IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-7 E-ISSN: 2347-2693

Probabilistic B-Tree Based Clustering Algorithm for Vehicular Ad-Hoc Network with Data Aggregation

B. Mukunthan[1*], B. Radha[2], S. Govindaraju[3]

[1*] Dept. of Computer Science, Jairams Arts and Science College, Karur, India
[2] Dept. of Computer Science, Jairams Arts and Science College, Karur, India
[3]Dept. of Computer Science, Sri Ramakrishna College of Arts and Science, Coimbatore, India

[*]Corresponding Author:dr.mukunthan.bmk@gmail.com, Tel.: +918870353613

Available online at: www.ijcseonline.org

Received: 12/Jun/2017, Revised: 25/Jun/2017, Accepted: 20/Jul/2017, Published: 30/Jul/2017

Abstract— Vehicular ad hoc network (VANET) is a kind of ad hoc network, where the wireless communication has been established between the moving vehicles. Recently, the clustering scheme is suggested as an effective solution to handle the fast topology changes of vehicular ad hoc networks. However, the stability of the existing clustering approaches shows poor performances due to highly dynamic scenario of VANET. Thus this paper proposes a probabilistic B-Tree based multi-hop clustering scheme for VANET. A probabilistic density function is computed based on the velocity, speed and acceleration of the vehicles in order to select the cluster head (Designated Node (DN)) and the Backup Designated Node (BDN).The clustering will be performed using the direction of vehicles. A B-tree has been constructed for each cluster and each node will keep and maintain the entire topology structure. Once the DN failed, the BDN will be the DN and new BDN will be selected and the B-tree will be rearranged, where the proposed scheme enables the faster convergence. Furthermore the data aggregation will be performed at the designated nodes to reduce data transmission and that makes the effective bandwidth utilization. The NS2 simulation has been used to evaluate the performance of the proposed scheme and identified that the scheme performs better than the existing clustering schemes in terms of packet delivery ratio, cluster stability, routing overhead.

Keywords— VANET, multi hop cluster, probabilistic density function, B-Tree, data aggregation

I. Introduction

Vehicular ad hoc networks (VANETs) are a kind of ad hoc networks, which provides several types of safety and non-safety applications on roads by enabling communication among the vehicles as well as between the vehicle and the road side unit [1]. Since VANETs are characterized by high mobility of vehicles. Because of high mobility, topology of VANETs changes rapidly; hence, high communication overhead will be the resultant for exchanging the new topology information. Several control mechanism for topology managements and media access have been proposed. One of these mechanisms is forming a hierarchical clustering structure within the VANET. In clustering schemes, the group of nodes will be formed as a cluster and can greatly enhance the network performance [2-11]. It permits the establishment of a virtual communication backbone which guarantees efficient data delivery in VANETs and also limits the consumption of limited resources such as bandwidth. In recent years, the clustering scheme has been well studied in Mobile Ad-hoc Networks (MANETs). Regarding to the VANETs characteristics, such as scale of the network, frequently changes of topology, high speed, etc., the traditional clustering schemes are not an appropriate choice for VANETs [4]. Hence, the new clustering technique should be designed effectively, particularly based on VANET characteristics.

Thus this paper proposes a probabilistic B-Tree based multi-hop clustering scheme for VANET. A probabilistic density function is computed based on the velocity, speed and acceleration of the vehicles in order to select the cluster head (Designated Node (DN)) and the Backup Designated Node (BDN). A B-tree has been constructed for each cluster and each node will keep and maintain the entire topology structure. Furthermore the data aggregation will be performed at the designated nodes to reduce data transmission and that makes the effective bandwidth utilization.

The rest of the paper is organized as follows: Section-II provides the recent related works based on the clustering schemes presented for VANET. The system model has been and the proposed probabilistic B-Tree based multi-hop clustering scheme has been described in the section-III. The results have been discussed in the section-V. Finally, section-V renders the conclusion.

II. RELATED WORK

Numerous clustering schemes for VANET have been presented in the literature. A novel clustering scheme has been proposed in [4] for highway based VANET. This approach tries to enhance the stability of VANET [14] by using the metrics such as speed difference among neighbouring nodes, position, and direction. Multi metric algorithms have been presented to elect the cluster head.

