Density Based Deceptive Data Detection and in VANETs

S.Sampath Kumar,C.Gnanavel,M.Dheepak
Department of EEE, AMET Deemed to be University ,Chennai

Abstract: **Deceptive Data Detection (DDD) in Vehicular Ad-hoc Network is a well- known research domain which holds the major concern on data integrity and data reliability. This problem comes under the category of data integrity checking which confronts a lifesaving schema when the data become sensitive. The traditional methods for targeting data integrity checking takes a concern amount of reliable checking time and the delay will get reflected in data transmission. Though the quality of solution seems an effective one, the delay has to be a major concern since even a millisecond delay causes disaster to a greater extent. This intended research interest helps the author to propose an effective way to discover a prominent deceptive data model which can detect the deceptive data based on density of vehicles in VANET simulation region to increase the reliability of deceptive detection**

Keywords:**Vanets, Data integrity Data reliability, Deceptive data**

1.Introduction

Pervasive Networking holds the capability to access the services of any type from anywhere for communication purposes. Pervasive networking evolved to the level of commercial aspect by which the use of such network takes place in commercial workplaces and in residential regions. Pervasive network consist of mobile nodes for which acts independently irrespective of its position and environment. The explicit use of pervasive networks are Wireless Mesh Network (WMN), Mobile Ad hoc Networks (MANET) and Vehicular Ad hoc Network (VANET). Pervasive networks can satisfy multiple agents from a fixed single access point. One such application is VANET where the RSU will be a fixed access point and the vehicles as mobile units. A network that connects the vehicles on roads with each other directly or indirectly through fixed access point are called as VANET. VANET technology uses moving vehicles as nodes in a network to create a mobile ad hoc network. In VANET, the node may be a vehicle or the road side units. The communication models for efficient data transfer is of two different forms namely, Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communication model. VANET is a decentralized architecture which also holds the property of self-organizing. The theme of VANET is to transfer the messages either emergency or entertainment messages to vehicles. The message can be of either broadcasting manner or on-demand request from respective vehicles.

2.Objective

VANET play a vital role by achieving safety to lives. A timely message on a complete upcoming road accident avoids a pileup. Avoidance of traffic congestions, injuries and other life threaten issues can be avoided. Achieving integrity of data in VANET, implies a clear picture of finding deceptive data. Deceptive data detection in VANET increases the complexity due to insufficient time delay to deliver the messages. A prominent and efficient message transfer on time with robust and reliable data improves the VANET consistency. Based on the comprehensive study made on VANET research issues, an efficient Deceptive Data Detection Model is proposed in Vehicular Ad-hoc Network. Although many security mechanisms are proposed in finding data integrity, still there exists loop holes in breaking those mechanisms and deception is still an active research region.

3.Related Work

The authors Kaur. H et al [2015] proposes a scheme, which has the capability for handling bogus information identification, marking of suspicious node, analyzing traffic pattern, packet flood identification modules. The proposed scheme will point out malicious nodes on the basis of identifying bogus information on the very first step. Once the above mentioned process is completed the traffic analyzing is carried from the nodes, and the pattern is incorporated. The flood identification modules over the traffic is generated and received from the different components of the VANET cluster. The adversary patterns or floods will be identified and their sources will be checked through the security mechanism. The performance of the proposed scheme will be calculated using several parameters such as throughput, end to end delay, packet delivery ratio, and data drop rate.

Shubham Gandhi et al [2014] discuss about various security protocols for carrying the secure message communication between the vehicles. This paper discusses the various security protocols mainly used for vehicular ad hoc networks. It provides information about the behavior of protocols by understanding their characteristics and challenges.In [Biswas et al 2010] authors discuss about the survey of the existing methods on security and privacy in VANETs. This paper projects distinct viewpoints, and categorize them depending on some of encryption and decryption techniques. The implementation of functionalities and other key characteristics are discussed. The identification of the characteristics for variety approaches and focused their comparative merits and demerits.

The authors [Aslam et al 2010] proposes information two novels for providing reliable traffic information propagation and two-directional verification of data verification, and time-based verification of data. The traffic message is passed through two channels and these two channels are defined either spatially or temporally spaced. A receiver vehicle verifies the integrity of message by checking the data received from both channels and checks the message is matched. The comparison is carried with the popular cryptographic based security systems. The proposed schemes are much easier and cost effective to implement, especially the initial transition level for VANET network infrastructure does not exist.

