

Configuration Manual

MSc Research Project
MSc In Cybersecurity
Practicum

Burhanuddin Shabbar
Student ID: 23142502

School of Computing
National College of Ireland

Supervisor: Khadija Hafeez

National College of Ireland
MSc Project Submission Sheet
School of Computing



Student Name: Burhanuddin Shabbar.....
Student ID: 23142502.....
Programme: MSc In Cybersecurity..... **Year:** 2024.....
Module: Practicum.....
Lecturer: Khadija Hafeez.....
Submission Due Date: 12/08/2024.....
Project Title: Detection of blackhole and DDoS attack in 5g VANET using multiple machine learning algorithms.
Word Count: 812 words **Page Count:** 15 pages

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Burhanuddin Shabbar.....
Date: 12/08/2024.....

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Burhanuddin Shabbar
23142502

1 INTRODUCTION:

The configuration manual explains the requirement to create the environment, implementation steps, necessary hardware, software and code snippets used for completion of research work. The main purpose of this manual is to demonstrate step by step coding procedure taken for this project. It will help to replicate and verify the results in future.

2 SOFTWARE AND HARDWARE REQUIREMENTS

2.1 Hardware requirements:

Following are the hardware requirements.

- CPU: Intel core i5 or i7 processor.
- 16 GB RAM
- Storage: 512 SSD

2.2 Software requirements:

Following are the software requirements.

- Windows 11
- Ubuntu 20.04
- OMNET++ version 5.6.2
- Venis version 5.0
- SUMO version 1.0.1
- Python 3.10

2.3 Python Libraries:

Following are the python libraries that are required to run the AI part of the project.

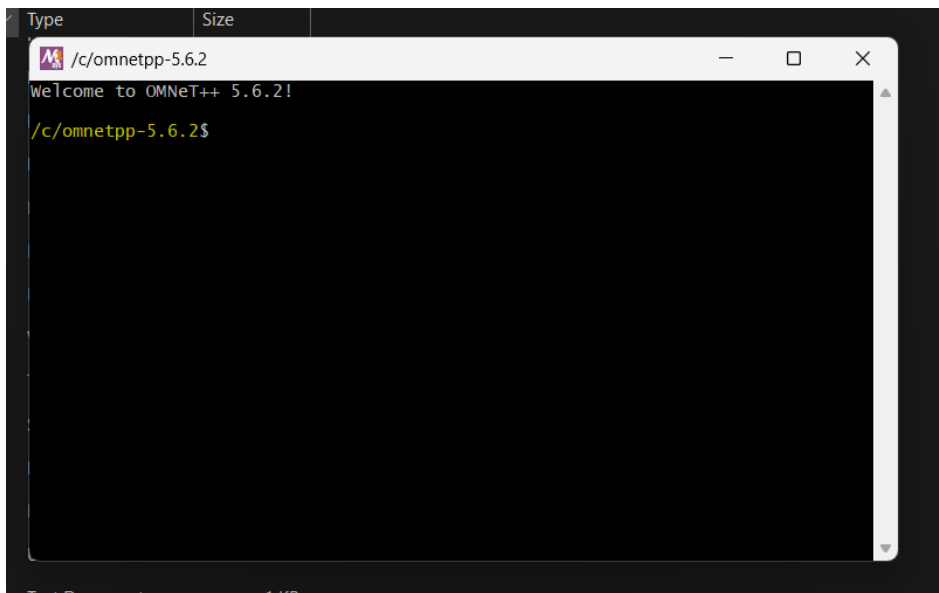
- Numpy
- Pandas
- Scikit-learn
- Matplotlib
- TensorFlow
- XGBOOST
- Datetime

3 SETTING UP THE SIMULATION

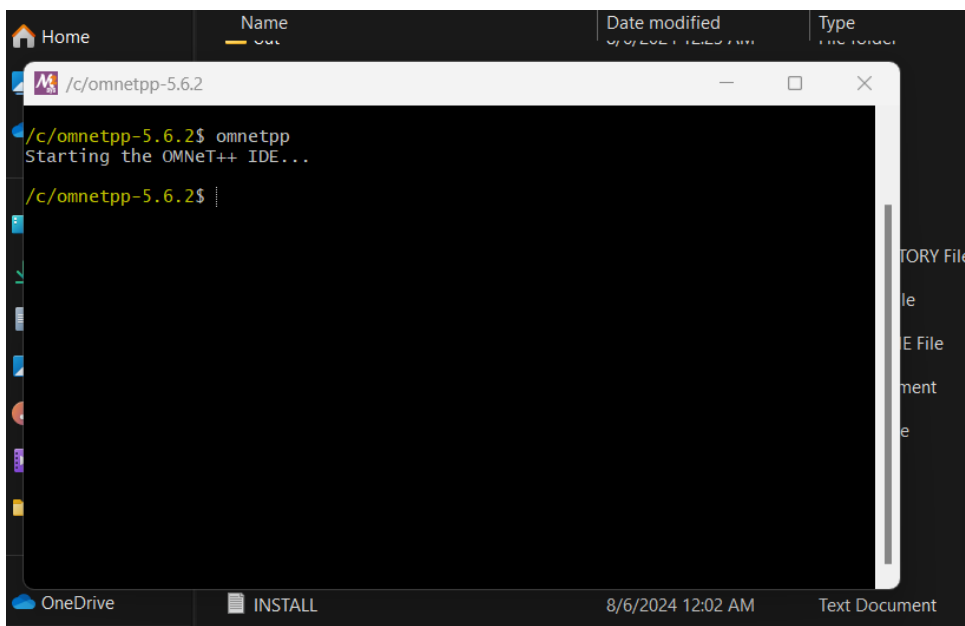
- First go to the folder where you have installed the simulator. Then find mingwenv file and run as administrator.