In [5], the author proposes a weighted clustering algorithm for VANET. This scheme uses several metric derived from number of neighbours, distrust value relied upon dynamic communication range and direction of vehicle movement to improve the stability of the cluster and connectivity of the cluster.

In [8], the author proposed a clustering technique namely, VMaSC relied upon choosing the node (cluster head) with the minimum mobility computed as a function of speed difference between the nodes.

Affinity based clustering routing protocol [8] has been proposed for VANET that divides the network into variable sized clusters relied upon the traffic type, infrastructure type, and node speed and affinity propagation technique. This protocol mainly concentrates on the minimization of routing overhead and delay for finding the effective routes for data transmission.

In [9], the author proposed an Agent Learning–based Algorithm for Clustering the VANET. The environments have been learned by the agents and with the learned information it takes the appropriate actions. The learned information will be shared with the neighbours in a collaborative manner. Based on the mobility and learned information, the vehicles have been clustered in the network.

In [10], the author proposes a traffic flow based (TFB) clustering scheme to form a stable clusters for VANET. The node position, speed difference and traffic flow has been considered in this scheme to choose the stable cluster head. Furthermore, only vehicles present inside the predicted speed difference between it and the cluster head is permitted to join the cluster.

A clustering technique has been proposed in [11] for large scale VANET. The structure of the cluster has been determined by the geographic position of the nodes and the vehicle traffic information. The size of the cluster has been controlled by a predefined distance between the vehicle and the cluster head. The clusters have been dynamically reconfigured whenever the nodes move out of the cluster.

Some of the objectives of clustering can be achieved using advanced neural network algorithms [3], which ensures the importance of computational methods [12] [7] such as Neural-Fuzzy mapping that are much cheaper and faster than conventional experimental methods. Also the role of wireless sensor networks is immensely needed in various areas of our daily life [13].

III. METHODOLOGY

a) System Model

A system model comprises of n number of vehicles and every vehicle is equipped with On-Board Unit (OBU). The OBU is composed of processing unit, sensors and storage capacity. Each vehicle can communicate with other vehicles as well as with the Road Side Unit (RSU). The Road side unit gathers the information such as crash, collision and so on from its surroundings and forwards this information to other vehicles. However some regions such as like highways, rural areas, the RSU may not accessible, so the vehicle itself transmits the information to each other.

Moreover, it has been assumed that all vehicles are equipped with a positioning system such as GPS from which it can obtain data about its present location. In this work, the vehicles are formed as a B-Tree based multi-hop clustered topology in each direction of the road side. The cluster members only communicate with their respective root vehicle, while the root vehicles communicate with all the cluster members' present in the cluster.

b) Probabilistic B-Tree based Multi-hop Clustering Scheme

In this section, the description of the cluster formation, cluster head selection and the data aggregation phase in the proposed Probabilistic B-Tree based Multi-hop Clustering Scheme (PBMCS) has been given.

i) Cluster Formation

Initially, the nodes (vehicles) are in the Undecided state. In order to partition the network into clusters, each node need exchange their Cluster-Head Candidate (CHC) state information by broadcasting a HELLO message containing all the essential information such as direction, position, speed, velocity and acceleration to its One-hop and two hop neighbours. In order to construct stable clusters, every vehicle needs to consider the neighbours moving in the same direction and pay no attention to broadcasts from vehicles in the opposite direction. Let us consider a static position (x_O, y_O), which is to be meeting point of the road sides and a vehicle node N. The position of the vehicle be (x_n, y_n), velocity of the vehicle be V_n, direction vector (dx, dy). A vector v signifying the direction of vehicle N to the road end can be computed as follows

The vector signifying the velocity of vehicle N can be computed as follows

$$v = (x_N - x_o, y_N - y_o) \qquad (1)$$

Let x be the direction value of the vehicle, which is to be similar with respect to the fixed position and it can be computed as follows

$$\vec{V_N} = V_N dx + V_N dy \qquad (2)$$

The vehicles which haves the comparable values of x will be clustered together in to one clusters.

