

# Performance enhancement of in-network probabilistic aggregation (PEPA) for vehicular ad-hoc networks

Rajeev Kumar<sup>1,2,\*</sup>, Dilip Kumar<sup>3</sup>  
and Dinesh Kumar<sup>4</sup>

<sup>1</sup>Department of Electronics and Communication Engineering, IKG-Punjab Technical University, Kapurthala, India.

<sup>2</sup>School of Electronics and Electrical Engineering, Lovely Professional University, Phagwara, India.

<sup>3</sup>Department of Electronics and Communication Engineering, Sant Longowal Institute of Engineering and Technology, Longowal, India.

<sup>4</sup>Department of Information Technology, DAV Institute of Engineering and Technology, Jalandhar, India.

\*E-mail: rajeev.patil@gmail.com

This paper was edited by Ashutosh Sharma.

Received for publication September 8, 2020.

## Abstract

In the current scenario of vehicular communication, it is very difficult to believe that the aggregated packet of information is not malicious. The algorithm is developed to allocate probability to the packet transmitted by the cluster head depending upon, the number of vehicles involved in creating a clustered information. The aggregated packet with the highest value of probabilistic correctness is considered by the receiving-end cluster head when multiple packets of similar information with related timestamps are received. The performance of the technique is tested under varying vehicular densities and data-sending rates. This scheme provides robust aggregation performance in comparison to the existing state-of-the-art structured and structure-free aggregation techniques.

## Keywords

Aggregation, Broadcast storm, Information dissemination, MANET, Structured aggregation, Structure-free aggregation, VANET.

Structure-free vehicular ad-hoc networks (VANETs) exchange information through roadside units (RSUs) using IEEE 802.11p at 5.9GHz. To make the journey experience safe and secure for the user, we need to ensure real-time information to road users on the highway or city scenario (S Al-Sultan et al., 2014; Schoch et al., 2008). Figure 1 indicates the basic communication module used for the propagation of information of the cluster to the end user. Cluster is required to overcome the issue of message storm in VANETs (Bali et al., 2014; Dietzel et al., 2014; Kumar et al., 2019). There exist many routing protocols in mobile ad-hoc networks, but they are not directly applicable to VANETs (Sharma and Kumar, 2016).

Researchers contributed to the area of structured aggregation (Bilal et al., 2014; Wischhof et al., 2003) but it does not find its scope in real-life situations. Smart data aggregation method aggregates on the

basis of vehicle direction, road speed limitation, and duplicate packet removal, but the mechanism does not support effective utilization of bandwidth (Allani et al., 2018, 2020). Data dissemination based on map splitting is created by building zones of relevance and extracting related data, but the basic issue of broadcast storm is not addressed (Allani et al., 2018). There were contributions in the area of structure-free aggregation, wherein there is a dynamic mechanism to propagate packets and make information available to the users through improved forwarding delay (Molina-Gil et al., 2014; Ibrahim and Weigle, 2008; Kumar and Dave, 2013). The information-centric networks are being replaced to ensure efficient communication in terms of efficient delivery and address the trust management issues of the network (Rathee et al., 2020). Another mechanism where a trust evaluation matrix is

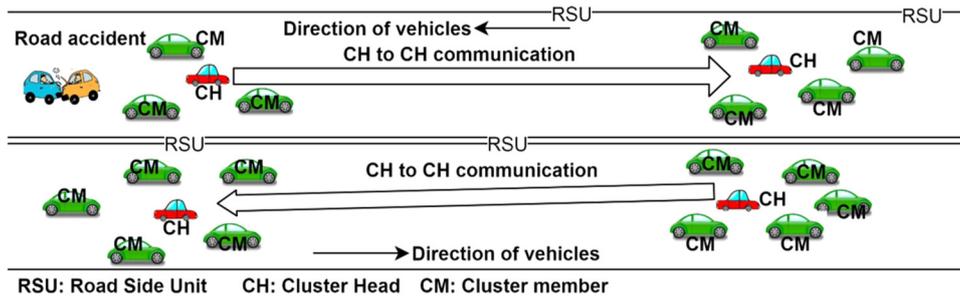


Figure 1: Basic clustering communication model in VANET.

created to overcome denial of service attacks on the VANETs, addresses bandwidth optimization without compromising security (Poongodi et al., 2019). Still, there is a scope to provide probabilistic correctness of the aggregated packet. Moreover, the authenticity of the aggregated packet remains an issue that needs urgent addressing. Once an aggregate of the original information is created, it is difficult to check whether the packet created by the cluster head is undoubtedly the correct one. Hence, whenever aggregated information is created, it must be ensured for its probable correctness, and the validity of the cluster member to send information to the cluster head needs to be checked at the time of the creation of an aggregated packet and not later.