4.Proposed System

Consider two ways, one lane road. Let the vehicle flow in the two ways one lane road be regular vehicle and public vehicle. Road Side Unit (RSU) will be placed at the side of the road within the particular range. Regular vehicle includes personally owned car, bus, lorry, etc.., and Public vehicles includes police cars, school bus, police vans, ambulance etc.., Each vehicle is equipped with an Onboard unit in which the license number of each vehicle will be fixed. The license number of each vehicle is burned in the onboard unit by the department of transportation, so no intruders can change that fixed license number. The license number is the unique identification number for each vehicle.

Here we are using vehicle to infrastructure communication i.e. vehicle to RSU communication to verify whether it is true or false accident report. In vehicle to infrastructure communication each vehicle broadcast its report to its nearby RSU. The RSU in turn will send their report to the Department of Transportation.Here in our proposed approach we are using vehicle to infrastructure communication. Once an accident happen the accident report is sent from accident vehicle through its own sensor to its corresponding nearby RSU. The sensor may be fixed anywhere in the vehicle. If the entire vehicle got damaged in an accident, at least there will be some parts (tire) from which the information is sent to the RSU. When a sensor is deployed, it picks up a role from the role set based on the specific features of the sensor, such as storage size, computation ability, communication ability, and trustable level, which specifies how much can this type of sensor be trusted. Then each sensor plays a different role in the system and has the functions assigned to that role. In our proposed approach, there are two roles in our system (i.e.) Regular vehicle and Public vehicle. Public vehicles are trustier than Regular vehicle because it cannot be easily compromised by malicious vehicle. Here, we assign higher priority to public vehicle than the regular vehicles.

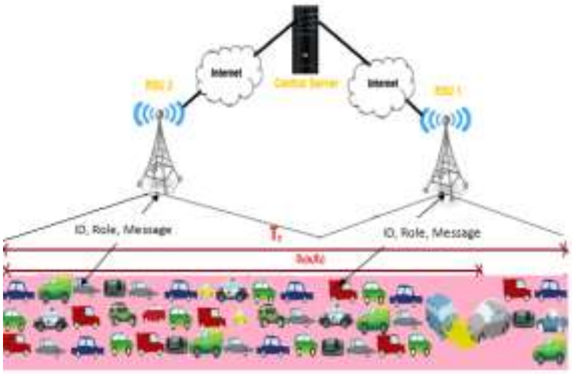


Figure.1 Density based approach within the range

Once an accident happen the accident report is sent from accident vehicle to its corresponding RSU. The report sent will be in the format of <ID, Role, Message> where ***ID*** denotes the License number of the vehicle which is the unique identification number, ***Role*** denotes role of the vehicle which is regular or public vehicle and ***Message*** denotes whether it is the true or false report. The constant value is given to both public and regular vehicles. High value is assigned to the public vehicle than the regular vehicle because public vehicles are more trustable than the regular vehicle.

5.Density Based Deceptive Data Detection

The proposed system detects the deceptive data based on the density of the traffic. This Module has been designed in order to reduce the computational time to find out the deception data and also to reduce the computational cost for evaluating data deception process. A glimpse of this module is explained as follows.A vehicular network consists of a set of vehicles, road side unit (RSU), and a number of remote servers. Usually, the vehicles that are moving on the road will collect information such as road conditions, accident report etc.., the moving vehicles are considered as nodes in a network to create a dynamic mobile network. As vehicles fall out of the signal range it gets dropped from the network, other vehicles can join in, connecting vehicles to one another so that a mobile Network is created.

Data in VANET are of two types: regular data and events data [Biswas et al 2010]. Regular data includes all normal data such as normal weather conditions, normal road conditions and some information about the vehicle itself such as location, direction and velocity. This type of data usually used for long-term analysis and it has low real-time requirements. The events data includes sudden brake, heavy traffic, car accident and wild weather conditions. This type of data usually has high real time requirements, which should be delivered to others vehicles on the road to give warning message. In VANET data can be transmitted using two ways of communication vehicle to vehicle communication (V2V) and vehicle to infrastructure communication (V2I). In V2V communication, vehicles will communicate with other vehicles within the dedicated short range communication.

Each vehicle is equipped with an On Board-Unit (OBU), which integrates the technologies of wireless communications, micro sensors, embedded systems, and Global Positioning System (GPS). In V2I communication the vehicles will communicate with RSU (Road Side Units). RSUs are the static fixed sensors that are deployed along the road within the particular distance. RSUs are more powerful than vehicles in terms of communication and computation. RSUs will then report to the remote servers using internet.