$$x = \vec{v}_N . v \quad (3)$$

ii) Cluster Head selection

After the cluster formation, the cluster head and the backup node have to be elected in the PBMCS. The probability density function (pdf) has to be calculated based on the speed and distance between the node and the road segment (fixed position). The current location of the vehicle will be measured using the location information provided by the GPS; the vehicle computes the time to live that means the residual time to cross the road segment and it has been calculated as follows

$$TTL = \frac{D}{Speed} \quad (4)$$

Where D is the distance between the node and the static position, speed is the speed of the node

The node stability S has been calculated based on the TTL, velocity v and the acceleration a of the node

$$S = \frac{TTL}{a} \quad (5)$$

The probability density function has been calculated for each node based on the S value.

$$pdf = \frac{1}{\sigma\sqrt{2\pi}} e^{\left(-\frac{1}{2}\right)\left[\frac{s-\mu}{\sigma}\right]^2} \quad (6)$$

Where μ is the mean of the S value obtained from the cluster members, is the standard deviation respected to the mean. The node with highest *pdf* will be elected as a cluster head (Designated Node (DN) and the node with the next maximum will be elected as a Backup Designated Node (BDN). The DN will always send beacon to BDN to detect the failure of DN. Once the DN failed the BDN will be the DN and new BDN will be selected.

The nodes which are not a 2 hop distance of the cluster head then that node are isolated from the cluster and they will form a new cluster and elect the cluster head and the Backup Designated Node (BDN) for that cluster.

The Binary tree (B-tree) has been constructed for each cluster to maintain the topology structure of the network. The cluster members and the cluster head will find the distance between its position and the static position. The cluster head will be keep it as a root and the node which has the distance minimum than the cluster head distance will be placed in the left of the tree, while the node which has the distance maximum than the cluster head will be placed in the right of the tree.

The algorithm for constructing the B-tree has been given in the algorithm 1.The entire topology will be formed by the gateway nodes of the clusters. The each node will keep and maintain the entire topology structure. This will maintain the consistency in communication by avoiding flood. Moreover, the CH will always send beacon to BDN to detect the failure of DN. Once the DN failed, the BDN will be the next DN and new BDN will be selected. This will enable faster convergence, if a node go away from the cluster the beacon signal will identify that and the B-tree will be rearranged. This will avoid the broadcast storm and performance issues. Transmission failure will be minimized due to alternate route backup.

In this work, Histogram based Data aggregation has been used to minimize the bandwidth consumption, where the data interval velocity is divided by an error bound ε; then, the data packet has been kept into the related interval. The data packet in one interval is aggregated data. The Data aggregation required to consider the division of the data zone and the grouping of the data. Relevant details should be given including experimental design and the technique (s) used along with appropriate statistical methods used clearly along with the year of experimentation (field and laboratory).

Algorithm 1 B-Tree construction for Clusters

Step 1:For i=1 to N, where n is the number of clusters
Step 2:Calculate the distance between the cluster head and the static position
Step 3:Place the cluster head as root
Step 4:For i=1 to m, where m is the number of cluster members
Step 5:Calculate the distance between node and the static position
Step 6:If
Step 7: If left child →NULL
Step 8: place the node in the Left hand side
Step 9: else
Step 10: Compare the distance with the occupied node as root and do the process from the step 6
Step 11: end if
Step 12:else If $D_{CM} > D_R$
Step 13: If right child→NULL
Step 14:place the node in the Right hand side
Step 15: else
Step 16: Compare the distance with the occupied node as root and do the process from the step 6
Step 17: end if
Step 18: end for

Table 1 Simulation setup

Parameters	Values
Simulation type	NS2
Network size	1000×1000
Number of nodes	20, 40, 60, 80, 100, 120
Speed	20-70 kmph
Physical standard	802.11/802.11p
Transmission range	200 m
Packet size	512 bytes
Data type	CBR

IV. Results And Discussion

The NS2 simulation is used to evaluate the performance of the proposed Probabilistic B-Tree based Multi-hop Clustering Scheme (PBMCS). Table 1 shows the Simulation setup. The proposed protocol is compared with the existing scheme Agent Learning–based Algorithm for Clustering (ALCA), VWCA. The metric used to evaluate these clustering schemes are packet delivery ratio, cluster stability, End to end delay and routing overhead.

