This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2732727, IEEE Access

ANTSC: An Intelligent Naïve Bayesian Probabilistic Estimation Practice for Traffic Flow to Form Stable Clustering in VANET

Amjad Mehmood[1], Akbar khanan[2], Abdul Hakim H. M. Mohamed[2], Houbing Song[3*]
[1]Institute of Information Technology, Kohat University of Science & Technology, Kohat, Pakistan
[2]Department of Management Information System, A'Sharqiyah University, Oman
[3]Department of Electrical and Computer Engineering, West Virginia University, Montgomery, WV 25136 USA
[*]Corresponding Author: H. Song (Senior Member, IEEE)

Abstract— The Vehicular Ad-hoc Network (VANET) is one of the promising and encouraging technologies, and it is going to make a splash in the near future. VANET has turned into a main module of the Intelligent Transport System (ITS). It is a self-controlled, wheeled network (also called Network on Wheels), a wider, stimulating class of Mobile Ad-hoc Network (MANET). VANETs raise many innovative challenges due to their high-class and unique features, such as high node mobility, dynamic topology changes, wireless links breakage, network constancy, and network scalability. A well-organized routing protocol is one of the most challenging matters of such networks. In this paper, we propose an intelligent naïve Bayesian probabilistic estimation practice for traffic flow to form a stable clustering in VANET, briefly named ANTSC. The proposed scheme aims to improve routing by employing awareness of the current traffic flow as well as considering the blend of several factors such as speed difference, direction, connectivity level, and node distance from its neighbours by using the intelligent technique. The proposed technique has proven to be more strong, stable, robust, and scalable than existing ones.

Keywords—VANET, MANET, Intelligent Transportation Systems, V2V, V2I, Routing, Clustering, Traffic flow, Naïve Bayesian, Scalability, Reliable

I. INTRODUCTION

Safety as well as luxury in travel has gained significant importance in social life. For the past few decades, every year many people have lost their lives, while others have been injured in highway accidents because the driver is unable to estimate the circumstances on the road ahead. Well-organized traffic control systems are also becoming abundant. On the other hand, traffic crowding is becoming an increasingly severe problem. Due to the crowding effect, vehicles waste millions of hours and gallons of fuel daily. Therefore, a clogged flow condition has a damaging influence on the economy, health, and environment. However, a huge quantity of accidents and congestion may be avoided by gathering and allocating information related to human safety. The information is collected using an intelligent network scheme among moving vehicles on urban roads or highway setups. As part of Intelligent Transportation Systems (ITS), it affects the financial and social causes of motor vehicle accidents.

Vehicular Ad-hoc Networks (VANETs) have attracted growing attention from both industry and theoretical scholars recently [1-4].

VANETs are a special case of MANETs, where the mobile nodes in these networks are vehicles that are fitted out with wireless communication devices, i.e. on board units (OBU), application units (AU), global positioning systems (GPS), e.g. digital maps, and thus move on roads at very great speed following traffic instructions, and spread or receive information [24-29].

Generally, we have two classes of communications in VANETs: vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I).

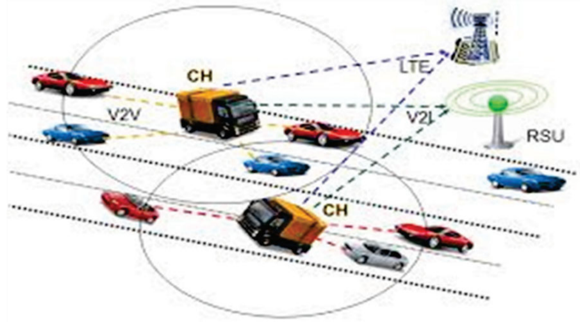

Fig. 1: VANET system architecture

There are two core units: the roadside unit (RSU) and OBU. In addition, standards in VANETs have been established by the IEEE, referred to as Wireless Access in Vehicular Environment (WAVE). In the IEEE 802.11p protocol, CSMA/CA is utilized for medium access control [5]. VANET permits vehicles to connect at about 100–300 meters from each other, and thus the created OBU consists of a wireless communication unit, such as dedicated short range communication (DSRC) or 3G, a memory set, and a processor in order to provide V2V and V2I communications. Fig. 1 above clearly demonstrates the VANET's organization. Usually, the VANET structural design facilitates two forms of communication unit, which are the OBU and RSU.

As specified above, OBUs are fixed in the vehicle, and RSUs are a static unit located on the roadside or near to traffic lights. The RSUs act similarly to an access point and are able to provide the communications with infrastructure (i.e., 2G, 3G, or microwave). The above-mentioned networks have an extensive range [6]. A vehicle acts as a router as well as a source and/or destination to exchange information packets in either one hop or multi-hop fashion.

Compared to other MANETs, VANETs are categorized by high vehicle movement, regular topology variations and interruption to the wireless connection. Furthermore, numerous applications have different QoS requirements. Mostly, safety applications need rapid and consistent message transfer, though non-safety applications need a high output and fair access to utilize the channels [10].