MIGRATION	8/6/2024 12:02 AM	Text Document	2 KB
mingwenv	8/6/2024 12:02 AM	Windows Comma...	1 KB
README	8/6/2024 12:02 AM	Text Document	4 KB

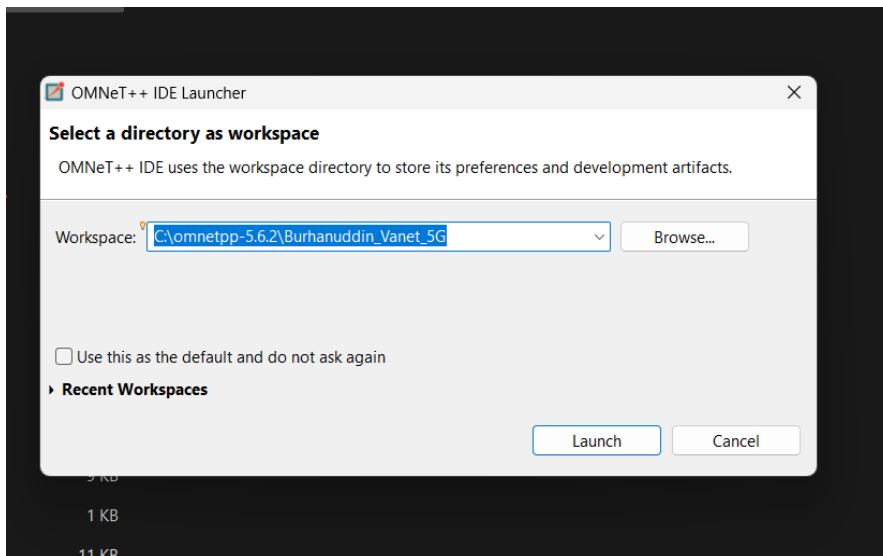
- A popup will open like this



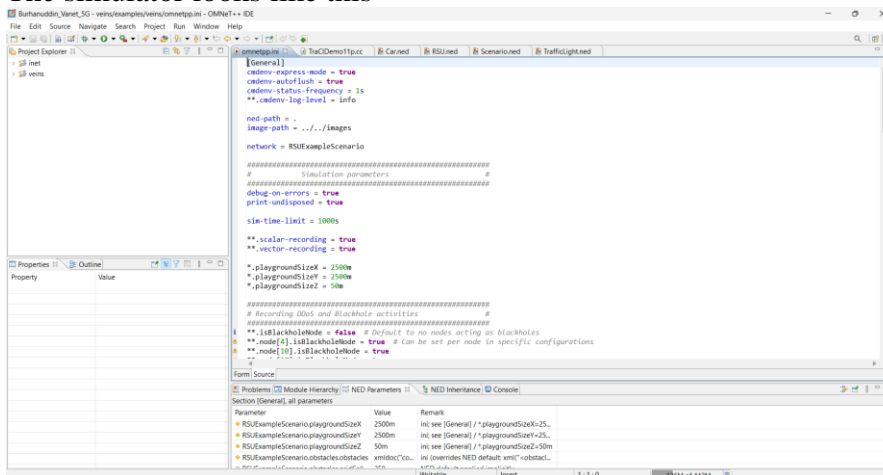
- Here type omnetpp and hit enter



- This popup will show up and now select the workspace you want to open. In our case it is the Burhanuddin_Vanet_5G and then hit launch

