig(2):Flow chart for Proposed Methodology

A 3 step method is followed for the analysis

1. Simulation of Network
2. Data collection
3. Model Training and Data Testing

1. Simulation of Network

In the present project, Mobile ad-hoc network(MANETs) is simulated in NS2 with 25 number of nodes as shown in fig(8).
3. MODEL TRAINING AND DATA TESTING

In the present project four algorithms were used to train and test the data
a. Support vector machine (SVM)
b. Random forest classifier
c. Decision tree classifier
d. Logistic regression

a. Support vector machine

The primary aim of support vector machine (SVM) is to separate the normal and abnormal nodes by choosing the best estimated hyperplane and it is selected such away that the distance from the hyperplane to the nearest node on each side is maximized.

Fig(6). Support Vector Machine Classification

In the present Project, the data set obtained from NS2 is fed into SVM algorithm. The black hole attack is detected in terms of accuracy and confusion matrix. The output is shown in Fig (7).

Fig(7). Accuracy Score and Confusion Matrix of SVM Classifier

From SVM Algorithm it is observed that an accuracy of 82.35 and the confusion matrix showing less false positive rate.
b. Random forest classifier

Random forest is an ensemble learning techniques used for predictive modelling and machine learning techniques. It classifies, create a set of decision tree and randomly selected subset of training set. It aggregates the votes from different decision tree to decide the final class of the test object.

Fig(8). Random Forest Classification

In the present Project, the data set obtained from NS2 is fed into Random forest classifier algorithm the black hole attack is detected in terms of accuracy and confusion matrix.

The output of Random forest classifier is shown in Fig(9).

![Random Forest Simplified](image)

\[ IG(D_p,f) = I(D_p) - \frac{N_{left}}{N} I(D_{left}) - \frac{N_{right}}{N} I(D_{right}) \]

Where

- \( f \) - Feature to perform the split
- \( D_p \) – Dataset of the parent node
- \( D_{left} \) – Dataset of the left child node
- \( D_{right} \) – Dataset of the right child node
- \( I \) – impurity criterion (Entropy)
- \( N \) – total number of samples
- \( N_{left} \) – Number of samples at left child node
- \( N_{right} \) – Number of samples at right child node

In the present project, the data set obtained from NS2 is fed into Decision tree classifier algorithm the black hole attack is detected in terms of accuracy and confusion matrix.

The output of Decision tree classifier is shown in Fig(10).

![Accuracy score and Confusion Matrix of Random Forest Classifier](image)

Fig(9). Accuracy score and Confusion Matrix of Random Forest Classifier

From Random forest classifier algorithm, it is observed that an accuracy of 88.23 and the confusion matrix showing less false positive rate

c. Decision Tree Classifier

Decision tree is a white box type of machine learning algorithms. It share internal decision making, logical which is not available in the black box type of algorithms. While building the decision tree the ID3 algorithm select the best feature at each step and it uses information gain to find best feature. Information gain is estimated using below formulae.

\[ IG(D_p,f) = I(D_p) - \frac{N_{left}}{N} I(D_{left}) - \frac{N_{right}}{N} I(D_{right}) \]

From Decision Tree classifier algorithm, it is observed that an accuracy of 82.35 and the confusion matrix showing more false positive rate.

d. Logistic Regression

Logistic regression is one of the techniques for data classification. Categoric dependent variables can be solved by Logistic regression using some Mathematical models. In the present project, the data set obtained from NS2 is fed into Logistic Regression classifier algorithm the black hole attack is detected in terms of accuracy and confusion matrix.

The output of Logistic Regression is shown in Fig(11).
Fig(11). Accuracy score and Confusion Matrix of Logistic Regression Classifier

From Logistic Regression classifier algorithm, it is observed that an accuracy of 88.23 and the confusion matrix showing less false positive rate.

By observing the above algorithms SVM and decision tree classifier obtained the same accuracy i.e, 82.35, but SVM shows more detection rate by compare with randomforest and logistic regression getting same accuracy by comparing four algorithms random forest getting more detection rate.

By observing the above algorithms SVM and decision tree classifier obtained the same accuracy i.e, 82.35, but SVM shows more detection rate by compare with randomforest and logistic regression getting same accuracy by comparing four algorithms random forest getting more detection rate.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Type of algorithm</th>
<th>Accuracy (%)</th>
<th>Detection rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Support Machine Vector</td>
<td>82.35</td>
<td>Less</td>
</tr>
<tr>
<td>2</td>
<td>Random forest classifier</td>
<td>88.23</td>
<td>More</td>
</tr>
<tr>
<td>3</td>
<td>Decision tree classifier</td>
<td>82.35</td>
<td>Less</td>
</tr>
<tr>
<td>4</td>
<td>Logistic regression</td>
<td>88.23</td>
<td>Less</td>
</tr>
</tbody>
</table>

4 CONCLUSION

From above results it is concluded that random forest classifier gives the higher accuracy and better detection rate compared to the SVM algorithm, decision tree classifier and Logistic regression. Therefore, the performance of IDS can be improvised by random forest algorithm.

REFERENCES

1. International Research Journal of R &D IN ENGINEERING SCIENCE AND MANAGEMENT, Study of MANET-Characteristics, Challenges, Application, Routing Protocol and Security Attacks Sandeep Kumar, Col. (Dr) Suresh Kumar


5. Springers Effective Classification and Handling of Incoming Data Packets in Mobile Ad Hoc Networks (MANETs) Using Random Forest Ensemble Technique (RF/ET) Anand Nayar, Bandana Mahapatra.