### Our contribution

A probabilistic aggregation scheme in which correctness of the aggregated packet is proposed, here, a new parameter, called the probability of correctness of the information aggregated, is created. Now, the receiving cluster head is able to make a decision, whether the packet is to be relayed further or not, depending upon the value of the probability of correctness. Moreover, we are not interested to keep information propagating further for more than twice the range of communication in VANET, since older information is not useful to make any decision. At the same time, the value of the timer that helps in regeneration of the aggregated packet is fixed. This helps in regeneration of the packet in case nearby vehicles are not available. The timer helped in reducing the packet loss during propagation of information. PEPA helps in making a correct decision on whether to keep using the aggregated packet or drop it and at the same time improves average delay, packet delivery ratio, and aggregation accuracy in comparison to other available state-of-the-art techniques.

### Related work

There are several aggregation schemes proposed in VANETs for fixed structures of road segments. As the distance from the segment increases, the accuracy of these packets is also reduced and applicability of the techniques is questioned. Moreover, it is not correctly measuring the situational information. To overcome this issue, structure-free aggregation schemes are developed using various methods such as semantic or syntactic.

Analysis of vehicular-density estimation through a fixed-length structure is not an accurate estimate even though all vehicles are in the communication range. If we take two continuous road segments, then we will find some area of the road without being covered in any road segment (Bilal et al., 2014). A fixed hierarchy traffic information system was developed based on average speed as a parameter (Wischhof et al., 2003). Here each vehicle is supposed to be equipped with a Global Positioning System (GPS) receiver and other related digital techniques that help in data processing. Data from the nearby nodes that have lots of similarities will be aggregated into one during aggregation in case it is a set of similar information (Nadeem et al., 2004). M. Caliskan et al. applied a road-segmentation approach to compare with different segments of parking spots and reduced the record size to 50 bytes per second (Caliskan et al., 2006). Moreover, this approach of hierarchical aggregation is as good as a tree-based approach. If the distance of information from the source is more, then the quality of information is doubtful. Lochert et al., (2008) implemented the usage of landmarks instead of quadtree structures. It is pointed out that the difficulty of implementing parking space applications is too large in comparison to building a traffic information system where the applicability of fixed hierarchy is questionable. Exchange of communication at

a minimum overhead is implemented by Sleet et al. (2010) using a clustering scheme for VANETs. Lee et al. (2014) reduced redundant information in urban and highway scenarios. The identity of vehicles is ensured through the cluster code. Multihop aggregation takes into consideration a larger area over which data is to be disseminated (Ucar et al., 2014). The issue of disconnected networks and broadcast storms is addressed by Ucar et al. (2016) at different vehicle densities. The authors were left to address the issue of multihop stability in an urban scenario.

Lochert et al. (2007, 2010) overcome the problem of scalability using (Flajolet–Martin) FM sketch-based aggregation in VANETS. If distinct elements are calculated in large databases without data being passed multiple times, FM sketches are being used. This helps in removing the older information. Here a time stamp is considered and the Time-to-Live (TTL) parameter is used (Flajolet and Martin, 1985). Consideration of the TTL parameter is not a practical solution. The price for this feature is the loss of exact counting. Zekri et al. (2012) also rely upon duplicate insensitive data fusion using FM sketches. The goal is not to save bandwidth but to get the exact knowledge of previous information. However, work does not use the TTL-adapted parameter and leads to high storage requirements. Many traffic information-related components and variables are made available through the Kalman filter (Milojevic and Rakocevic, 2015) and these variables varied from application to application. It helped in identifying single observations and removing old aggregations in real time. Usage of clustering for the exchange of beacons and optimization in the last 2-hop route is conducted to achieve better results (Ji et al., 2016). It is come up that the major problem in VANET is to restrict the flow of several packets that leads to a broadcast storm. This problem can be overcome using clustering to create clusters and data aggregation. This work is also useful in the area of signal disruption in optical fiber communication (Kumar and Sharma\*, 2020)

## Proposed approach

The goal is to create a system for generating an aggregated packet of information that contains a parameter with probable correctness of the packet. Such a system makes the existing aggregation scheme more adaptable in real-life processing of packets as shown in Figure 2.