A. Performance metrics

Packet delivery ratio is the fraction of CBR packets given to the destination to those disseminated by the sources. It is computed by the number of received packets by destination through the number of packet generated from the source.

TP = The Total Number of packets successfully delivered.

$$PDR = \frac{TP}{\text{Total Number of packet sent}} X\,100 \quad (7)$$

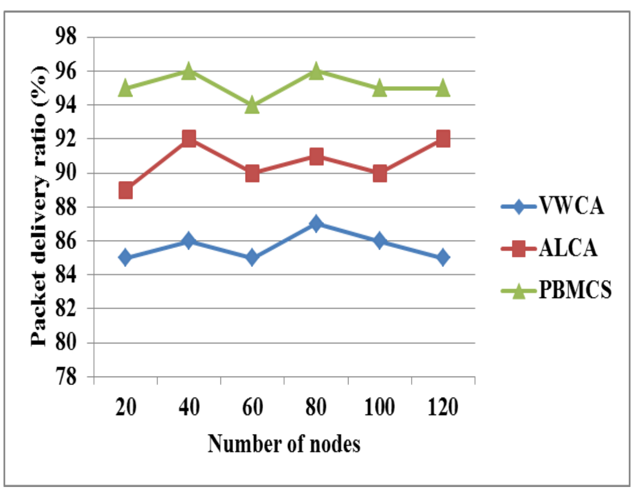

Figure 1 Packet delivery ratio with respect to number of nodes

Figure 1 shows the packet delivery ratio with respect to the number of nodes in the network. The proposed PBMCS attains the maximum packet delivery ratio with the help of the topological information using the B-tree. Moreover, by performing the data aggregations, the congestion will be less in the network. So the packet drop due to the congestion will be less in the proposed PBMCS, where it increases the packet delivery ration when compared to the existing VWCA and the ALCA.

In order to measure the cluster stability, cluster head duration need to be computed. Cluster head duration is defined as the time duration from when a node changes its state to CH to when a node transfers from state CH to or CM. Figure 2 shows the average cluster head change with respect to the maximum speed of the vehicle. The PBMCS select the most stable cluster head by using the probability density factor. The PBMCS acquires 12 cluster head changes for 70 m/s maximum speed, while, the VWCA, ALCA acquires, 19, 14 cluster head changes respectively.

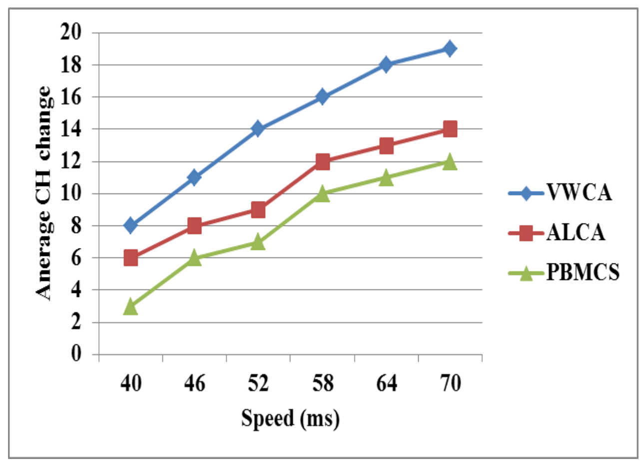

Figure 2 Cluster stability with respect to maximum speed of the vehicle

Routing overhead is the ratio between the total number of created control packets and the time interval. With the help of the topological structure, the PBMCS can generate a minimum number of control packets. Moreover with the BDN, the PBMCS can minimize the transmission error and thereby the control packet generated for the retransmission will be minimized. The PBMCS generates 150 control packets for 300 s, while VWCA, ALCA generates 260, 350 control packets respectively.

It should include important findings discussed briefly. Wherever necessary, elaborate on the tables and figures without repeating their contents. Interpret the findings in view of the results obtained in this and in past studies on this topic. State the conclusions in a few sentences at the end of the paper. However, valid colored photographs can also be published.

