An Intrusion Detection System Against Black Hole Attacks on the Communication Network of Self-Driving Cars

By Khattab M. Ali Alheeti

Presented by Danning Ma

vehicular ad hoc networks (VANETs)

Communicate with each other and road side units within radio coverage

high mobility, fast changing network topology, absence of fixed security infrastructures and open communication medium

---

V2V, V2I, I2I:

networks provide security and safety to passengers, drivers and vehicles by exchanging CAMs and emergency messages

Attack: black hole, gray hole, rushing and DoS attacks

Self-driving and semi self-driving vehicles:

- equipped with communication devices in the form of On Board Units (OBU) and an array of sensors and embedded systems

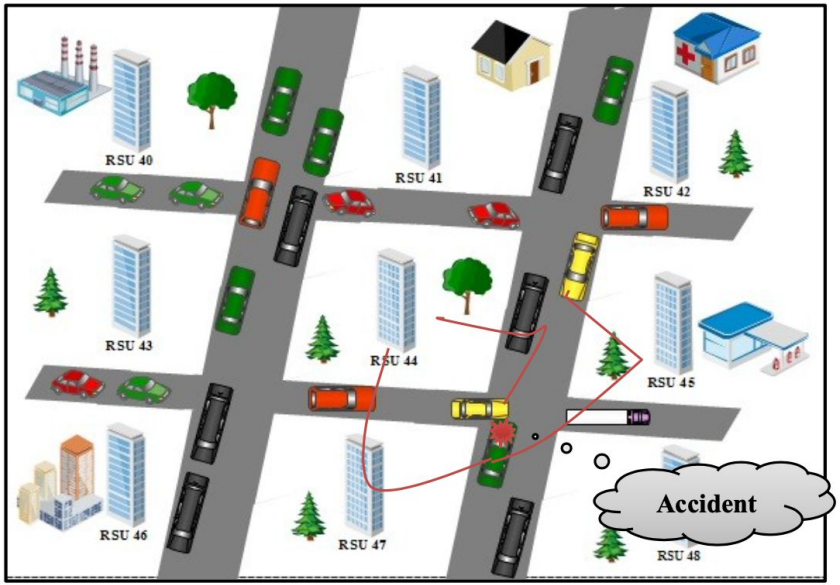


Fig. 1. An example of the process of responding to cases of emergency on the road

black hole attacks:

1) Inhibit forwarding of packets from one vehicle to its neighbor's "destination node".
2) Inhibiting the reception of packets from other vehicles.
3) Dropping all received packets.

Research:

simulate an intelligent intrusion detection system IDS that is mounted directly in the vehicles rather than the (RSUs) road side units