6.Density Calculation

Density of the vehicles plays the major role in detecting the deceptive data which includes false data. Density refers to the number of vehicles still remains in the particular area. By the amount of vehicles in the particular area we can determine whether it is true or false report. In our proposed work we assign role for each vehicle. Role of the vehicles can be categorized into public vehicle and regular vehicle. Each vehicle is assigned with the constant value. Public vehicle is assigned with the highest value because it is trustier than the regular vehicle. Public vehicles include police car, school vans, police van, ambulance etc., regular vehicle includes personally owned cars, buses etc.

Our system assumption is one lane two-way road. There may be n number of vehicle flow which may be a regular vehicle and public vehicle. Here we are using the vehicle to Infrastructure communication. So the vehicles communicate its report to its nearby RSU. For density based approach let T_R be the total range covered by two RSU. Let $D_1(R_1,a)$ be the distance between the starting point of first RSU to accident place. Our aim is to find the distance between accident place to second RSU which is denoted as $D_2(a,R_2)$.Once an accident happens the accident report is first sent from the accident vehicle to its corresponding RSU which is denoted as A_r. let V_f be the actual number of vehicle flow within the range of T_r meters. The actual flow of vehicles includes both regular and public vehicles. Let X_{veh} be the Average flow of vehicles for meters T_r which includes both regular and public vehicles. As our system is a vehicle to Infrastructure communication. Each vehicle sends the message to its corresponding RSU. From that, we can determine the number of vehicle flowing in that area. The first step for finding deceptive data in density calculation is to find the distance between accident places to its nearby RSU. Once an accident happens the accident report is sent from the accident vehicle which is referred as (A_r) to its corresponding, then take its next nearby RSU range which is defined as D and find the distance that the RSU covers.The distance that RSU cover is considered as T_R. The distance between starting point of RSU to accident place is denoted as $D_1(R_1,a)$. Now find the distance from accident place to D_2 which is denoted as $D_2(a,R_2)$.For calculating the distance from accident place to D_2as $D_2(a,R_2)$ subtract total distance covered by RSU (T_R) with the distance from the D_1 to the accident place $D_1(R_1,a)$.

After finding $D_2(a,R_2)$, if distance is $D_2(a,R_2)$ then next step is to find the average vehicle flow for total distance covered by RSU (and also finds the average vehicle flow for particular distance $D_2(a,R_2)$.Let us consider the average vehicle flow for (R) meters is X_{veh}, then the average vehicle flow for $D_2(a,R_2)$ meters will be calculated by average vehicle flow for T_R meters with the distance of $D_2(a,R_2)$ meters which should be divided by the total distance covered by RSU which is denoted as T_R meters which is given as (X_{veh}* $D_2(a,R_2)$)/ T_R .X_{veh} includes both public and regular vehicles.After finding the average flow of vehicle for the particular distance $D_2(a,R_2)$ meters from the total distance, then find the actual flow of vehicle in $D_2(a,R_2)$ meters. Each vehicle has a unique ID called as license number which cannot be altered by anyone and it will be fixed on the On board unit. This ID denotes the vehicle count because of its unique nature. Count the ID of the vehicle till it reaches zero. ID is equal to zero then no more vehicle is flowing in that particular distance of $D_2(a,R_2)$ meters. V_f be the actual flow of vehicle in that particular area.After finding the average flow and actual flow of the vehicle of the vehicle in $D_2(a,R_2)$ meters, then compare both actual glow and average flow of the vehicles in$D_2(a,R_2)$ meters to find whether the accident really happens or not. If V_f is the actual flow of vehicle for $D_2(a,R_2)$ at the time of accident report received. If V_f(actual flow of the vehicle for $D_2(a,R_2)$ meters) is lesser than the average flow of the vehicle for distance $D_2(a,R_2)$ meters then we conclude it as a true accident else it is consider as a false accident. From these density calculations we can come to a conclusion whether the accident is happened or not. About 80% of true accident will be confirmed with the density calculation.