As stated directly above, since VANETs have been used in numerous applications such as safety and sensitive information related to road conditions, both applications need a highly unchanging topology, so improved solutions for clustering in VANETs are needed. Clustering is an eminent method to form groupings of vehicles and thus organize ad hoc networks, and is also an effective method to make the VANET global topology less dynamic. It allows the creation of a virtual

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2732727, IEEE Access

communication that supports well-organized distribution of data in VANETs [7]. The clustering approach is a well-organized and efficient key to the scalability issue [8]. The hidden terminal problem can also be reduced by clustering [9]. Clustering in VANETs is an operative method to shorten some significant functions, such as vehicle management, routing, medium access management, the provision of resources and bandwidth [10].

The stable clustering methods have been well considered in Mobile Ad-hoc Networks (MANETs) in recent years. However, considering the high-class features of VANETs, these traditional and outdated clustering methods are inappropriate for VANETs. Several stable clustering-based schemes in VANETs have been proposed in the literature. However, it is a significant issue with all these old schemes that they do not keep the long life of clusters and the stability of the cluster head, due to the high movement of vehicles, the continuously changing topology of vehicles and other related factors in VANETs.

In this paper, we proposed an intelligent clustering algorithm whose purpose is to extend the lifetime of a cluster head. We will take advantage of being aware of the lane of automobiles on the road and then transmit this awareness to further adjacent vehicles to determine the best and ideal cluster head. Our technique of choosing the cluster head is the essential way to attain a more stable cluster. Our scheme takes the traffic flow lane use information, speed alteration, direction, connectivity level, cluster size and vehicle distance into consideration, for selecting a long-lasting cluster head.

The remainder of the paper is arranged as follows: Section I, sums up and reviews the background and a few associated techniques that are reported in the literature. Section II describes the proposed Naïve Bayesian based cluster head (CH) selection method, the highlighted model of CH, and the cluster formation and selection process. Section IV includes the simulation results of the presented technique. The conclusion and future of the current scheme are drawn in section V.

II. Background & Related Work

A cluster is an abstract structure in which vehicles having some characteristics in common are merged in a group and thus form a cluster. For ease, every cluster has its own cluster head, or there can be more than one, depending on the scenario. The vehicle nominated to be a cluster head will manage the whole cluster as well as communication-related tasks. A number of clustering techniques and methods for VANETs are found in the literature. However, some of the proposed methods use the MANET clustering techniques to form the clusters, though these works cannot be suitable for VANETs due to the major difference between VANETs and MANETs.

In order to improve the standard cluster lifetime and number of cluster heads per vehicle [7], the Intelligent Based Clustering Algorithm in VANET (IBCAV) is proposed. This technique seeks to enhance the algorithms of routing in VANETs by using inter-layered approaches, having knowledge of the current traffic flow as well as the mixture of numerous attributes including the size of the cluster, the vehicle speed and density of vehicles. The core technique for selecting the cluster head is based on the smart technique that is called an artificial neural network.

To provide stability in the cluster lifetime for VANETs, [10] proposed a clustering algorithm that is based on the flow of the traffic, aiming to establish stable clusters in urban scenarios. The flow of the traffic, vehicle speed variation and vehicle position are taken into consideration in this paper. A vehicle with a heavy flow of traffic in its environment is given higher importance to be selected as the cluster head.

To extend the lifetime of the cluster head, [11] discusses a technique where the head of the cluster is selected depending on the lane having the maximum traffic flow. Vehicles have knowledge of the traffic in every lane, i.e. lane recognition, lane direction and map matching on the road, and they broadcast the gathered information to adjacent vehicles. This helps in selecting an efficient and stable cluster head.

The authors of [12] proposed a new VANET cluster creation algorithm in order to extend the cluster lifetime and reduce the vehicle alteration per cluster. This technique groups vehicles that are moving fast in high-speed lanes into one cluster, while vehicles that are moving slowly are grouped into another cluster. This procedure takes into consideration the speed alteration between vehicles as well as their location and direction throughout the cluster creation procedure.

The cluster head selection process is one of the major challenges in VANETs. In [5] a fuzzy logic based cluster head selection algorithm is proposed. The technique used four fuzzy input functions that are direction, speed, acceleration, and distance, leading to better decisions on cluster-head selection that can improve the overall network lifetime. The proposed algorithm selects the CH from both low-speed and high-speed vehicles on a two-way multi-lane highway.

To improve the connectivity and reduce the interruption when a vehicle changes its cluster head, [13] offered a new clustering metric, the 'vehicle interconnection metric', which discloses the movement resemblance of vehicles in the neighborhood. It is constructed on beacon frames directed among vehicles and aims to find the communication capabilities between each pair of them, rapidly adjusting to the alterations in the network. In this technique, a clustering algorithm with double cluster head connectivity is proposed, thus eliminating the need for the periodic selection of a new cluster head as the old cluster head leaves the cluster.

In [20], the authors offer a novel clustering technique for determining cluster configuration parameters by using the physical position of the vehicles. A group of vehicles selects a CH, whereas the size of cluster is determined by the CH.

The authors of [21] presented a multi-hop clustering procedure and announced a new movement metric according to the comparative mobility between vehicles in a multi-hop distance, selecting the vehicle with the minimum cumulative mobility value as the cluster head.

Furthermore, a multi-agent driven dynamic clustering system is presented in [22,23]. It presented a dynamic cluster creation system that contains hefty-weight fixed and light- weight mobile agents and clusters the vehicles that show related mobility patterns, directions, and speeds.