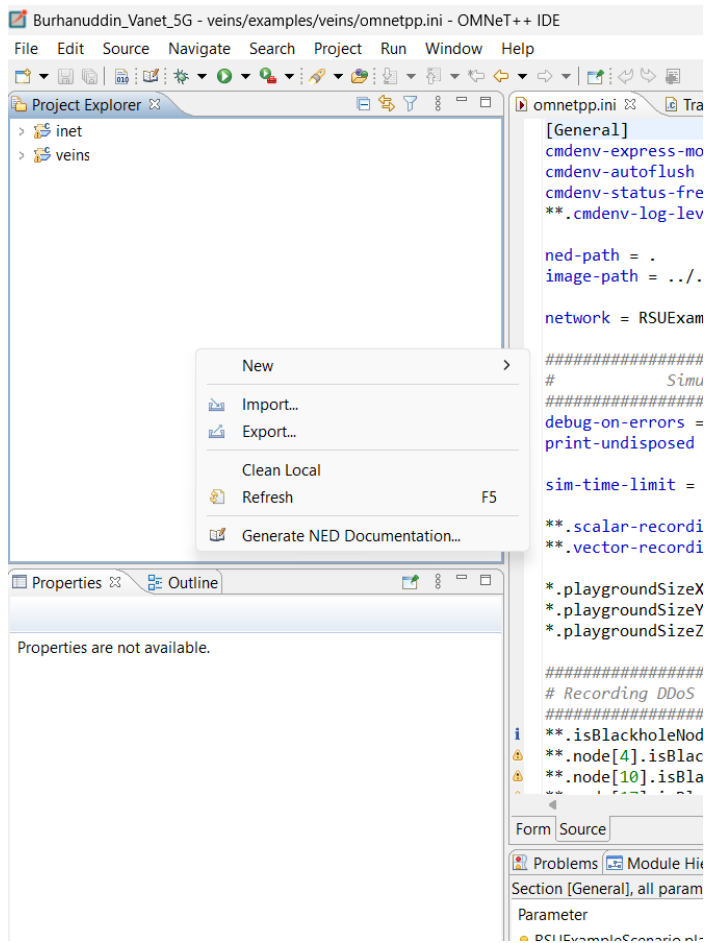


- The simulator looks like this

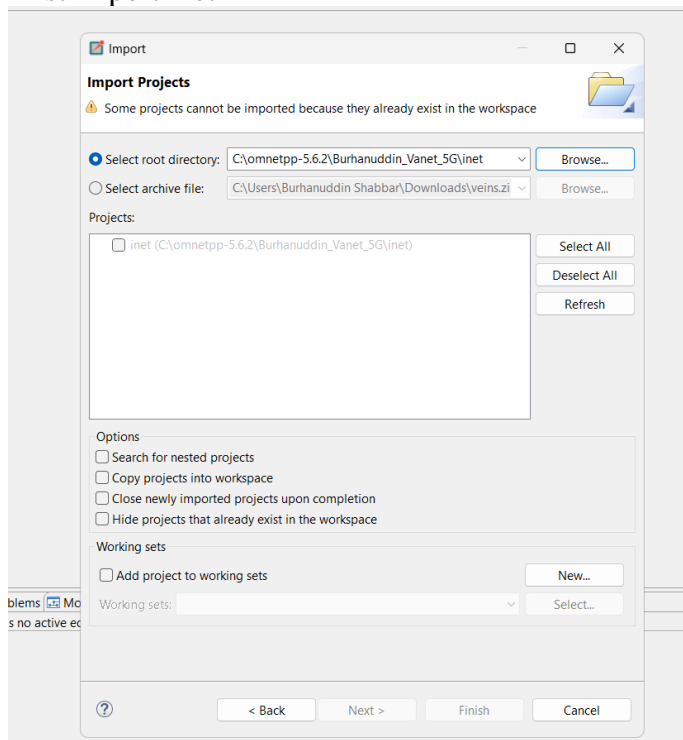


- First step is to install inet and viens frameworks. This is done by downloading both the frameworks from their official site and import it into the OMNET++ simulator.

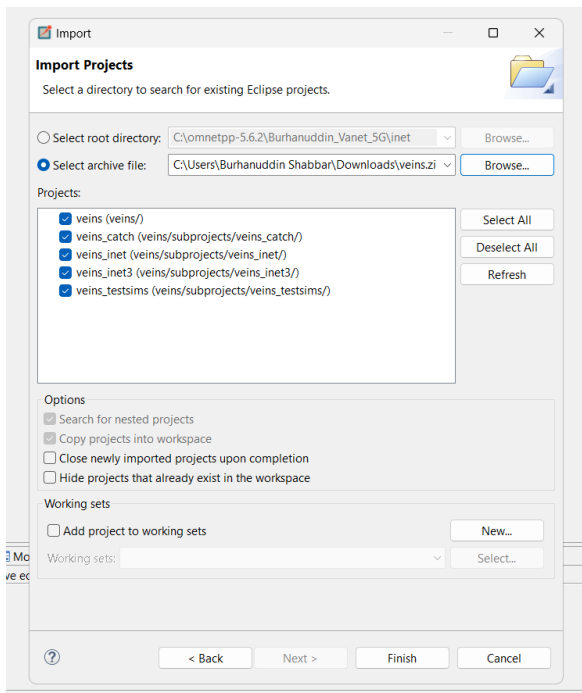
- Right click the taskbar to the left it will give you the option to import



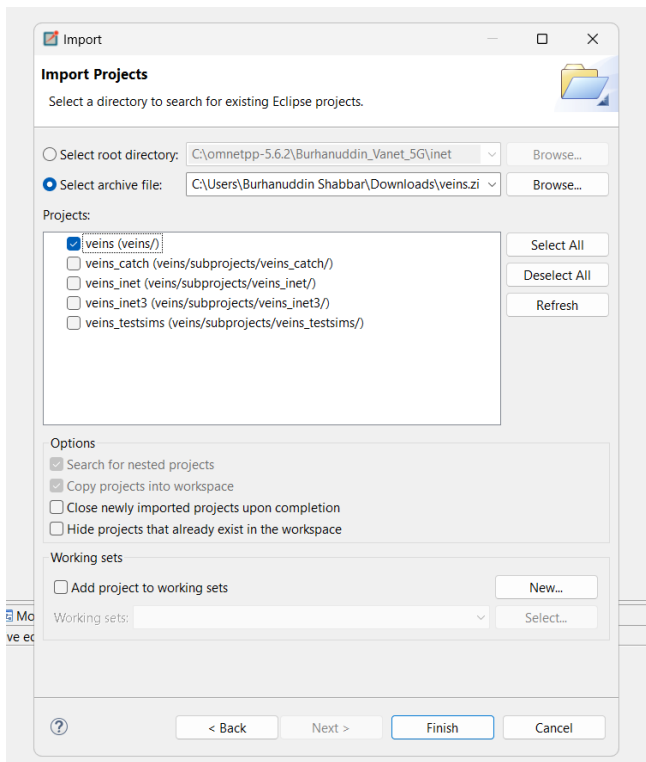
- First import inet



- After importing inet built the project
- After the built is completed. Now import viens into the simulator.



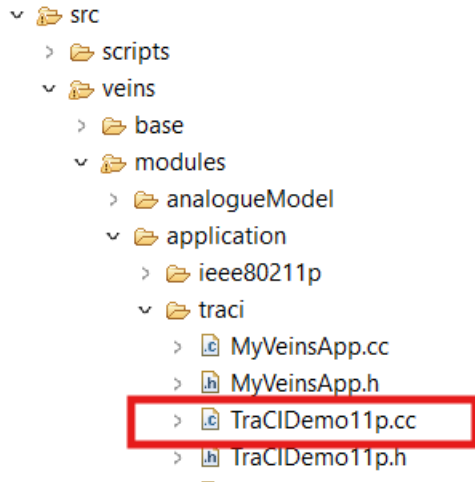
- Now unselect the other projects that is associated with veins. After doing that hit finish.



- Now select veins and built the viens framework.
- Setting up the simulation is completed.