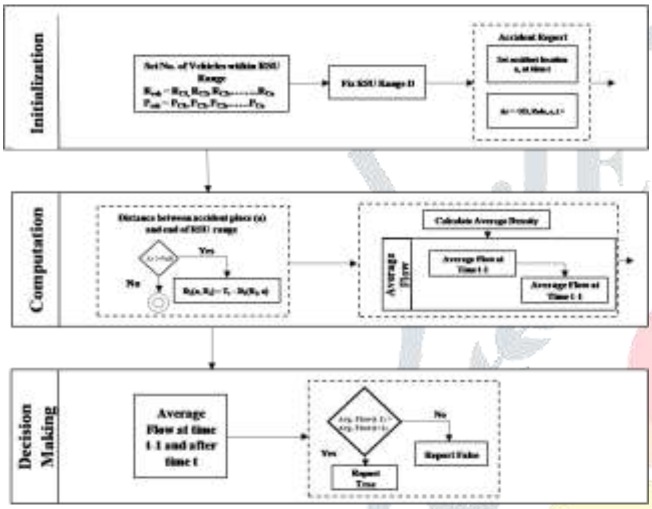

Initialization
Set No. of Vehicles within RSU Range
R_{veh} = R_{C1}, R_{C2}, R_{C3}.......R_{Cn}
P_{veh} = P_{C1}, P_{C2}, P_{C3}.......P_{Cn}
Fix RSU Range D
Accident Report
Set accident location a, at time t
Ar = <ID, Pos, a, t>
Computation
Distance between accident place (a) and end of RSU range
Yes
No
$D_2(a,R_2) = T_r - D_1(R_1,a)$
Calculate Average Density
Average Flow
Average Flow at Time t-1
Average Flow at Time t-1
Decision Making
Average Flow at time t-1 and after time t
Avg. Flow(t-1) > Avg. Flow(t+1)
No
Yes
Report False
Report True

Figure.2 Block diagram for density based deceptive data detection

Our proposed algorithm can be explained in detail with an example. As our proposed approach is based on vehicle to Infrastructure communication, consider each vehicle sends its report to its corresponding RSU. Let us consider the first message will be sent from the accident vehicle to its corresponding RSU about the accident report. Let us consider A_r be the accident report from accident vehicle. As soon as the accident report is sent from accident vehicle to its corresponding RSU. Let us consider the RSU which receives the accident report. Our aim is to find the distance from accident place to RSU. Let us consider the range of RSU be 5km. Now the distance from RSU to accident place be 2km. So the distance from accident place to D_2 can be found by subtracting the total range covered by RSU with the distance from D_1 to the accident place $D_1(R_1,a)$. which is given as 5km-2km=3km. So, the distance from accident place to D_2 be 3km. After finding the distance from accident place to D_1 next step is to find the average flow and actual flow of vehicles within that particular range.

Let us consider the average flow of vehicles for the total range be X_{veh}. X_{veh}includes both regular and public vehicles, then the average vehicle flow for $D_2(a,R_2)$ will be calculated by the equation(2) as (X_{veh}* $D_2(a,R_2)$)/ T_R where X_{veh} includes both regular and public vehicles, $D_2(a,R_2)$ be the distance from RSU2 to accident place and T_R be the total range covered by RSU. Let us consider X_{veh} be 300 vehicles. Let the distance from D_2 to accident place be $D_2(a,R_2)$ meters is 8km. the total distance covered by RSU is 5 km so we get the average flow as (300* 8)/ 10=240 vehicles. so Let us consider the average flow of vehicle is 240 vehicles. Now we have to find the actual flow of vehicles within the range. The actual flow of vehicles can be found by the ID which is sent by each vehicle by its report. ID is the unique identification number known as license number which cannot be altered by anyone and it is fixed in the onboard unit. With this ID unique number, we can find the number of vehicle flow in a particular range. So let us consider the actual flow of vehicles within the distance from accident place to D_2 be 100 vehicles. After finding the actual flow of vehicle and the average flow of vehicles within the distance from accident place to D_2. Now next is to find whether the report is a true report or it is a false report. First we have to find the accuracy of the received report if the actual flow of vehicle flow is less than the average vehicle in that particular range then it is true accident else if the actual flow of vehicle flow is greater than the average vehicle in that particular range then it is false accident report here the average flow of vehicle within the distance is 240. The actual flow of vehicle within the range is 100.Now we found that the average flow of vehicles within the range is greater than the actual flow of vehicles within the particular distance. That is 240 vehicles is greater than the 100 vehicles. So we come to the conclusion that true accident has happened. With the help of this techniques, we can attain the result as with the Density calculation 80% of true accident reports will be confirmed.

6.1.Variable used for density based deceptive data detection

T_R	←Total range covered by RSU
$D_1(R_1,a)$	←Distance between starting point of RSU to accident place
$D_2(a,R_2)$	← Distance from accident place to end point of RSU Range
A_r	←Accident report from accident vehicle
V_f	←Actual number of vehicle flow within range
R	←Average Density

Assumptions:

Two lane one way road
Regular vehicles are personally owned car, bus, lorry, etc.
Public vehicles are police cars, school bus, etc.
Average flow of vehicles for meters is X_{veh}.