TECHNIQUES	CH SELECTION PARAMETER	PURPOSE OF CLUSTERING	PERFORMANCE METRICS	SIMULATOR	SCENARIO
IBCAV (2013) [8]	Velocity, density, cluster size, smart method called artificial neural network)	Scalability	Throughput, End-to-end delay, Packet delivery ratio	NS3	Highway
Traffic flow based clustering scheme for VANETs (2014) [10]	Flow of the traffic, speed alteration, vehicle position	Vehicle, resource & medium access management, routing, bandwidth allocation	Mean cluster head lifetime and number of clusters altered per vehicle	NS-2 simulator (version 2.35), SUMO	Urban
Lane-based detection clustering algorithm (2011)- [11]	Traffic flow, speed, distance [head will be selected from the lane having heavy traffic flow]	Scalability, hidden terminal problem	Cluster head, stability	NS3, Intelligent Driver Model (IDM), MOBIL lane change model	Urban
Cluster-based traffic information generalization in VANET (2014) [17]	Benefit factor: Time to leave, relative average speed, neighbourhood degree	Network stability & scalability	Accuracy of CH by low, medium, and high flow of traffic	OMNet++, SUMO	Highway
Fuzzy logic based clustering strategy for improving VANET performance (2014) [5]	Speed, acceleration, distance, direction	High mobility & scalability, control of the topology	Check CH variations for low & high speed vehicles	OPNET, MATLAB Fuzzy Inference System editor	Highway
Multi-head clustering algorithm in Vehicular Ad-hoc Networks (2013) [19]	Relative position and mobility	Transmission collisions & increased resource utilization	Average number of clusters, average number of CMs, average cluster lifetime, average idle time, average residence time	NS-2, VanetMobiSim	Not mentioned

Table 1: Comparison between different state-of-the-art protocols

It also presents the method to predict similar cluster participants based on their movement patterns for future suggestion.

III. Proposed Technique

A. Stable Clustering Structure based on Traffic Flow

In this section, we consider the whole core process of the clustering algorithm based on traffic flow. As well as the system scenarios and scheme model, we consider the cluster head selection algorithm and the clustering procedure.

B. System Scenario's Synopsis and Suppositions

All vehicles should have the same processing power and storage in our assumptions. The concept of CH rounds per vehicle in a cluster is not used in our scenario, which is mostly used in WSN CH selection, as the node in this network is a vehicle that has no battery issue. When the CH is leaving, it will inform ahead for selecting a new cluster head or give this role to the second winning cluster head. When two heads have the same result, the head will be chosen randomly. Moreover, we consider urban scenarios with intersections in our proposed technique. In the given scenario, the flow of traffic separates at each intersection. Every single intersection is based on three traffic flows: left-turn flow, right-turn flow and straight-flow movements. The cluster head will be selected from the lane having the heaviest traffic flow, aiming to increase both the lifetime of the cluster head and the stability of the cluster. For example, if we have five lanes on the road and three of them are straight-flow lanes, the cluster head will be selected from the lanes that have a straight flow because these lanes have more vehicles and have the same lane. Fig. 2 displays a scenario with six lanes. Lane 1 is the lane that turns left with two vehicles. Lanes 2, 3, 4 and 5 have straight-flow with 10 vehicles, overall. Lane 6 turns to the right with three vehicles. Observably, the cluster head will be elected from the vehicles on Lane 2, 3, 4 or 5 as these lanes have heavy flows of traffic and fewer vehicles will leave the cluster after crossing the intersection, which in turn increases the lifetime of both the cluster head and the clusters. As mentioned earlier in section 2, each vehicle is fitted with a GPS device, with the help of which every vehicle is capable of recognizing its exact position and immediate movement information, such as its route, locality, information on the lane to which it belongs, and speed. However, GPS has a 5 meters' fault which is greater than the distance between lanes [11]. Various additional methods are found in the literature to help identify lanes on the road. One technique is to combine the GPS with a wheel odometer [14]. The wheel odometer can identify vehicle movement of about tenths of millimeters [15]. Moreover, other techniques studied in the literature did not use GPS but used other methods that were able to detect lanes on the road [15, 16]. The following figure shows the system scenario.

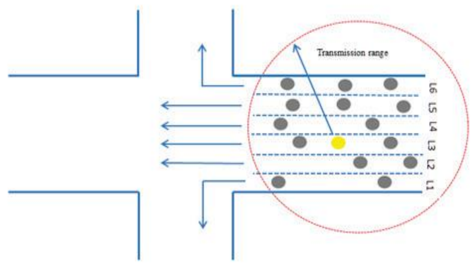

Fig. 2: System scenario

Clustering Process Transition Model Each vehicle in the cluster is allocated a unique ID and each vehicle from time to time broadcasts its immediate movement information including its speed, distance, cluster size, direction and connectivity level and lane number to its surrounding neighbouring vehicles. 'M' is the time between two successive broadcasts. The importance of a vehicle to be selected as a CH is based on its neighbours' movement information as we are following the traffic flow technique. In the cluster formation level, vehicles may have one of four roles: UN (unknown vehicle), CH (cluster head), CM (cluster member), and TM (temporary member):

- UN is a vehicle that is not currently a member of any cluster

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2732727, IEEE Access

NOTATIONS	DESCRIPTIONS
BI (Broadcast Interval)	Time interval occupied by a vehicle to broadcast a HELLO packet
TI (Timeout Interval)	Time after an unreachable/temporary member is removed from a neighbour table.
HELLO	Periodically broadcast packet by each vehicle that contains the vehicle movement information and the current role of this particular vehicle
JOIN-INVITE	Packet broadcast by a CH to invite an unknown vehicle to join a cluster
JOIN-REPLY	Broadcast packet sent by an unknown vehicle to CH to acknowledge a join invitation
JOIN-REQUEST	If an unknown vehicle receives no invitation message from CHs due to certain conditions, this particular vehicle will send this message to its nearest CHs for joining the cluster.
$\pm\triangle V_{th}$	Threshold to categorize neighbouring vehicles as stable or non-stable neighbours.