4 IMPLEMENTATION OF THE ATTACK:

The following file contains the attack code:



4.1 DDoS Attack:

The following is the code for DDoS attack.

```
void TraCIDemo11p::handleSelfMsg(cMessage* msg) {
    if (msg->isName("DDoSAttack")) {
        int attackMessages = uniform(0, 4);
        double fakePacketSize = uniform(512, 2048);
        double fakeTransmissionTime = uniform(0.005, 0.02);

        for (int i = 0; i < attackMessages; ++i) {
            TraCIDemo11pMessage* attackMsg = new TraCIDemo11pMessage();
            populateWSM(attackMsg);
            attackMsg->setDemoData("DDoS Packet");
            sendDown(attackMsg->dup());

            // Record additional attack parameters
            recordScalar("Packet Size (bytes)", fakePacketSize);
            recordScalar("Transmission Time (s)", fakeTransmissionTime);

            // Increment attack statistics
            attackMessagesSent++;
        }

        EV << "DDoS attack executed, sending " << attackMessages << " messages.\n";
        recordScalar("DDoS Attack Messages Sent", attackMessagesSent);

        // Schedule the next attack
        scheduleAt(simTime() + uniform(1, 5), msg);
    } else if (auto* wsm = dynamic_cast<TraCIDemo11pMessage*>(msg)) {
        sendDown(wsm->dup());
        wsm->setSerial(wsm->getSerial() + 1);
        if (wsm->getSerial() >= 3) {
            stopService();
            delete wsm;
        } else {
            scheduleAt(simTime() + 1, wsm);
        }
    } else {
        DemoBaseApplLayer::handleSelfMsg(msg);
    }
}
```

4.2 Blackhole Attack:

The Following is the code for blackhole attack.


```

void TraCIDemo11p::onWSM(BaseFrame1609_4* frame) {
    TraCIDemo11pMessage* wsm = check_and_cast<TraCIDemo11pMessage*>(frame);

    if (isBlackholeNode == true) {
        EV << "Blackhole node dropping received message.\n";
        recordScalar("Messages Dropped by Blackhole", ++blackholeMessagesDropped);
        return;
    }
    EV << "Blackhole attack detected\n";

    findHost()->getDisplayString().setTagArg("i", 1, "green");

    std::string demoData = wsm->getDemoData();
    if (demoData != "DDoS Packet" && mobility->getRoadId()[0] != ':') {
        traciVehicle->changeRoute(demoData.c_str(), 9999);
    } else {
        EV << "DDoS Attack detected(link failure): " << demoData << "\n";
    }

    if (!sentMessage) {
        sentMessage = true;
        wsm->setSenderAddress(myId);
        wsm->setSerial(3);
        scheduleAt(simTime() + 2 + uniform(0.01, 0.2), wsm->dup());
    }
}
}

```

5 Run the simulation:

- To run the simulation first we have to launch sumo-gui.

```

Welcome to OMNeT++ 5.6.2!

/c/omnetpp-5.6.2$ omnetpp
Starting the OMNeT++ IDE...

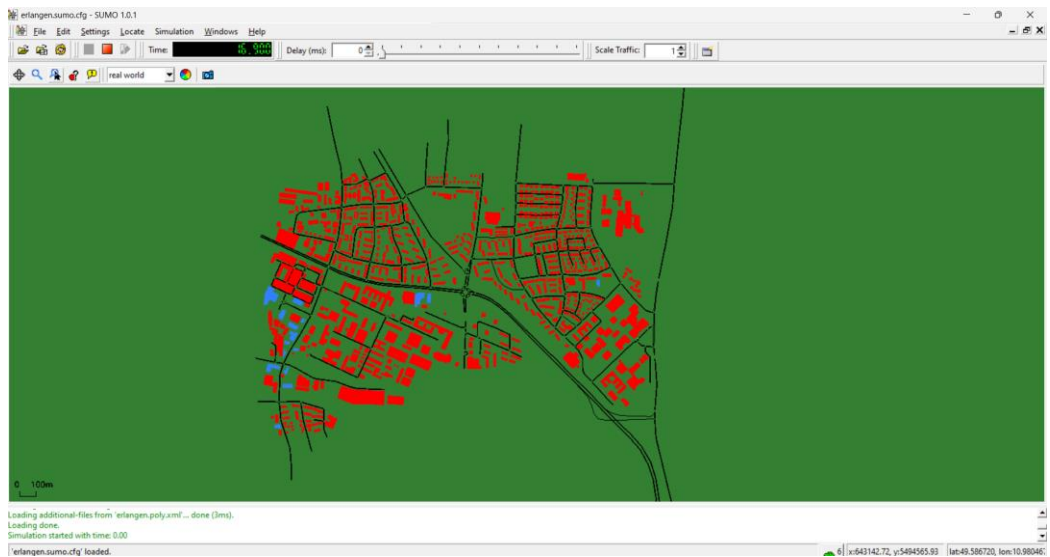
/c/omnetpp-5.6.2$ cd Burhanuddin_Vanet_5G/veins/

/c/omnetpp-5.6.2/Burhanuddin_Vanet_5G/veins$ ls
configure  doxy.cfg      images  print-veins-version  subprojects
COPYING    examples      Makefile  README.txt           sumo-launchd.py
doc         format-code.sh  out      src

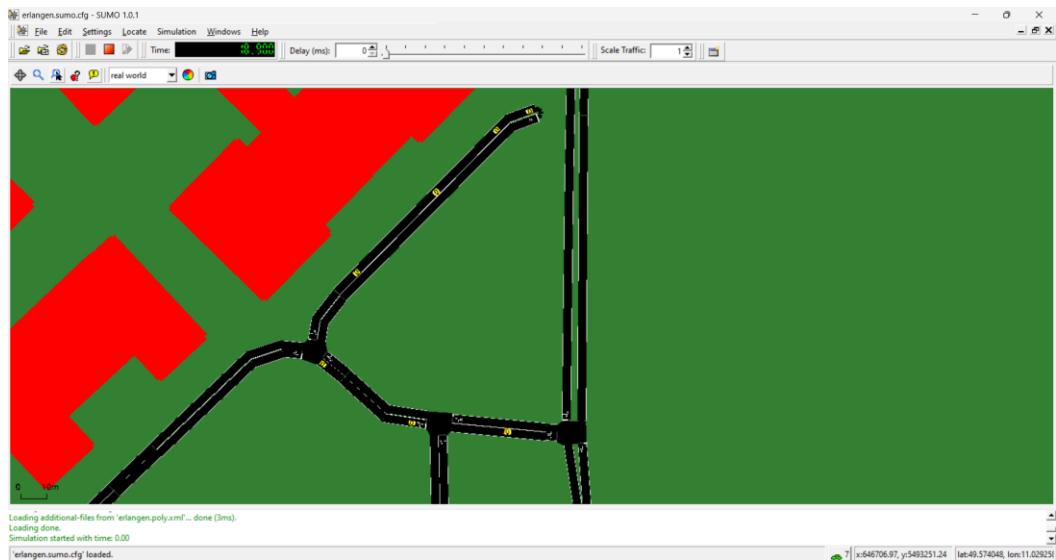
/c/omnetpp-5.6.2/Burhanuddin_Vanet_5G/veins$ sumo-launchd.py -vv -c 'C:\Program
Files (x86)\Eclipse\Sumo\bin\sumo-gui.exe'
Logging to c:/users/burhan~1/appdata/local/temp/sumo-launchd.log
Listening on port 9999
Connection from 127.0.0.1 on port 57531
Handling connection from 127.0.0.1 on port 57531
Got TraCI message of length 2
Got TraCI command of length 1