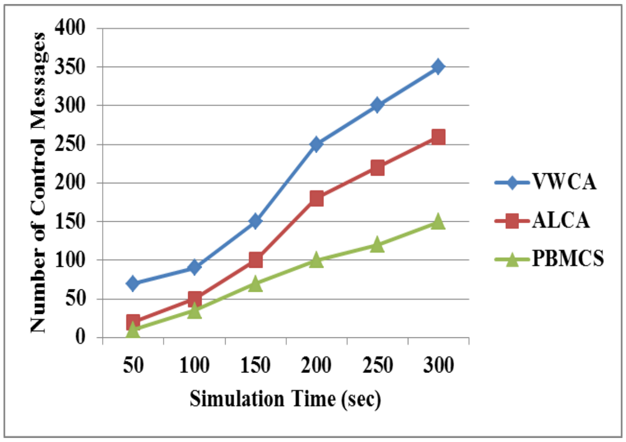

Figure 3 Routing overhead with respect to the simulation time

V. CONCLUSIONAND FUTURE SCOPE

This paper proposes a probabilistic B-Tree based multi-hop clustering scheme (PBMCS) for highly dynamic scenario of VANET. The PBMCS forms the most stable clusters heads ((Designated Node (DN)) and the Backup Designated Node (BDN) by computing the probability density function based on the velocity, speed and acceleration of the vehicles. A B-tree has been constructed for each cluster and each node will keep and maintain the entire topology structure. Once the DN failed, the BDN will be the DN and new BDN will be selected and the B-tree will be rearranged, where the proposed scheme enables the faster convergence.

The data aggregation using the Histogram has been performed at the cluster head to reduce data transmission and that makes the effective bandwidth utilization. The NS2 simulation has been used to evaluate the performance of the proposed scheme and identified that the scheme performs better than the existing clustering schemes VWCA, ALCA in terms of packet delivery ratio, cluster stability, routing overhead. In future the performance of the above work can be improvised by applying suitable neural network algorithms from artificial neural networks.

ACKNOWLEDGEMENT

I owe my sincere thanks to my friends Dr.N.Nagaveni and Dr.A.Pushpalatha for their support in this Research work. I also extend my sincere gratitude to Mrs.B.Radha, for her support in preparing this paper.

REFERENCES

[1]. A. F. Santamaria, C. Sottile, and P. Fazio, *"PAMTree: Partitioned Multicast Tree Protocol for Efficient Data Dissemination in a VANET Environment,"* International Journal of Distributed Sensor Networks, vol. 11, no. 5, p. 431492, Jan. 2015.

[2]. Mohamed Hadded, RachidZagrouba, AnisLaouiti, Paul Muhlethalery and Leila AzouzSaidane, *"A Multi-Objectif Genetic Algorithm-Based Adaptive Weighted Clustering Protocol in VANET"*, Congress on Evolutionary Computation (CEC), IEEE, Sendai, pp. 994-1002.

[3]. Mukunthan B, "*A Neural Network Approach for Precise Pattern Identification of Human DNA*", International Journal of Neural Networks and Applications, Vol. 5, Issue 1, pp.21-27, 2012.

[4]. Zaydoun Y Rawashdeh and Syed Masud Mahmud, *"A novel algorithm to form stable clusters in vehicular ad hoc networks on highways"*, EURASIP Journal on Wireless Communications and Networking, Vol. 1, Issue. 15, 2012, doi.org /10.1186 /1687-1499-2012-15.

[5]. Daeinabi A, Rahbar AGP, Khademzadeh, *"A. VWCA: An efficient clustering algorithm in vehicular ad hoc networks"*, Journal of Network and Computer Applications, Vol. 34, Issue. 1, pp.207–222, 2011.

[6]. Obulapu Hiteshreddy, Pardeep Singh and Siddharth Chahuan, "*A Review on Cluster Based Data Aggregation Protocols in Wireless Sensor Network*", International Journal of Computer Sciences and Engineering, Vol.3, Issue.8, pp.37-45, 2015.

[7].Mukunthan. B and Pushpalatha. A, *"Automation of DNA Finger Printing for Precise Pattern Identification using Neural-Fuzzy Mapping Approach"*, International Journal of Computer Applications, Vol. 13, Issue. 3, pp.16-24, 2011.