Table 2: Notations and descriptions

- CH is a vehicle that is selected to handle cluster-related responsibilities
- CM is a vehicle that is a part/member of a cluster
- TM is a vehicle that does not get a HELLO packet from the cluster head within a TI interval.

A vehicle will play one of these four roles and may transit from one role to another if the cluster structure is reformed or changed due to road conditions. Fig. 3 shows the whole clustering procedure model.

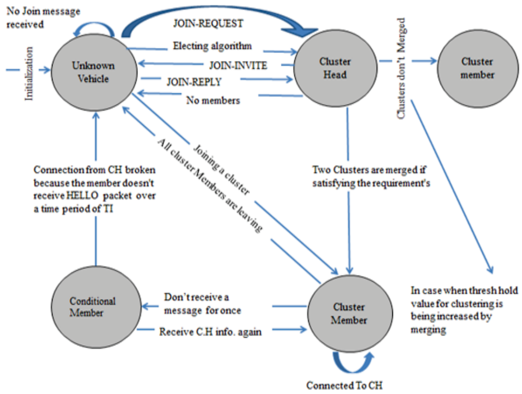

Fig. 3: Clustering Creation Procedure

C. *Clustering Creation Phase*

The clustering model consists of the following steps: initialization, joining, cluster head selection algorithm, short-term disconnection from CH, leaving procedure, merging procedure, non-merging procedure, all CMs' leaving procedure, and the procedure for giving up the CH role.

The notations given below will be used in the clustering formation phase:

CREATION PROCEDURE

i. Initialization: When a vehicle joins the road. Each vehicle is initialized to be a UN. Such a vehicle does not yet receive any JOIN message or HELLO message.

ii. Every vehicle broadcasts a HELLO packet per BI. Joining Procedure: All CHs broadcast the *JOIN-INVITE (JI)* message along with their movement information every M time period. If an unknown vehicle that is not part of any cluster receives a JI message, that vehicle will first check whether the direction of the CH is the same as its own or not. Secondly, it will check that its current speed difference from the CH is within $\triangle\pm V_{th}$. If both conditions are satisfied, then the unknown vehicle will send a *JOIN-REPLY (JR)* message to the relevant CH. When this CH receives the JR message, it sends an ACK packet with the unique ID of this new cluster member. If the unknown vehicle does not receive any invitation message from a CH, this particular vehicle will send a *JOIN-REQUEST (J-REQ)* message to its nearest CH for joining the cluster.

iii. CH selection procedure: If an unknown vehicle does not receive any (JI) message during M interval of time, the vehicle will start the cluster head selection method. It collects the information of nearer surrounding vehicles. The proposed algorithm to select the cluster head will be discussed below.

iv. Short-term disconnection with CH: When a CM does not receive any HELLO packet from its CH over a time of TI, the role of this vehicle changes from cluster member to temporary member, as its membership within the cluster depends on certain conditions. It does not lose its membership instantly, as the interruption might be due to poor performance of the wireless connection. If the TM again receives packets broadcast by the CH in the next TI duration, the role reverts to CM.

v. TM leaving procedure: When a temporary member fails to receive any CH packets repeatedly over a time period of TI, the role of this vehicle changes to an unknown vehicle. In the meantime, the CH will also remove this unreachable member from the membership list.

vi. Merging procedure: If two CHs are in direct communication range, a merging procedure is performed. The cluster head with fewer members will give up the role of cluster head and join the cluster that has more members.

vii. Non-merging procedure: In the situation of two clusters merging, if the size of the new cluster would exceed the predefined threshold value, the merging procedure will not take place, as managing a larger cluster is also a major issue for the CHs. In this case, both CHs will continue to perform their own roles and manage their own clusters.

viii. All CMs leaving procedure: In the situation when all the cluster members move to separate lanes, the cluster will no longer exist and all vehicles will become unknown vehicles and will again search for another cluster.

ix. CH role giving up procedure: When all the members leave the cluster for whatever reasons, the CH will also revert to being an unknown vehicle again.