```

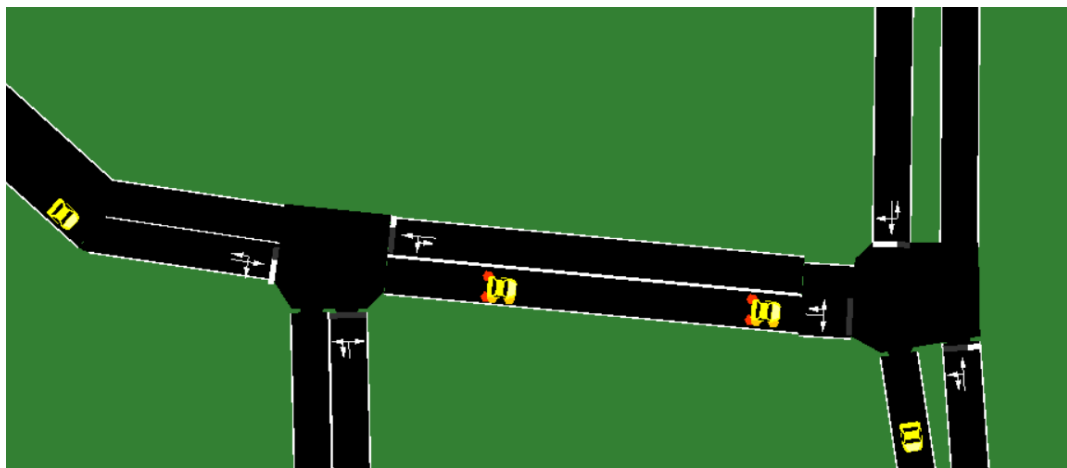
- After the sumo is launch. Hit the play button on the OMNET++
- Sumo Gui will be launched.



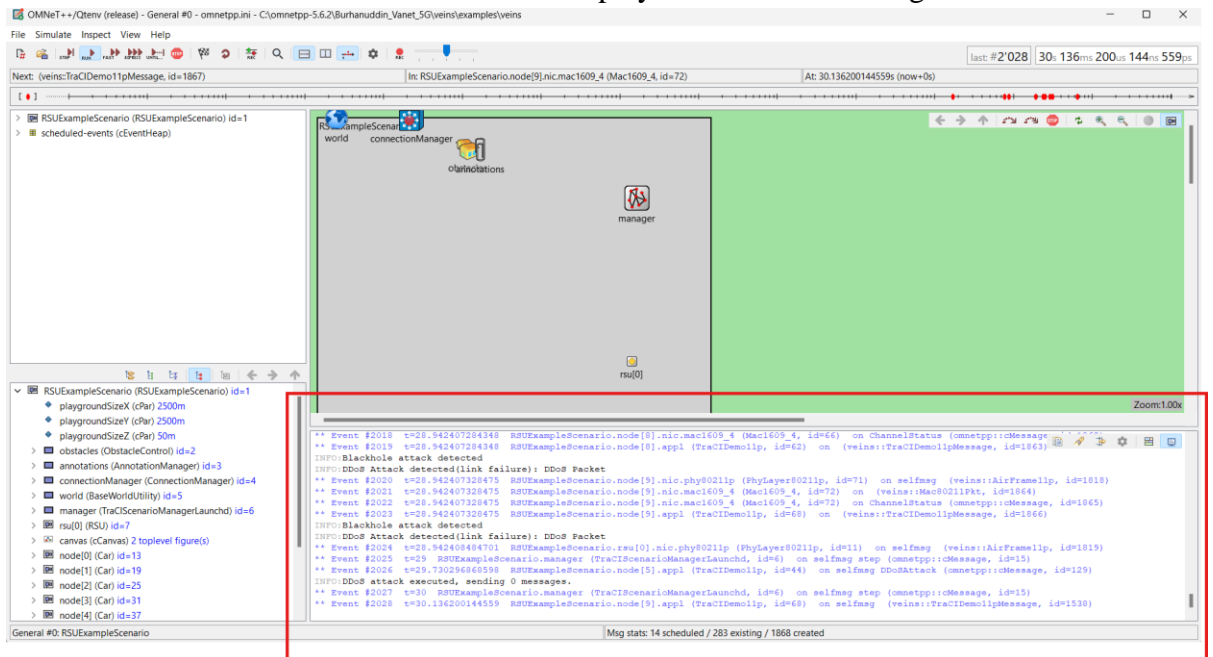
- The vehicles are automatically moving within the map. Some of these vehicles are malicious.