[8]. Atul Rana, Manju Bala and Varsha , "*Review Paper on MSEEC: Energy Efficient Clustering Protocol in HWSN*", International Journal of Computer Sciences and Engineering, Vol.4, Issue.5, pp.71-75, 2016.

[9]. Neeraj Kumar, Naveen Chilamkurti and Jong Hyuk Park, *"ALCA: agent learningbased clustering algorithm in vehicular ad hoc networks"*, International journal on Personal and Ubiquitous Computing, Vol. 17, Issue .8, pp. 1683-1692, 2013.

[10]. Fan Yang, Zhijian Lin, Yuliang Tang, 2014, *"A Traffic Flow Based Clustering Scheme for VANETs", Sensors & Transducers*, Vol. 180, Issue 10, pp. 110-116, 2014.

[11]. Zhigang Wang, Lichuan Liu, MengChu Zhou, and Nirwan Ansari, 2008, *"A Position-Based Clustering Technique for Ad Hoc Intervehicle Communication"*, Transactions On Systems, Man, And Cybernetics, IEEE, Vol. 38, Issue .2,pp.201-208,2008.

[12] Mukunthan. B and Nagaveni. N, *"Identification of Unique Repeated Patterns, Location of Mutation in DNA Finger Printing Using Artificial Intelligence Technique"*, International Journal of Bioinformatics Research and Applications, Vol. 10, Issue. 2, pp. 157-176, 2014, doi.org: 10.1504/IJBRA.2014.059516.

[13] Manish B.Giri, Ravi Singh Pippal, "*Agricultural Environmental Sensing Application Using Wireless Sensor Network forAutomated Drip Irigation"*, International Journal of Computer Sciences and Engineering, vol.4, Issue .7,pp. 133-137, 2016.

[14] Krishnakumar K.G, Dr.B.Mukunthan, "*Cross Layer Based Adaptive Routing Approach for VANET*", International Journal of Control Theory and Applications, vol .9, Issue 40, pp. 161-169, 2016

AUTHORS PROFILE

Dr.B.Mukunthan Ph.D pursed Bachelor of Science(Computer Science) from Bharathiar University, India in 2004 and Master of Computer Applications from Bharathiar University in year 2007 and Ph.D from Anna University-Chennai on 2013,He is currently working as Research Advisor in Department of Computer Science, Bharathidasan University, Tiruchirapalli since 2016. He is a member of IEEE & IEEE computer society since 2009, a life member of the MISTE since 2010. He has published more than 10 research papers in reputed international journals including Thomson Reuters (SCI & Web of Science) .He is also Microsoft Certified Solution Developer. His main research work focuses on Algorithms,Bioinformatics,Big Data Analytics, Data Mining, IoT and Neural Networks. He also invented a Novel and Efficient online Bioinformatics Tool and filed for patent. He has 11 years of teaching experience and 9 years of Research Experience.

Mrs B.Radha pursed Bachelor of Commerce and Master of Computer Application from Bharathiar University, India in year 2009. She is currently pursuing M.Phil. and member of f IEEE & IEEE computer society since 2008, a life member of the MISTE since 2012. She has published many research papers in reputed international journals including Thomson Reuters (SCI & Web of Science) and conferences. Her main research work focuses on Data Mining, Big Data, Algorithms, Network Security, Cloud Security and Privacy, IoT and Computational Biology based education. She has 6 years of teaching experience and 1 year of Research Experience.

Mr S Govindaraju MCA, MPhil pursed Bachelor of Commerce, in the year 2002 and Master of Computer Application from Bharathiar University, Coimbatore in the year 2005 and completed Mphil in Computer Science in the year 2011. He is pursuing Ph.D in Bharathiar University , Coimbatore since 2014, and currently working as an Assistant Professor in School of Computing, Department of Computer Science, Bharathiar University, Coimbatore since 2014. He has published more than 5 research papers in reputed international journals including Thomson Reuters (SCOPUS & Web of Science) and conferences and it’s also available online. His main research work focuses on Image Retrieval using Medical Images. He has 10 years of teaching experience and 9 years of Research Experience