D. *CH Selection Parameters*

i. Traffic weight

The significant approach in our work is to consider the neighbours surrounding the vehicle, which we call the traffic flow. Different traffic flows lead in different directions, and will then divide after the intersection. Vehicles in different traffic flows must be grouped into one cluster to reduce the number of vehicles that leave the cluster and to select the most stable CH; a traffic weight (TW) metric will apply for each traffic flow. We divided the flow of the neighbours heading in the same direction as vehicle 'a' into three sets: SL: left- turn vehicles,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2732727, IEEE Access

SR: right-turn vehicles and SS: straight-flow vehicles. The number of vehicles in each set is symbolized as VL, VR and VS correspondingly. The total number of neighbouring vehicles with the same direction as vehicle 'a' is denoted as V. The flow in which vehicle 'a' travels must be allocated the maximum weight. The weight of the flow of traffic is calculated as the number of neighbouring vehicles in the flow by the total number of vehicles. Then, the weight for each traffic flow can be calculated as Eq. (1):

Notation	*Description*
F	Describes the flow in which vehicle "a" exists
I	Index of the flow $k \in (L\ R\ S)$
S_R	Right-turn vehicles
S_L	Left-turn vehicles
S_S	Straight-flow vehicles
V	Total no. of neighbouring vehicles
V_F	No. of vehicles in each traffic flow

Table 3: Notations and descriptions

$$TW_aK = \begin{cases} 1 & K = J \\ \dfrac{Vk}{V}, K \neq J, K \in (KRS) & \end{cases} \quad ------(1)$$

For example, take six lanes, as shown in Fig. 2. Lane 1 is the left-hand lane, Lanes 2, 3, 4, 5 are straight-flow lanes, and Lane 6 the right-hand lane. For vehicle i, its communication range covers all the further vehicles in Fig. 2. Then, each traffic flow has the following weight: 2/16, 1 and 3/16 respectively.

Vehicle density

The size of the cluster is unfixed and depends upon the road conditions. The vehicles' density determines the size of the cluster. The vehicle speed varies according to the conditions; if the vehicle speed is low this means the density is high, so the density is inversely proportional to the speed. The size of the cluster is big if the density is high and it is small if the density is low. The vehicle density is shown in the following equations:

Density (vehicles per mile) = Flow (vehicles per hour) / Speed (miles per hour)

ii. Connectivity level

In order to find the total connectivity level (TCL), we have to compute the overall connectivity level and the connectivity level for the flow of traffic. The overall CL, α, shows the average number of vehicles directly associated with vehicle 'a' and is defined as

$$\alpha_a(T) = \sum N(a, b, T)$$

where 'b' is the neighbouring vehicle of vehicle 'a'. N (a, b, T) is equal to 1 if a connection between 'a' and 'b' is present at time 'T' or else equals 0. Now, after calculating the overall CL for a flow of traffic, we have to calculate the connectivity level for a vehicle 'a' and the vehicles in the flow of traffic to which the particular vehicle belongs. This is defined as:

$$\beta_a(T) = \sum N(a, b_{TF}, T)$$

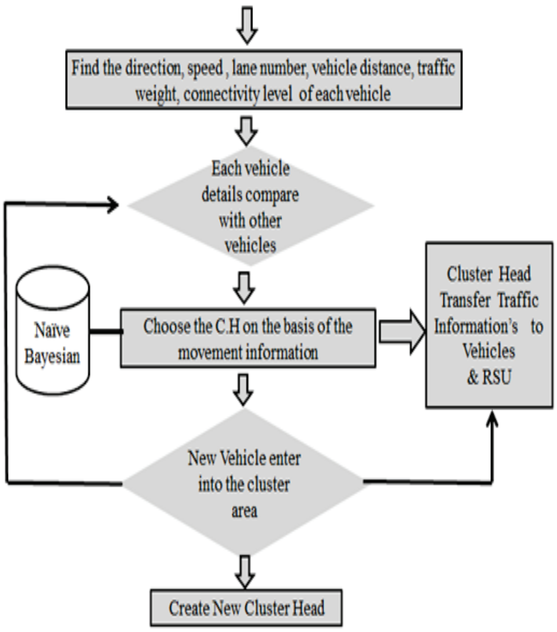

Fig. 4: CH selection process

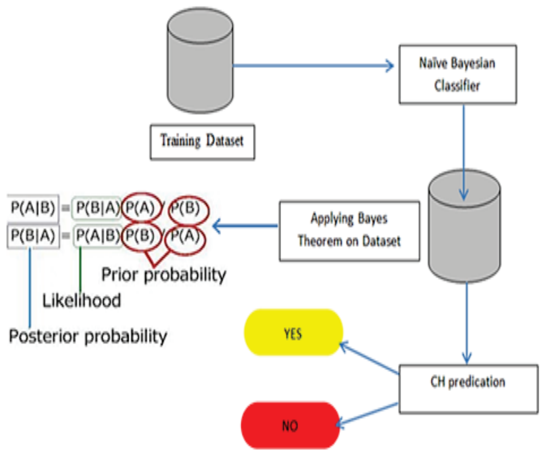

Fig. 5: Bayesian model

b_{TF} is a vehicle in the same flow of traffic as vehicle 'a'. The total connectivity level for vehicle 'a' is now defined as:

$$TCL_a(T) = \alpha_a(T) + \beta_a(T) + TW_a^k$$

iii. Speed and distance

Both parameters are calculated by using their standard formulas.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2732727, IEEE Access

E. Naïve Bayesian Technique

Bayesian networks can represent complex probability distributions well, and have received much consideration in recent years. Bayesian (also named as Belief) networks (BN) are an influential illustration of facts and a reasoning system. The familiar machine learning algorithm, Naïve Bayes is a particular example of a Bayesian network. The Bayesian classification characterizes a supervised learning technique as well as a statistical means of classification. In machine learning and statistics, classification is the problem of recognizing to which of a set of classes a new observation belongs, on the basis of a training set of data. An example would be predicting a given email as "spam" or "non-spam" or, on a given dataset of the weather, being able to predict whether tomorrow we will play or not. In terms of machine learning, classification is counted as an example of supervised learning, i.e. learning where a training set of properly predictable observations is accessible.