- The vehicles with red nodes are malicious ones.

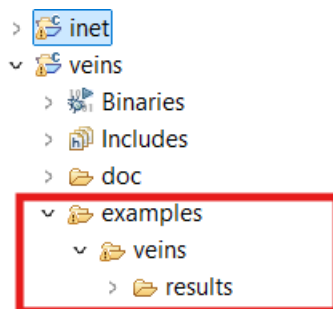


- When ever the attack is launched it is been displayed on the console.log

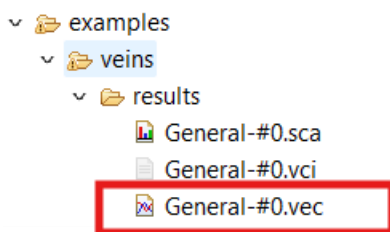


6 RESULTS OF ATTACK SIMULATION:

- The results of the attacks are being stored in the results folder.



- The simulation generates a new dataset each time it runs, replacing the previous set of findings.
- This is the file where the data is generated



- So, this is what data set looks like

Browse Data

Here you can see all data that come from the files specified in the Inputs page.

All (1186 / 1186) Vectors (100 / 100) Scalars (1086 / 1086) Histograms (0 / 0)

module filter statistic name filter

Experiment	Measurement	Repli...	Module	Name	Value
General		#0	RSUExampleScenari...	isBlackholeNo...	0.0
General		#0	RSUExampleScenari...	isBlackholeNo...	0.0
General		#0	RSUExampleScenari...	isBlackholeNo...	0.0
General		#0	RSUExampleScenari...	isBlackholeNo...	0.0
General		#0	RSUExampleScenari...	isBlackholeNo...	0.0
General		#0	RSUExampleScenari...	Packet Size (b...	1790.9040786...
General		#0	RSUExampleScenari...	Transmission T...	0.0193573273...
General		#0	RSUExampleScenari...	DDoS Attack ...	1.0
General		#0	RSUExampleScenari...	isBlackholeNo...	0.0
General		#0	RSUExampleScenari...	Packet Size (b...	759.77140915...
General		#0	RSUExampleScenari...	Transmission T...	0.0199294934...
General		#0	RSUExampleScenari...	Packet Size (b...	759.77140915...
General		#0	RSUExampleScenari...	Transmission T...	0.0199294934...
General		#0	RSUExampleScenari...	DDoS Attack ...	2.0
General		#0	RSUExampleScenari...	isBlackholeNo...	0.0
General		#0	RSUExampleScenari...	Packet Size (b...	1231.8482406...
General		#0	RSUExampleScenari...	Transmission T...	0.0126443655...
General		#0	RSUExampleScenari...	Packet Size (b...	1231.8482406...
General		#0	RSUExampleScenari...	Transmission T...	0.0126443655...
General		#0	RSUExampleScenari...	Packet Size (b...	1231.8482406...
General		#0	RSUExampleScenari...	Transmission T...	0.0126443655...
General		#0	RSUExampleScenari...	DDoS Attack ...	3.0
General		#0	RSUExampleScenari...	Packet Size (b...	1361.1337996...
General		#0	RSUExampleScenari...	Transmission T...	0.0053016132...
General		#0	RSUExampleScenari...	Packet Size (b...	1361.1337996...

- To export the data, select the data and click right and select the option for export data and click on csv records. The is now save on the system in csv format.

leNo...	0.0
leNo...	0.0
leNo...	0.0
le (b...	1790.90
ion T...	0.01935
ack ...	1.0
leNo...	0.0
le (b...	759.771
ion T...	0.01992
le (b...	759.771
ion T...	0.01992
ack ...	2.0
leNo...	0.0
le (b...	1231.8482406...
ion T...	0.0126443655...
le (b...	1231.8482406...
ion T...	0.0126443655...

Plot

+ Add Filter Expression to Dataset...

 + Add Selected Data to Dataset...

 Export Data >

CSV Records...

 CSV for Spreadsheets...

 JSON...

 OMNeT++ Scalar File...

 OMNeT++ Vector File...

 SQLite Scalar File...

 SQLite Vector File...

7 DATA PRE-PROCESSING:

7.1 Packages:

First install the necessary packages

```

# Install necessary packages
!pip install -q xgboost scikit-learn tensorflow matplotlib

# Import Libraries
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.metrics import precision_score, recall_score, f1_score, accuracy_score
from sklearn.tree import DecisionTreeClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from xgboost import XGBClassifier
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout
from tensorflow.keras.callbacks import TensorBoard
import matplotlib.pyplot as plt
from datetime import datetime

```

7.2 Dataset:

the NSL-KDD dataset being loaded and printed.

```

# Download the NSL-KDD dataset
!wget -q -O nsl_kdd.zip "https://github.com/defcom17/NSL_KDD/archive/master.zip"

# Unzip the dataset
import zipfile
with zipfile.ZipFile('nsl_kdd.zip', 'r') as zip_ref:
    zip_ref.extractall('nsl_kdd')

# Load the dataset
df = pd.read_csv('nsl_kdd/NSL_KDD-master/KDDTrain+.txt', header=None)

# Display the first few rows of the dataset
print("Dataset head:")
print(df.head())

```

Import the simulation dataset from the google drive

```
from google.colab import drive
drive.mount('/content/drive')

file1_path = '/content/drive/home/scalars2.csv'
file2_path = '/content/drive/home/vectors2.csv'

# Read the CSV files into DataFrames
df1 = pd.read_csv(file1_path)
df2 = pd.read_csv(file2_path)

# Display the first few rows of each DataFrame
print("First DataFrame:")
print(df1.head())
print("\nSecond DataFrame:")
print(df2.head())
```