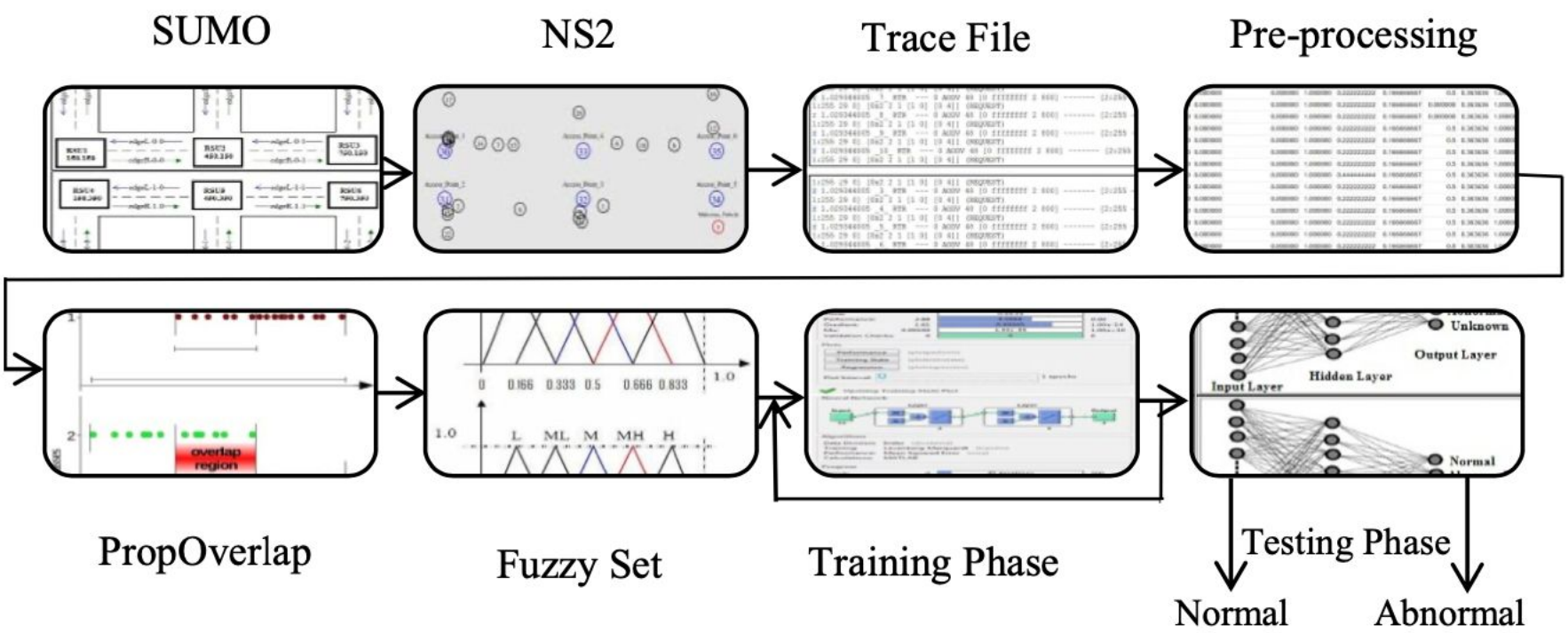
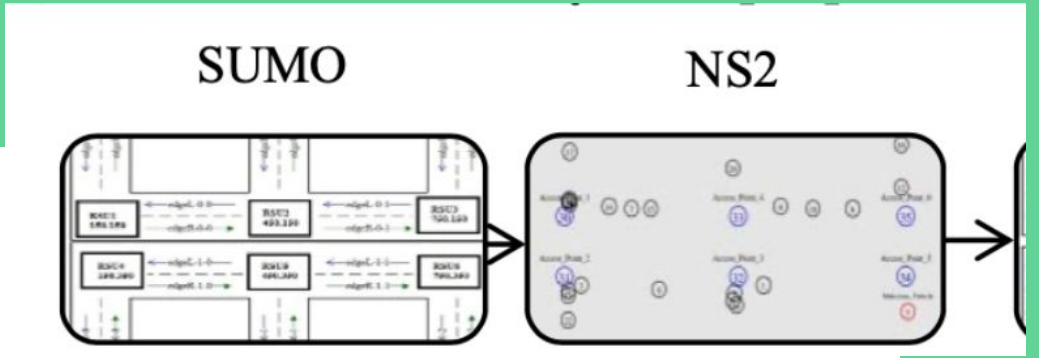
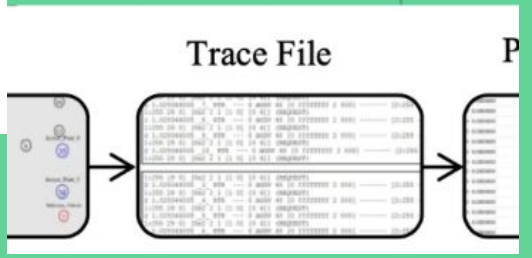


Fig. 3 Architecture of IDS



Simulation of Urban Mobility Model

- NS2 simulations
    - Simulation of Urban Mobility Model (SUMO) and MObilty VEhicles (MOVE)
        - SUMO provides efficient computation even in various sizes of scenarios
        - MOVE receives the files produced by SUMO by converting them to the NS2 format and immediately using in network simulation.



Feature Sets

The trace file generated in NS2 is divided into three groups: “basic trace”, “internet protocol trace” and “AODV trace”

$$X = \frac{X - MIN}{MAX - MIN} \qquad (2)$$

Feature Extraction

The distinguishing extracted features are singled out by analyzing the overlap among the expression values across two classes.

Proportional Overlapping Scores (POS) method

algorithm pseudo code:

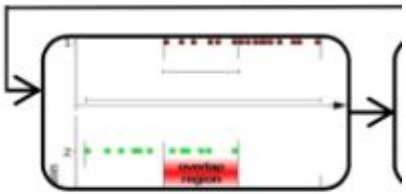
Algorithm POS Method

```
1.  Inputs: "data1.csv".
2.  Output: Sequence of the selected features.
3.  install.packages("propOverlap").
4.  source("http://bioconductor.org/biocLite.R").
5.  biocLite("Biobase").
5.  library(propOverlap).
6.  ?propOverlap.
7.  getwd().
8.  data <- read.csv("data1.csv",header=T).
9.  str(data).
10. data <- t(data).
11. G <- data[1:21,] # define the features matrix 21.
12. G <- jitter(G). # to avoid the noise in data
13. Class <- as.factor(data[22,]) #define the
    observations' class labels.
14. set.seed(1234).
15. Selection <- Sel.Features(G, Class,
    K=21,Verbose=TRUE) # the main function.
16. Selection$Features. # extract the number of
    features
17. Selection$Measures. # extract name of features.
```

employed all extracted features from trace file [22].