- *Bayesian Theorem*

The Naive Bayes algorithm is based on the Bayesian theorem. Assuming training data, A, the posterior probability of a hypothesis B, P (B|A), follows the Bayesian theorem. It uses the knowledge of earlier events to predict future events.

$$P(B|A) = P(A|B).P(B)/P(A)$$

Vehicles	Cluster Size	Speed	Connectivity level (<=300m)	Distance	C.H
Car1	30	High	True	Near	No
Car2	24	High	False	Far	No
Car3	13	Low	True	Normal	Yes
Car4	17	Normal	False	Normal	Yes
Car5	44	High	True	Normal	Yes
Car6	9	High	False	Near	No
Car7	24	Normal	True	Normal	Yes
Car8	33	High	True	Far	No
Car9	22	High	True	Near	Yes
Car10	12	Low	False	Far	No
Car11	21	High	False	Normal	Yes
Car12	11	Normal	True	Near	Yes
Car13	23	Normal	True	Near	Yes
Car14	14	High	False	Far	No

Table 4: Sample date-set for CH selection

Consider the case scenario:

Situation 1: (22, high, true, normal)
Computing results for: (22, high, true, normal)
Compute for P(y): Probability of YES
P(y) = 8 / 14
P (22|y) = 1 / 8
P (high |y) = 3 / 8
P (true |y) = 6 / 8
P (normal |y) = 5 / 8
Result: (8/14)*(1/8)*(3/8)*(6/8)*(5/8) =0.57*0.125*0.375*0.75*0.625 = 2.4
Compute for P (n): Probability of No
P (n) = 6 / 14
P (22|n) = 0 / 6
P (high |n) = 5 / 6
P (true |n) = 2 / 6
P (normal |n) = 0 / 6
Result: (6/14)*(0/6)*(5/6)*(2/6)*(0/6)
= 0.43 * 0* 0.83 * 0.33 * 0 = 1.59
Winning result is: CH = Yes as 2.45 > 1.59

Notations	*Description*
TW_a^k	Traffic weight where vehicle ‘a’ exists
CL	Connectivity level
A	Overall connectivity level
B	CL for vehicle ‘a’ and the vehicles in the flow of traffic
TCL	Total connectivity level
T	Time interval where connection exists between ‘a’ and ‘b’

Table 5: CL notations

IV. SIMULATION RESULTS

We used the NS2 simulator to perform the simulation of the presented clustering technique. We also used SUMO (Simulation of Urban Mobility) to generate the proposed scenario and the flow of traffic. We compare our simulation outcome with the LBC technique proposed in [11] and the TFB technique proposed in [10].

Fig. 6 demonstrates the ordinary lifetime of the cluster head by using the new proposed technique. The outcome illustrates that the lifetime of the cluster head is improved because of applying the proposed algorithm. This is because we are selecting the CH from the heavy traffic flow, thus minimizing the chances of the CH leaving. The range of transmission also affects the stability of the clusters as if more members are in the cluster this in turn reduces the chances of the cluster being dismissed. Thus, we set the transmission range in the middle of the range (150-500), so more members can be part of a particular cluster.

The average numbers of clusters per vehicle of the LBC and TFB techniques are compared in Fig. **7**. The number of clusters per vehicle is reduced by applying our proposed technique. The reason is that we increase the range of transmission, which in turn reduces the number of members that leave the cluster, and more member are able to be part of the cluster, so the cluster lifetime is extended.

Cluster stability means reconfiguration of the cluster because of the continuously changing topology, as the nodes in this network are vehicles that have a high mobility parameter. We should have a good algorithm of clustering for cluster stability that minimizes the number of changes in a cluster by maximizing the proportion of the CH mobility pattern with its CMs. **Figs.** 8, 9, and 10 demonstrate the stability of the clusters as percentages of three different scenarios as LOW flow, MEDIUM flow and HIGH flow along with the vehicles velocity, and compares these with the existing techniques of LBC [11] and TFB [10]. As displayed in the figure, the high flow of traffic has the maximum value of cluster stability among all the scenarios. Also, our technique performs well compared with the old techniques, as also shown by the simulation outcome.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2017.2732727, IEEE Access

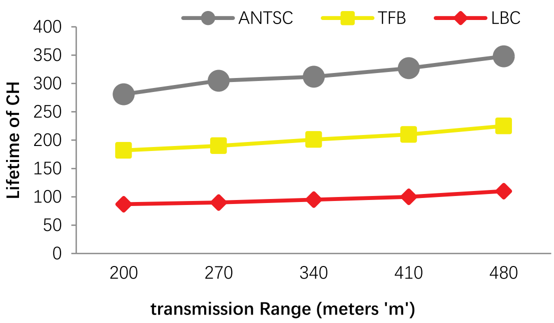

Figure 6: Lifetime of CH with its transmission Range, with ± Vth = 30km/h

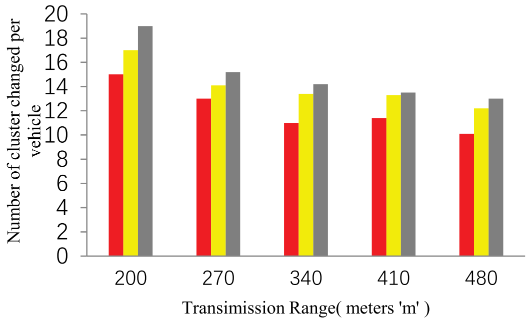

Fig. 7: No of vehicles changed per vehicle along with transmission range, with $\pm \Delta V_{th}$ *=30km/h*