7.3 Data pre-Processing:

```
# Data Preprocessing
# Column names based on the NSL-KDD dataset documentation
column_names = [
    "duration", "protocol_type", "service", "flag", "src_bytes", "dst_bytes", "land",
    "wrong_fragment", "urgent", "hot", "num_failed_logins", "logged_in", "num_compromised",
    "root_shell", "su_attempted", "num_root", "num_file_creations", "num_shells", "num_access_files",
    "num_outbound_cmds", "is_host_login", "is_guest_login", "count", "srv_count", "error_rate",
    "srv_error_rate", "rerror_rate", "srv_rerror_rate", "same_srv_rate", "diff_srv_rate",
    "srv_diff_host_rate", "dst_host_count", "dst_host_srv_count", "dst_host_same_srv_rate",
    "dst_host_diff_srv_rate", "dst_host_same_src_port_rate", "dst_host_srv_diff_host_rate",
    "dst_host_error_rate", "dst_host_srv_error_rate", "dst_host_rerror_rate",
    "dst_host_srv_rerror_rate", "label", "difficulty_level" # Added missing column
]
df.columns = column_names

# Encode categorical features and label
categorical_columns = ["protocol_type", "service", "flag", "label"]
for column in categorical_columns:
    df[column] = LabelEncoder().fit_transform(df[column])

# Define features and labels
features = df.drop('label', axis=1)
labels = df['label']
scaler = StandardScaler()
features_scaled = scaler.fit_transform(features)

# Train-Test Split
X_train, X_test, y_train, y_test = train_test_split(features_scaled, labels, test_size=0.2, random_state=42)
```

7.4 Algorithms:

Applying the machine learning algorithms.

```
# Machine Learning Algorithms
# 1. Decision Tree
dt_model = DecisionTreeClassifier()
dt_model.fit(X_train, y_train)
y_pred_dt = dt_model.predict(X_test)

# 2. Logistic Regression
lr_model = LogisticRegression(max_iter=1000)
lr_model.fit(X_train, y_train)
y_pred_lr = lr_model.predict(X_test)

# 3. Support Vector Machine (SVM)
svm_model = SVC()
svm_model.fit(X_train, y_train)
y_pred_svm = svm_model.predict(X_test)

# 4. XGBoost
xgb_model = XGBClassifier(use_label_encoder=False, eval_metric='mlogloss')
xgb_model.fit(X_train, y_train)
y_pred_xgb = xgb_model.predict(X_test)

# 5. Neural Network
log_dir = "logs/fit/" + datetime.now().strftime("%Y%m%d-%H%M%S")
tensorboard_callback = TensorBoard(log_dir=log_dir, histogram_freq=1)

nn_model = Sequential([
    Dense(64, input_dim=X_train.shape[1], activation='relu'),
    Dropout(0.5),
    Dense(32, activation='relu'),
    Dense(1, activation='sigmoid')
])

nn_model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
nn_model.fit(X_train, y_train, epochs=10, batch_size=32, validation_split=0.2, callbacks=[tensorboard_callback])
y_pred_nn = (nn_model.predict(X_test) > 0.5).astype("int32").flatten()
```

7.5 Generating results:

Evaluating and plotting the results.

```
# Evaluate Performance
def evaluate_model(y_true, y_pred, model_name):
    precision = precision_score(y_true, y_pred, average='weighted')
    recall = recall_score(y_true, y_pred, average='weighted')
    f1 = f1_score(y_true, y_pred, average='weighted')
    accuracy = accuracy_score(y_true, y_pred)
    print(f"Model: {model_name}")
    print(f"Precision: {precision:.4f}")
    print(f"Recall: {recall:.4f}")
    print(f"F1 Score: {f1:.4f}")
    print(f"Accuracy: {accuracy:.4f}")
    print("-" * 30)
    return precision, recall, f1, accuracy

# Collecting metrics for each model
metrics = {}
metrics['Decision Tree'] = evaluate_model(y_test, y_pred_dt, "Decision Tree")
metrics['Logistic Regression'] = evaluate_model(y_test, y_pred_lr, "Logistic Regression")
metrics['SVM'] = evaluate_model(y_test, y_pred_svm, "SVM")
metrics['XGBoost'] = evaluate_model(y_test, y_pred_xgb, "XGBoost")
metrics['Neural Network'] = evaluate_model(y_test, y_pred_nn, "Neural Network")

# Plotting the results
metrics_names = ['Precision', 'Recall', 'F1 Score', 'Accuracy']
for metric_index, metric_name in enumerate(metrics_names):
    plt.figure(figsize=(10, 6))
    plt.title(f'Model Comparison - {metric_name}')
    plt.bar(metrics.keys(), [metrics[model][metric_index] for model in metrics.keys()])
    plt.ylabel(metric_name)
    plt.xlabel('Model')
    plt.show()

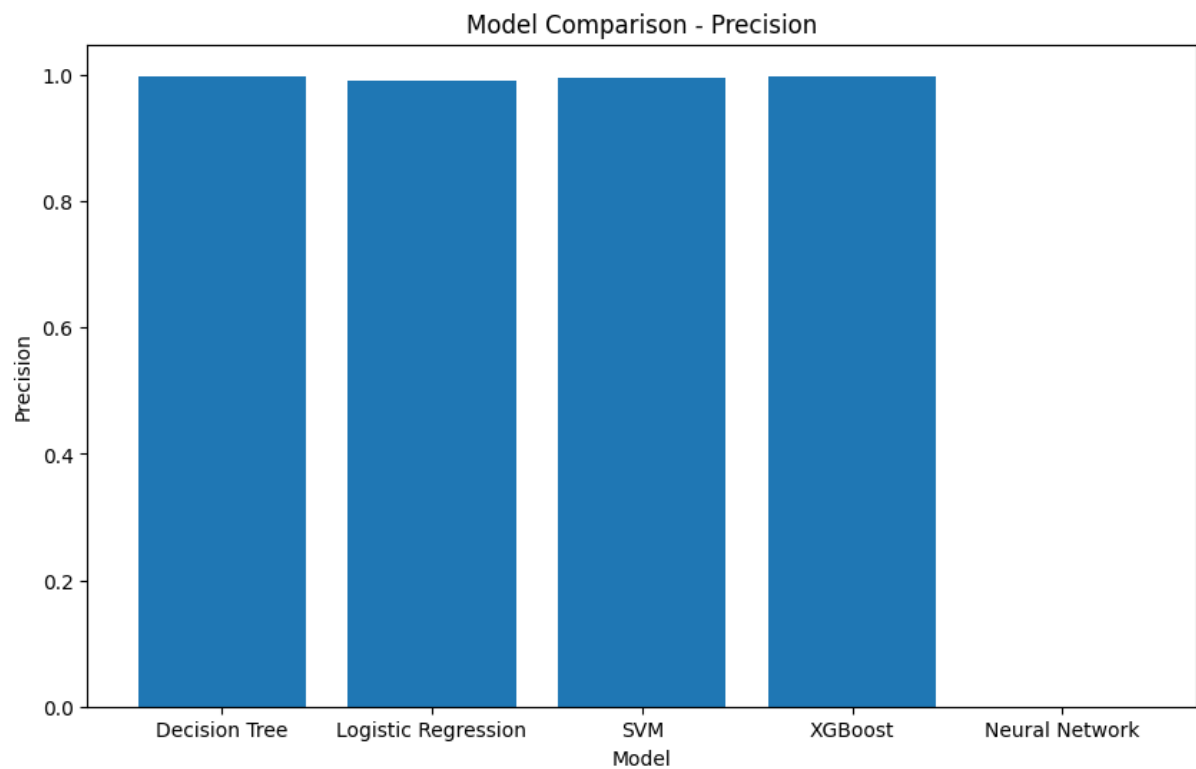
# Launch TensorBoard for Neural Network
%load_ext tensorboard
%tensorboard --logdir logs/fit
```