Table 1 Performance Metrics

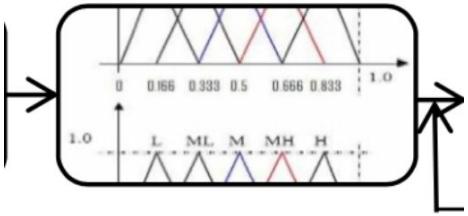
	IDS with all Features	IDS with 15 Features
Training Rate	98.97%	99.86%
Average False Alarm	6.21%	0.53%
Error Rate	2.05%	0.15%
TrainParam.Epochs	68	15



PropOverlap

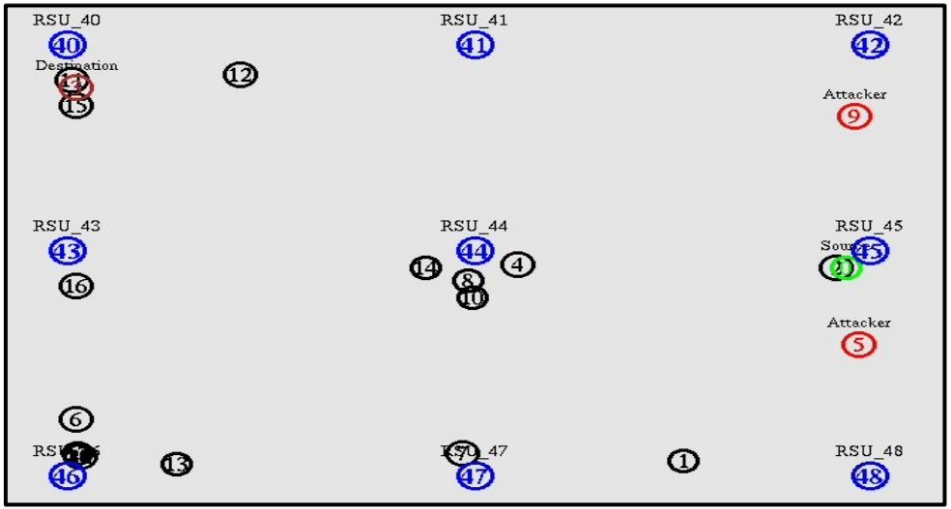
Fuzzy Membership

$$f(x,a,b,c) = max(min(^{x-a}/_{b-a}, ^{c-x}/_{c-b}),0) \quad (1)$$



Fuzzy Set

Simulation Environmental and Parameters



Intelligent Intrusion Detection System

Table 4 Classification Rate

IDS					
Class	Original Records	ANN	Match Records	Miss Records	Accuracy
Normal	19285	19288	19261	27	99.87%
Abnormal	10715	10698	10685	13	99.72%
Unknown	0	14	0	14	NaN

training phase are TrainParam. Epochs=15, TrainParam.lr=$1*10^{-7}$, TrainParam.goal=0 and TrainParam.min_grad = $1*10^{-12}$.

Generating Malicious behavior

Training and Testing Neural Network with (Misuse Detection)

Table 2 Classification Rate

IDS					
Class	Original Records	ANN	Match Records	Miss Records	Accuracy
Normal	6382	6381	6375	6	99.89%
Abnormal	3618	3617	3611	6	99.80%
Unknown	0	0	0	2	NaN

Table 3 Alarms Rate

Alarm Type	**Accuracy**
True positive	99.90%
True negative	99.83%
False negative	0.09%
False positive	0.16%

Training and Testing Neural Network (Anomaly Detection)

Table 4 Classification Rate

IDS					
Class	Original Records	ANN	Match Records	Miss Records	Accuracy
Normal	19285	19288	19261	27	99.87%
Abnormal	10715	10698	10685	13	99.72%
Unknown	0	14	0	14	NaN

Table 5 Alarms Rate

Alarm Type	Accuracy
True positive	99.86%
True negative	99.87%
False negative	0.14%
False positive	0.12%

Discussion

Any problems and costs in this design ?

Eg. Performing a fuzzy set “fuzzification” on the dataset which was extracted from the trace file. This approach has a direct positive impact on the result by increasing the detection rate, decreasing the false alarm rate and error rate. However, the main drawback is that the system needs extra memory resources to store data and the approach is more computationally heavy.