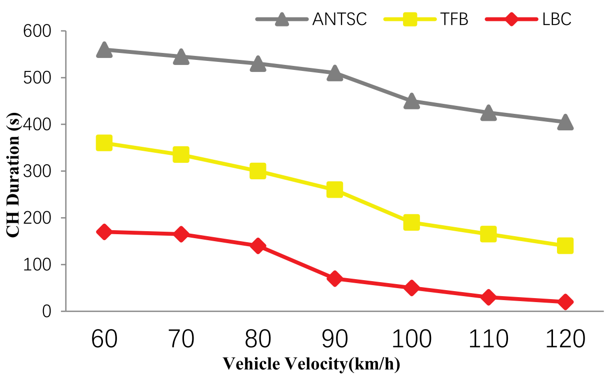

Fig. 8: Low Traffic Flow

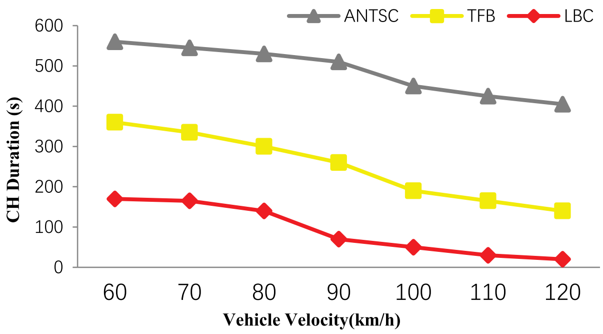

Fig. 9: Medium Traffic Flow

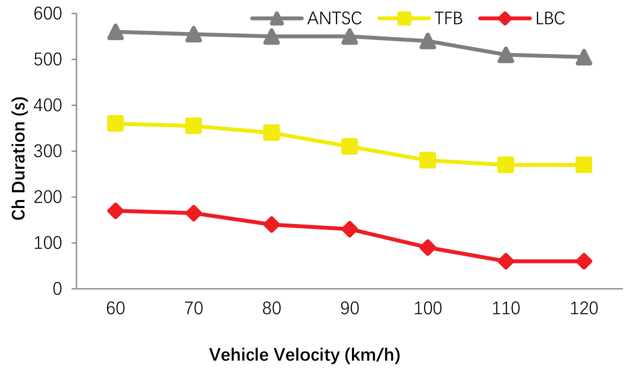

Fig. 10: High Traffic Flow

From Fig. 9, medium flows of traffic also have the average stability of CH as compared to low flows of traffic, nevertheless, our technique has much more positive results for CH variation as demonstrated below.

V. CONCLUSION AND FUTURE WORK

In this paper, we presented a traffic movement based stable clustering algorithm using the Naïve Bayesian method, with the aim of creating stable clusters in urban scenarios. The parameters of the vehicles such as traffic weight, speed, cluster size, connectivity level, direction and vehicle distance are used in our technique for cluster head selection. The cluster head will be selected from the lane having the heaviest traffic flow, aiming to enhance the stability of the cluster as well as the lifetime of the cluster head. The proposed algorithm is compared to the current algorithms and it expressively performs well and increases the stability and lifetime of both the cluster and the CH, as the simulation outcomes also show the performance of the suggested algorithm.

REFERENCES

[1] Mehmood, A., Lloret, J., Noman, M., & Song, H. (2015). Improvement of the Wireless Sensor Network Lifetime using LEACH with Vice-Cluster Head. Adhoc & Sensor Wireless Networks, 28.

[2] Mehmood, A., Lloret, J., & Sendra, S. (2016). A secure and low-energy zone-based wireless sensor networks routing protocol for pollution monitoring. Wireless Communications and Mobile Computing, 16(17), 2869-2883.

[3] Mehmood, A., Umar, M. M., & Song, H. (2016). ICMDS: Secure inter-cluster multiple-key distribution scheme for wireless sensor networks. Ad Hoc Networks. C. Li, M. Ye, G. Chen, J. Wu, An energy-efficient unequal clustering mechanism for wireless sensor networks, 2nd IEEE Int. Conference on Mobile Adhoc and Sensor Systems, Washington, USA, 7 Nov 2005.

[4] Mehmood, A., Nouman, M., Umar, M. M., & Song, H. (2016). ESBL: An Energy-Efficient Scheme by Balancing Load in Group Based WSNs. KSII Transactions on Internet & Information Systems, 10(10).

[5] Ali Çalhan, "A fuzzy logic based clustering strategy for improving vehicular ad-hoc network performance", Sadhana Vol. 40, Part 2, April 2015, pp. 351–367.