The detailed working of Figure 2 is illustrated through a flow diagram of Figure 3. The authorization phase involves only those nodes in the network that are authorized by the certificate authority (CA). This phase is accomplished by authenticating the offline-registered nodes through the RSU. The RSU is responsible for examining the authenticity of the vehicle. If the vehicle proves to be authorized by CA, the vehicle is allowed to participate in the network. Otherwise, the vehicle is not certified to join the operating network. First, each vehicle is required to perform registration with the CA offline and then to get authorization from RSU in order to get operated within the network.

The second phase of the proposed mechanism involves the clustering of nodes. To overcome problems of power consumption where consideration of extra battery level or the cost limitations are very much mitigated by selecting the group heads from the authenticated vehicles. In this phase, all authorized nodes are required to form the cluster based on the information registered in the database of the CA. Moreover, cluster heads are selected based on average speed, trust factor, neighborhood degree, and time-to-live parameters (Kumar et al., 2019). The nodes with the higher value of cluster head factor and high-performance results from previous communications are selected for clustering. There is at least one group head for each cluster. The formation of these clusters marks an advantageous feature of decreasing the burden on the RSU.

The next phase of the process is to collect the information packets from the cluster nodes and create an aggregated packet at the cluster head level. Here, a parameter probability of correctness

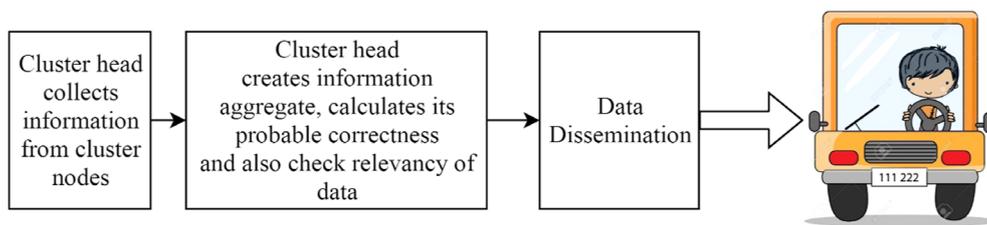


Figure 2: Basic model for information flow.

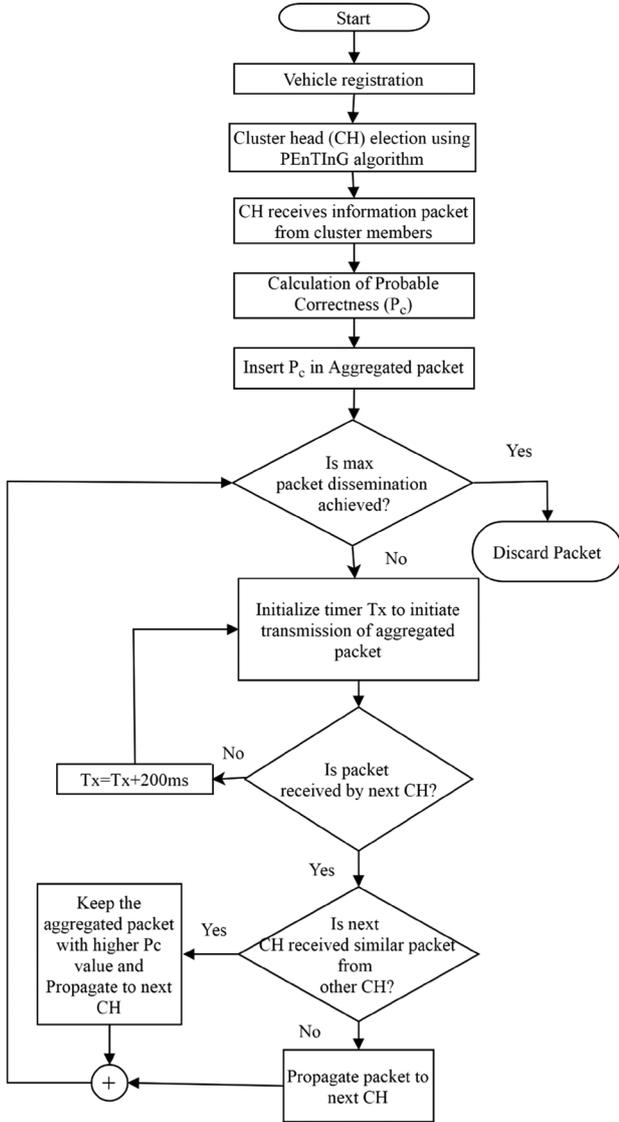


Figure 3: Data aggregation and dissemination process.