7.6 Results:

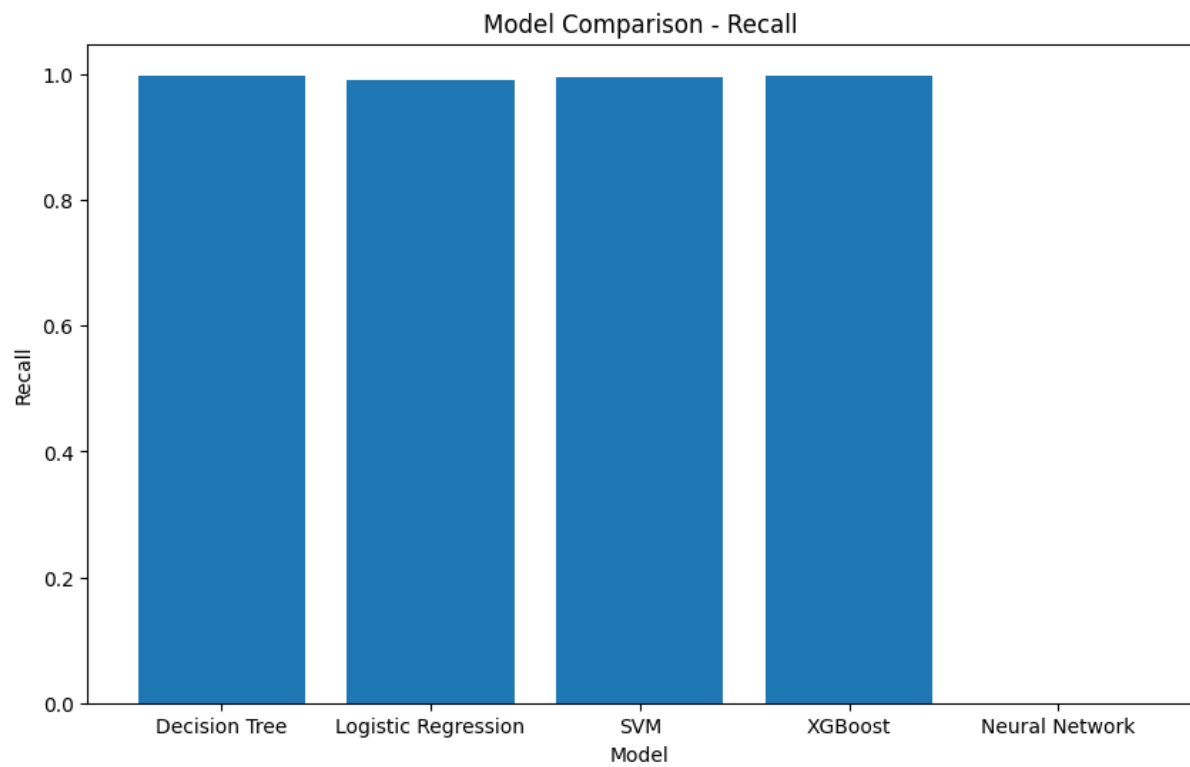
For are the results we got from this research

```
Model: Decision Tree
Precision: 0.9976
Recall: 0.9973
F1 Score: 0.9974
Accuracy: 0.9973
-----
Model: Logistic Regression
Precision: 0.9905
Recall: 0.9903
F1 Score: 0.9903
Accuracy: 0.9903
-----
Model: SVM
Precision: 0.9951
Recall: 0.9954
F1 Score: 0.9951
Accuracy: 0.9954
-----
Model: XGBoost
Precision: 0.9985
Recall: 0.9985
F1 Score: 0.9985
Accuracy: 0.9985
-----
Model: Neural Network
Precision: 0.0000
Recall: 0.0004
F1 Score: 0.0000
Accuracy: 0.0004
```

For precision:



For Recall:



For accuracy:

