Project Closure Report For

Preventing Black Hole

Attacks In Manets Using Dynamically Generated Audit Data

**Audience:**

Project sponsors, project mentors, unit coordinators and all others in the university and ITC community interested in this project.

**Purpose:** The purpose of this document is to clarify the scope, methodology and objectives of this capstone project. The main stakeholders should evaluate the project proposal plan and provide feedback and suggestions that could improve any part of the plan.

|  |  |  |
| --- | --- | --- |
| **Authors** | **Version** | **Date** |
| Thi Doan,  Alan Gaugler,  Tristan Hore,  Johnson Sangah | 1.0 | 03 May 2023 |

**Table of Contents**

**Dataset Creation**

The main behavioural characteristics of black hole nodes in AODV networks that were identified in the literature review were some of the key features to be incorporated into the dataset to be used in the machine learning process in the detection of black hole nodes.

The data is created from running simulations, so collecting a sufficient amount of data for training and testing the machine learning models is not an issue. With proper simulation files available It is recommended to run several simulations with different configurations and merge them together to create the training set (ensure there are no duplicate Node ID names) and this can also be done for the test set. Once the model has been tuned, this can be applied to an individual network simulation and used to detect the black hole nodes within it.

The output of the trace files consisted of messaging files from each node. These were in .pcap file format and were loaded into Wireshark where they could be observed. Wireshark is a free and open-source network protocol analyzer which is available from here:

<https://www.wireshark.org>

A screenshot of a computer

Description automatically generated

Figure 1 – Example display of AODV messaging in Wireshark

A screenshot of a computer

Description automatically generated

Figure 2 –AODV message details in Wireshark

From Wireshark, these files were exported as .json files where they can be loaded into a text editor or an IDE to be viewed an analyzed. The tool used in this project was Visual Studio. From here, the key message components of the AODV messages could be located.

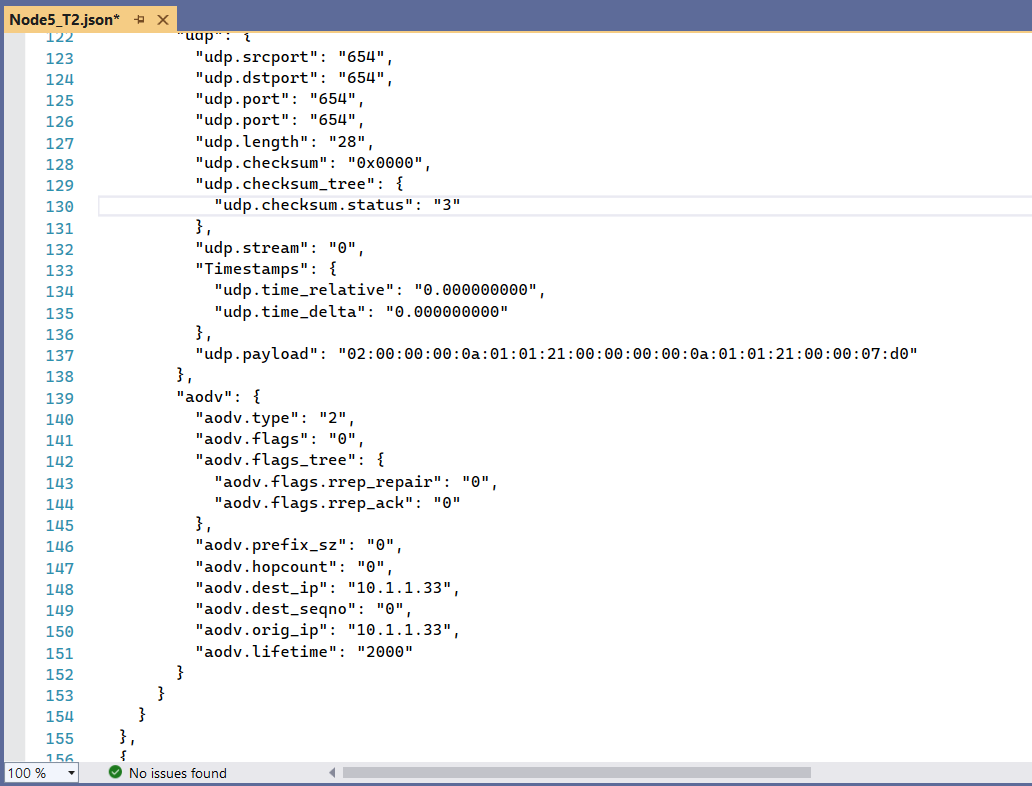


Figure 3 –Example of AODV messaging in JSON format view in Visual Studio

From this, the next stage was to extract the relevant features of the AODV messaging being sent and received from each node to and from its tier 1 neighbours and convert them into a dataset that could be used for the machine learning process.

Our team developed a Python script in Jupyter Notebooks to do this. This script is included in the project deliverables with the name of “AODV\_to\_Dataset.ipynb”.

The list of black hole nodes was manually entered into the script. When run, the target variable “Black\_Hole\_Node” was modified to True if the neighbour node was in the list. The rows in the dataset where the black hole node was the subject node were removed from the dataset because the purpose of this process is for a normal node to learn how to detect the behaviour of a black hole node.

The process of converting the individual node trace files into a useable dataset was basically split into three stages.

**Stage 1:** Read in the relevant AODV information from each node (each JSON file) and store them in separate data frames. If it is desired in the future to use another source of AODV messaging as the input into the dataset creation process, only stage 1 needs to be modified to get the relevant message features from the other source. Stages 2 and 3 can be left unchanged.

Looking at the example Stage\_2.csv files, each row refers to each extracted message marked by a frame time. It contains information of the Subject Node, Neighbour Node, the AODV Message type and many other relevant features. Columns B to O are the data frames for each node produced by Stage 1.

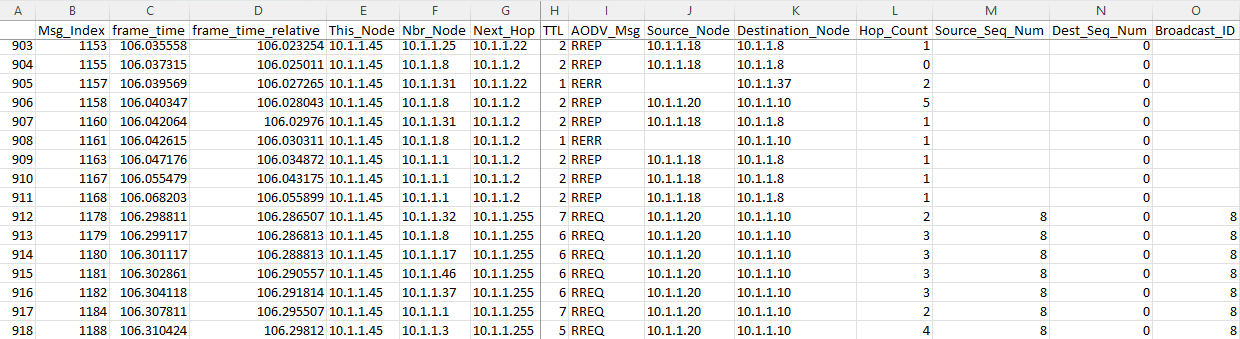


Figure 4 – Example Output of a Data Frame after Stage 1

**Stage 2:** Some processing is made for each AODV messages, creating new features in the data frame which are columns P to AA in the Stage\_2.csv files. Many of these are Boolean values of if the message is a certain type or is addressed to or originates from the neighbour node.

If the message is a RREP message responding to a RREQ message, the RREQ message index is found in Column U and the response time between the RREQ and the RREP is noted in columns V & W. If it is an RREP, the destination sequence number increment is also determined in Column Y.

As currently there are some errors with the NS-3 outputs in emulating the AODV protocol, some rows were deleted from the data frames including for now RREP-ACK messages and RREP messages from which no corresponding RREQ message could be found.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 5 – Example Output of Additional Features of a Data Frame added after Stage 2

**Stage 3:** This stage will convert the data frames from Stage 2 into datasets for each Subject Node and then merge them all into one combined dataset. Every row represents features between the subject node and each of its 1st tier neighbouring nodes. Each subject to neighbour node relation is only ever one row. The features are mainly counters, percentages as well as Boolean values. All the features are described in the excel file “Dataset Features.xls” which is attached with the deliverables.A screenshot of a computer

Description automatically generated with medium confidence

Figure 6 – Example Output of the Dataset produced after Stage 3

A screenshot of a computer

Description automatically generated with medium confidence

Figure 7 – Example Output of More Features from the Dataset produced after Stage 3

**Machine Learning from Black Hole Node Detection**

A machine learning script was developed to use the datasets created by the conversion script as input. The ML script would be used to develop models to detect which nodes within the AODV network were malicious black hole nodes.

An example training and test csv files have been attached in the deliverable package to demonstrate the machine learning script working. Again, it must be highlighted that many AODV characteristics have not been accurately implemented into the datasets. Black hole nodes were not implemented accurately either and so the datasets have been modified for the rows where the target variable “Black\_Hole\_Node” is set to True. Many of the final features’ values have been altered so that the black holes behaviour is considered “dumb” and its characteristics are extremely obvious. This would be the behaviour of a black hole node that is not expecting the network nodes to have any intelligence in detecting and mitigating malicious nodes as most black hole nodes are.

For example, for black hole nodes, RREQs\_From\_Nbr is always set to 0 and its corresponding flag feature Nbr\_Never\_Sends\_RREQ is always set to True. This is the behaviour that is expected from a black hole as it would always be responding to an RREQ with an RREP claiming to have the shortest route to the destination, so it would never broadcast the RREQ. Likewise, RREP\_Resp\_Pct would always be 100% because the BHN always response to any RREQ it receives. It would also respond with a short hop count to make it more likely to be the chosen path, so Hop\_Cnt\_Over\_1\_Cnt would always be 0. It is likely that the BHN would respond with a high destination sequence number increment to give its path the highest priority and so the flag High\_Dest\_Seq\_Num\_Inc\_Pct would likely always be 100% because the BHN always responds to the RREQ with a high sequence number in the RREP.

To summarise which features have been manually altered for black hole nodes, they are Columns I to V, Y, Z & AB.

Only the train and test datasets have been modified. The attached examples of Stage\_2 files have not been modified but they do not exhibit any black hole node behavioural characteristics.

Most of the process in the Machine learning script is explained in the actual Jupyter Notebook. So I won’t repeat it in much detail here.

The training and test datasets are loaded, and there is some cleaning of the datasets. The Boolean values are converted to integers, 1 or 0 and any rows with missing values (mainly due to neighbours not carrying any traffic) are removed. It is not expected that any black hole neighbours will have rows with missing values due to the nature of their behaviour.

The features with the highest correlation to the Black\_Hole\_Node target variable are selected for the train and test sets that will be used in the modelling. After evaluation, a final set of 10 features was selected to be used in the modelling.

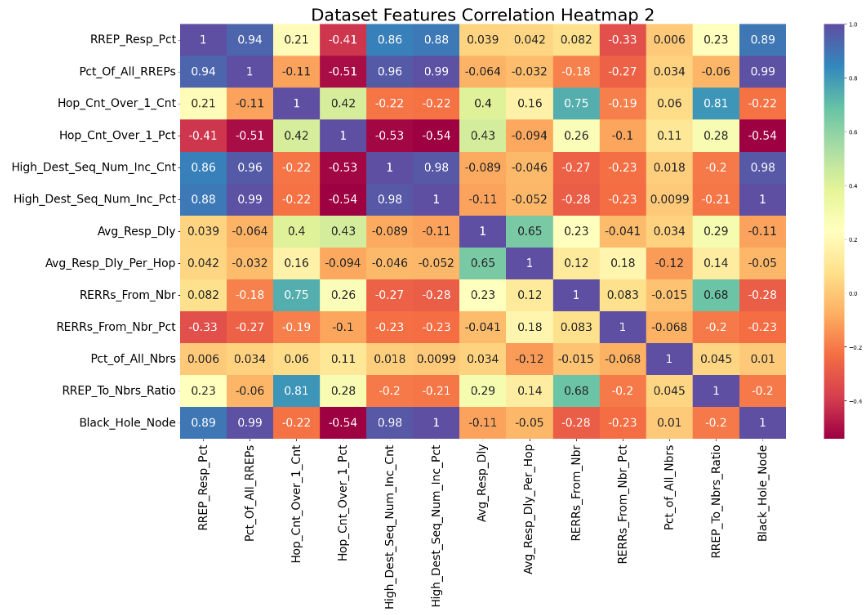


Figure 8 – Example Output of the correlation heatmap demonstrating the relationship between the features and the target variable (bottom row)

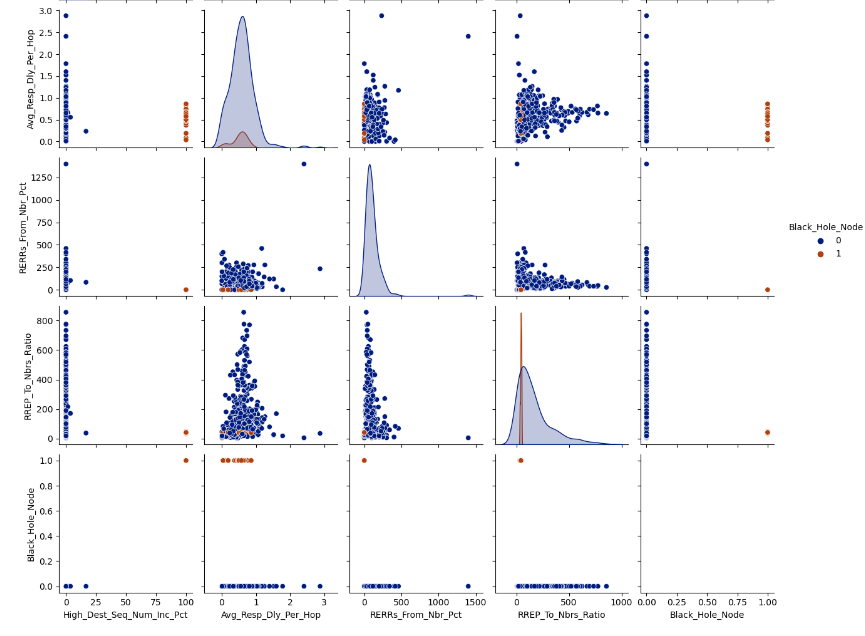


Figure 9 – Example Output of Pair Plots showing the relationships between variables.

Black hole node is brown, Normal node is Blue.

Two models are used to detect the black hole nodes. Model 1 is a random forest classifier and Model 2 is a Support Vector Machine (SVM) classifier. A grid search is carried out on both of these models to determine the optimum hyperparameter settings. Standardization is also applied to the data for the SVM so that all features carry an equal weight in the decision process. The models are trained and are then applied to the test set to determine their accuracy. In both models, their accuracy is 100%. This is confirmed by displaying the node IDs of the predicted rows which match those of the simulated black hole nodes.

A picture containing text, screenshot

Description automatically generated

Figure 10 – Example of the Confusion Matrix from the SVM Model Classifying Black Hole Nodes.

As mentioned, much more work needs to be done once the NS-3 simulator is simulating AODV networks accurately. The future work section in the final report mentions what work should be carried out if anyone continues to work on this project.

This script is included in the project deliverables with the name of “ML\_for\_Malicious\_Node\_Detection.ipynb”.

**Outstanding Issues**

**Current observed issues in the network simulation output accurately simulating AODV network and black hole node behaviour:**

After running the NS-3 simulations. A detailed examination was carried out to determine if the output conforms to AODV protocol behaviour. Unfortunately, numerous issues were found. The time frame of the project did not permit us enough time to correct these issues.

The issues discovered with the AODV simulations are as follows:

1. The Destination Sequence Number (DSN) in the RREP message is not incremented or is not greater than the Destination Sequence Number in the corresponding RREQ. This should be incremented to either the be set to destination sequence number in the RREQ + 1 or to the current sequence number in the destination node if it is already higher. This will be complex to implement as a routing table must be kept for each node which increments every time it sends a RREP.