7.Algorithm For Density Based Deceptive Data Detection

```
begin
//* Initialization *//
Set the number of Vehicles in RSU Range
$N = |Rveh, Pveh|$
a.  $R_{veh}$    $R_{C1}$,$R_{C2}$ ,$R_{C3}$ ……..$R_{Cn}$
b.  $P_{veh}$    $P_{C1}$,$P_{C2}$ ,$P_{C3}$ ……..$P_{Cn}$

//* Accident Report $A_r$ from position 'a' at time t *//
$- < a, t, Di >$

//* Computing distance between accident place and end of RSU range *//
If ($\neq \emptyset$ ) then
    $(_2) \leftarrow T_r$ - ( ,
else
    No Accident reported
    exit()
End

//* Calculating Average Vehicle Density between 2 RSU's per unit time*//
$Avg_{density} = \frac{=_0 R_{veh} P_{veh}}{nit\ Time}$

//* Average flow of vehicle at time t-1 in $D_2(a,R_2)$ *//
Compute rage $Flow^{t-1}_{D_2(a,R_2)} \leftarrow \frac{R*D_2(a,R_2)}{t2-t1}$

//* Average flow of vehicle at time t+1 in $D_2(a,R_2)$ *//
```

```
Compute $\text{rage Flow}_{D_2(a,R_2)}^{t+1} \leftarrow \frac{R*D_2(a,R_2)}{t4-t3}$

//* Evaluation of Reported Accident *//

if ($\text{rage Flow}_{D_2(a,R_2)}^{t+1} < \text{Average Flow}_{D_2(a,R_2)}^{t-1}$ ) then
    return Accident report is True
else
    return false
end

end
```

8.Testbed Design

In order to implement our proposed system, we use VANET simulator called a traffic simulator. We design our simulator by extending a traffic simulator that simulates the movements of the vehicles, such as acceleration, deceleration, speed and lane changing. In addition, we simulate the scenario of accidents, as well as malicious vehicles. Simulation ran in Intel i7 core processor 3rd generation of 2.4GHz and 2 GB RAM, IDE Eclipse Kepler version 4.3 to produce GUI for our proposed system.

8.1Traffic Micro Simulation:

In our simulation, we use traffic micro simulation called traffic simulator. Micro simulation is a term used in traffic modeling and is typified by software packages such as TransModeler, PTV VISSIM, TSIS-CORSIM, Cube Dynasim, LISA+, QuadstoneParamics, SiASParamics, Simtraffic and Aimsun. Empirical modeling software such as LINSIG, TRANSYT, TRANSYT-7F or SIDRA INTERSECTION represents a different class of models based on deterministic methods. Traffic micro simulation models simulate the behavior of individual vehicles within a predefined road network and are used to predict the likely impact of changes in traffic patterns resulting from changes to traffic flow or from changes to the physical environment. Micro simulation has its greatest strength in modeling congested road networks due to its ability to simulate queuing conditions.

Table.1 Parameter Setup for Simulation

Parameters	Range
Number of Vehicles	400, 700 and 1000
RSU Transmission Range	0.5 to 0.7
Vehicle Speed Limitation	120Km/hr
Accident Report generation	At 10th second
Average Flow (Time Period)	100 to 600 milliseconds
Number of Simulations	30

This capability makes these types of models very useful to analyze traffic operations in urban areas and city centers, including interchanges, roundabouts, unsignalized and signalized intersections, signal coordinated corridors, and area networks. Micro simulation also reflects even relatively small changes in the physical environment such as the narrowing of lanes or the relocation of junction stop lines.Vehicles are generated in equal interval time during simulation. The total simulation period for a single run takes 30 seconds. All the vehicles follow a random path around the simulated scenario.

9.Performance Metric

9.1Recall:

Recall can be defined as the ratio between the number of accurate identification of deception to the total number of simulations with respect to time. It can be formulated as

$$Recall = \frac{\# \varphi_1}{\# \varphi}$$

where φ refers the total number of simulation runs and φ_1 refers accurate identification of accident report during simulation.

10.Experimental Results

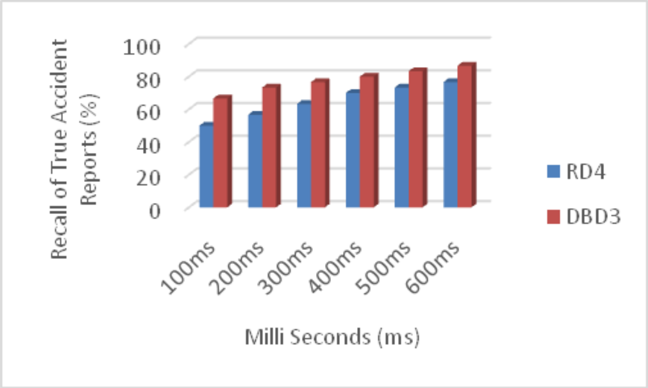
After the simulation process gets started the average flow of vehicle on each RSU will be calculated. Once the accident report is initiated the RSU is instructed to calculated the average flow of vehicles after $t+1$ where t is the time of accident. The proposed module has been compared with existing RD4 technique in the form of accuracy.

Table.2 Experimental result analysis

Run	Flow at time t in D₂		Flow after time t in D₂		Final Report
	Actual Flow	Average Flow	Actual Flow	Average Flow	
1	344	207	192	117	True
2	319	164	280	175	False
3	334	220	241	149	True
4	291	157	227	159	False
5	311	177	264	134	True
6	258	160	218	150	True
7	339	192	226	134	True
8	330	168	267	142	True
9	267	165	201	152	True
10	309	191	203	164	True
11	313	202	249	145	True
12	302	216	214	122	True
13	275	195	218	168	True
14	321	174	226	157	True
15	252	178	172	171	True
16	323	209	234	140	True
17	283	185	251	142	True
18	288	160	248	169	False
18	258	196	264	152	True
20	333	181	198	126	True
21	318	196	224	149	True
22	313	166	250	136	True
23	315	173	197	166	True
24	311	166	205	135	True
25	326	218	242	145	True
26	285	154	232	169	False
27	278	185	211	117	True
28	267	197	202	165	True
29	338	186	236	133	True
30	264	214	256	154	True

11.Performance Analysis

Results of the proposed system and existing system is taken from the simulation for every 100 milliseconds after the accident report generated. proposed Density based Deceptive Data Detection provides better results (when compared to the existing algorithm RD4. At the 100thand 200th millisecond of accident report our proposed system works better than RD4 with 17% increased accuracy. At the 300thmillisecond the proposed system works better than RD4 with 14% of increased accuracy.



On 400, 500 and 600th millisecond proposed system results shows a performance increase of 10% than existing algorithm. This result that the proposed algorithm works better than existing with an overall of 15% for low dimensional vehicles.

Conclusion

A comprehensive survey has been made over the recent related works and it has been concluded with the necessity for having an improved model for data integrity in VANET using efficient security mechanisms. An efficient Deceptive Data Detection using density based approach in VANET, which contemplates the vehicle density between RSU's for effective detection of deceptive data. The promising experimental results demonstrate the impact of the proposed models in terms of accuracy and defends the arguments of proposed model over other existing techniques. The constructive and encouraging results justify the significance and necessity of the proposed line of research and of course it may encourage further enhanced investigation in the identified area of research.

References

[1] Abdulkader.Z.A, Abdullah.A, Abdullah.M.T, ZukarnainZ.A (2017), "Vehicular ad hoc networks and security issues: survey", Modern Applied Science 11(5), pp. 30-41.

[2] Khan.U, Agrawal.S, Silakari.S (2015), "Information Systems Design and Intelligent Applications. A detailed survey on misbehavior node detection techniques in vehicular ad hoc networks." (Springer, 2015), pp. 11–19.

[3] Raiya Rashmi, and Shubham Gandhi (2014), "Survey of Various Security Techniques in VANET." International Journal of Advanced Research in Computer Science and Software Engineering 4, no. 6, pp. 431-433.

[4] Shubham Gandhi, Shalini (2014), "Security Protocols for Vehicular Adhoc Networks: A Review", IJCSMC, vol. 3, no. 5, May 2014, pp. 948 – 953.

[5] Yang.W (2013), "Security in vehicular ad hoc networks", Wireless network security, Berlin Heidelberg, Springer, pp. 95–128.

[6] Yousefi.S, Mousavi.M.S, Fathy.M (2006), "Vehicular ad hoc networks (VANETs): challenges and perspectives", Proc. of the Sixth Int. Conf. on ITS Telecommunications, pp. 761-766.

[7] Zhang, Jie, Chen Chen, and Robin Cohen (2010), "A scalable and effective trust-based framework for vehicular ad-hoc networks", Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, vol.1, no. 4, pp. 3-15