[6] H. Saini, R. Mahapatra "Implementation and Performance Analysis of AODV Routing Protocol in VANETs", International Journal of Emerging Science and Engineering (IJESE), ISSN: 2319–6378, Vol. 2, Issue 3, January 2014

[7] Bali, RS, Kumar, N, Rodrigues, JJ. "Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions." Vehicular Communications 2014; 1(3):134–152

[8] M. Mottahedi, S. Jabbehdariand Sepideh, "IBCAV: Intelligent Based Clustering Algorithm in VANET", IJCSI International Journal of

Computer Science Issues, Vol. 10, Issue 1, No 2, January 2013 ISSN (Print): 1694-0784 | ISSN (Online): 1694-0814

[9] S. Gajare, P. Deore, S. Wagh, “Traffic Management in VANET Using Clustering”, International Journal of Engineering and Technical Research (IJETR) ISSN: 2321-0869, Vol. 2, Issue 5, May 2014

[10] F. Yang, Z. Lin, Y. Tang “A Traffic Flow Based Clustering Scheme for VANETs”, Sensors & Transducers, Vol. 180, Issue 10, October 2014, pp. 110-116

[11] M. S. Almalag, M. C. Weigle “Using Traffic Flow for Cluster Formation in Vehicular Ad-hoc Networks”, in Proceedings of the 35th Annual IEEE Conference on Local Computer Networks, 2010, pp. 631 -636.

[12] Z. Y. Rawashdeh and S. Masud Mahmud, “A novel algorithm to form stable clusters in vehicular ad hoc networks on highways”, EURASIP Journal on Wireless Communications and Networking 2012,

[13] S. Vodopivec, M. Hajdinjak, J. Bešter and A. Kos, “Vehicle interconnection metric and clustering protocol for improved connectivity in vehicular ad hoc networks”, EURASIP Journal on Wireless Communications and Networking 2015,2014:170

[14] A. J. Dean, S. N. Brennan, Terrain-based road vehicle localization on multi-lane highways, in Proceedings of the American Control Conference, 2009, pp. 707–712.

[15] M. Jabbour, P. Bonnifait, V. Cherfaoui, Enhanced local maps in a GIS for a precise localisation in urban area, in Proceedings of the 9th IEEE Conference on Intelligent Transportation Systems (ITSC), 2006, pp. 468 - 473.

[16] T. Oskiper, Z. Zhu, S. Samarasekera, R. Kumar, Visual odometry system using multiple stereo cameras and inertial measurement unit, in Proceedings of the Computer Vision and Pattern Recognition, 2007, pp. 1 – 8

[17] H. Reza Arkian, R. Ebrahimi Atani, A. Pourkhalili, S. Kamali “Cluster-based traffic information generalization in Vehicular Ad-hoc Networks”-2015

[18] Y. Mohan Sharma, S. Mukherjee “A Novel Cluster Based Efficient Routing Approach for VANET”, International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)-2013

[19] S.-C. Lo, Y.-J. Lin, and J.-S. Gao, “A Multi-Head Clustering Algorithm in Vehicular Ad Hoc Networks”, International Journal of Computer Theory and Engineering, Vol. 5, No. 2, April 2013

[20] Z. Wang, L. Liu, M.-C Zhou, N. Ansari, 2008, A Position Based Clustering Technique for Ad hoc Inter Vehicle Communication, IEEE Transactions on Systems, Man and Cybernetics Vol. 38, No. 2

[21] Zhang, Z., Boukerche, A., Pazzi, R. A novel multi-hop clustering scheme for vehicular ad-hoc networks. Proceedings of the 9th ACM International Symposium on Mobility Management and Wireless Access, MobiWac ’11, ACM, 2011;19–26.

[22] Kakkasageri, M., Manvi, S. Multiagent driven dynamic clustering of vehicles in VANETs. Journal of Network and Computer Applications 2012; 35(6):1771–1780.

[23] Mehmood, A., & Song, H. (2015). Smart Energy Efficient Hierarchical Data Gathering Protocols for Wireless Sensor Networks. Smart Computing Review, 5(5), 425-462.

[24] L. Huang, J. Wu, F. You, Z. Lv and H. Song, "Cyclist Social Force Model at Unsignalized Intersections With Heterogeneous Traffic," in *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 782-792, April 2017. doi: 10.1109/TII.2016.2597744

[25] W. Li and H. Song, "ART: An Attack-Resistant Trust Management Scheme for Securing Vehicular Ad Hoc Networks," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 4, pp. 960-969, April 2016. doi: 10.1109/TITS.2015.2494017

[26] S. H. Ahmed, S. H. Bouk, M. A. Yaqub, D. Kim, H. Song and J. Lloret, "CODIE: Controlled Data and Interest Evaluation in Vehicular Named Data Networks," in *IEEE Transactions on Vehicular Technology*, vol. 65, no. 6, pp. 3954-3963, June 2016. doi: 10.1109/TVT.2016.2558650

[27] L. Nie, D. Jiang, L. Guo, S. Yu and H. Song, "Traffic Matrix Prediction and Estimation Based on Deep Learning for Data Center Networks," *2016 IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, 2016, pp. 1-6. doi: 10.1109/GLOCOMW.2016.7849067

[28] Blanco, Jorge I., and Houbing Song. "Simulation of Communications and Networking in Vehicular Ad Hoc Networks." In Simulation Technologies in Networking and Communications: Selecting the Best Tool for the Test, pp. 547-570. CRC Press, 2014.

[29] LIN, Yong, and Houbing SONG. "DynaCHINA: Real-time traffic estimation and prediction." IEEE Pervasive Computing 4 (2006): 65.