2. The Destination Sequence Number was mostly set to 0. 0 is a special value that means the DSN is not known. This is expected at the beginning of the network simulation but the DSNs should increment afterwards.

3. The RREP does not send the Origination Sequence Number. A field needs to be created in the RREP message for this. This should be easy to implement.

4. RREPs appear to be broadcast rather than unicast. Nodes are receiving RREPs without having had sent an RREQ with the same source and destination nodes.

5. Clarification needs to be made about the RREQs and RREPs from the current node. Are the messages we are seeing just what is received observed from the current node? Or does it include sent messages too?

6. Hellos should end in 255.255.255 not just .255. This is not a big issue.

7. Neighbour node is never the destination. Nor is the subject node. According to destination of the RREQ

8. All nodes have high neighbour counts 31 – 49. This is not an error, but it is a high density of nodes. Various network configurations should be made to test for black hole nodes in several network scenarios.

9. Hello Count from all neighbours should be roughly even. However, more Hellos will be generated from those nodes carrying less traffic. A HELLO message should be generated after every 1 second of inactivity (no other AODV messages sent). This latter part appears to be functioning correctly.

10. In the RREQ messages, the neighbour is never the destination. It is however observed as a destination in the RREP messages which is inconsistent.

11. Often the neighbour sees an RREQ sent from the Source Node, however the source node’s pcap file does not send it. This needs to be investigated as a high priority.

12. Similar to 11, there are inconsistencies in the messages sent in different node files. i.e. the same RREQ message consistent across nodes. This can be determined from its timing and Broadcast ID.

13. Black hole nodes have not been implemented yet.

It is highly likely that there are also more issues with the simulation accuracy that have not yet been identified. As the simulations are upgraded and more work continues, further testing would be carried out.

**Future Work**

There is so much more interesting work that could be done to improve the scope of this project and its effectiveness at nullifying black hole nodes in real life scenarios. This is beyond the scope of a one semester project. Many important steps include:

* Improve the NS-3 simulations so that the simulations are behaving as expected according to the AODV protocol.
* Improve the black hole node simulations. For the 1st stage, the black hole nodes should be “dumb” where their behaviour is obvious, i.e. the act exactly as is described above.
* More features could be extracted from the trace files and added to the dataset that go beyond the AODV protocol which include amount of traffic data (kilobytes, megabytes) received from 1st tier neighbour node.
* Once these are working accurately, create several datasets with various network configurations, including:
  + Network size
  + Number of nodes,
  + simulation run time
  + mobility of nodes
  + activity of nodes
  + number of black hole nodes.
* Combine all of these datasets of the various network configurations into one large training dataset. This larger set will used to train the models to various configurations of networks, so that it may more accurately be able to identify black hole nodes in certain network conditions. Blackhole nodes should not be the subject node in the dataset, as it is desired to train normal nodes to detect black hole nodes.
* The test dataset should also be a large and consist of a wide variety of network conditions.
* After the models have been trained, a new network simulation can easily be configured and run and used to detect the model’s accuracy at detecting the black hole nodes.
* After the final models have been trained, the aim is to deploy these algorithms into the normal nodes within the network. These normal nodes will monitor their 1st tier neighbours behaviour for a certain period of time and should be able to detect if a 1st tier neighbour is a black hole node. If so, the subject node should be trained to find alternative routes to the desired destination, effectively avoiding and isolating the black hole node.
* Further work from the previous step could be to tune how much activity data is required for the subject node to accurately detect the black hole node. The idea is to minimise the amount of time to detect the black hole node so that disruption to network traffic is minimised.
* Total network traffic could be monitored for both:
  + 1. normal nodes
  + 2. All nodes made “smart” by incorporating the ML algorithms to detect and bypass black hole nodes

The difference in total network traffic degradation could be monitored and compared from the two scenarios.

* Progressively make the black hole nodes “smarter” so that their behaviour is not consistently so obvious and repeat the network simulations to test the models’ accuracy of detection.