is introduced to inform other cluster heads about the usefulness of the information packet. The desired value of this parameter should be as close to one as possible. To guarantee the correctness of an aggregated packet, the scheme verified the signature of each contributing vehicle as part of the cluster, and if any vehicle is found using wrong signatures, then the probability of correctness will be reduced, otherwise, we strive to get the probability of correctness as close to one as possible. In case eight vehicles are contributing to the aggregate, then, if a vehicle is found with the correct signature, probable correctness will be calculated as 1/8; in case one vehicle does not have

the correct signature, then net probable correctness will be (1-1/8); this parameter will apply to both cities and highways. The packet forwarded by the cluster head contains the field of the number of vehicles and probable correctness parameters. This will help the user to utilize the information in case the user is willing to do so, thereby decision-making becomes easier.

In case  $n$  be the number of vehicles involved in the creation of a packet, then the probability of correctness of each vehicle is given as  $\frac{1}{n}$ . If  $k$  is the number of vehicles with incorrect signatures, then, the total probable correctness  $P_c$  of the aggregated packet is given as the algebraic sum of all probabilities of all correct signature vehicles, hence

$$P_c = \left( \frac{n-k}{n} \right). \quad (1)$$

Once a packet is disseminated to the next cluster head with correctness information as part of the aggregated packet, then it happens that the next cluster head might have received similar packets from other existing cluster heads also, so the next CH has to decide which packet to keep and which one to drop. This decision is taken based on the probable correctness level of the received packet. In case multiple packets of similar information are received, then higher the value of probable correctness, the larger is the probability of selection of that aggregated packet in order to disseminate the information further. This dissemination of the information is also made probabilistic: transmission of the aggregated packet will be less in case the distance of the aggregate from the source cluster head is increased. This follows equations (2) and (3) as given below.

In case  $D$  is a random variable, indicating the distance from source vehicle  $S$  and in case  $d$  denotes the distance from source aggregator vehicle to the current location of the aggregated packet with some cluster head, then:

$$P[\{D \leq d\}] = \frac{d}{R}, \quad (2)$$

where  $R$  is the range of communication using IEEE802.11p.

$T_x$  is the timer value related to the vehicle with distance  $D$  (Kumar and Dave, 2013), then:

$$T_x = T_{max} \left[ 1 - \left( \frac{d}{R} \right)^\epsilon \right], \quad (3)$$

where  $\epsilon > 0$  and  $T_{max} = 200ms$

The aggregated packet consists of the average speed of the cluster, location of origin of the packet, information, time stamp, and correctness value as part of it. The aggregated packet will not be rebroadcasted further in case the maximum propagation distance is reached. Moreover, in case distance from the source of the packet is increased, then several transmissions will also be reduced, or in other words, this is made probabilistic. Equations (2) and (3) put a limit on the number of retransmissions of the aggregated packet. It is found that older information is not useful in VANETs, and hence as and when packets become older, they are no more needed in the communication process and communication is continued with fresh packets. The loop ends when the maximum distance of dissemination is achieved for an aggregated packet, or the aggregated packet becomes older. Now, fresh aggregated packets take part in the communication process and the older ones keep on getting eliminated from the dissemination process. The aggregated packet dissemination remains in the communication process as long as the maximum dissemination range is not reached, i.e., twice the range of communication. The packet gets discarded if the condition becomes true, and hence the loop ends if there are no packets to disseminate further.

## Simulations and results

This section evaluates the performance of PEPA through simulations. The prime objective of the data aggregation process is to save bandwidth. To analyze the performance of the probabilistic aggregation in terms of packet delivery ratio, average delay, and aggregation accuracy, several NS 2 simulations are conducted and Table 1 is used to carry out the simulation process. This section presented a detailed simulation analysis of the proposal for a road of a length of 2.5km. There are six lanes considered for simulation, with an equal number of lanes in the forward and backward direction.

The performance of the PEPA scheme is analyzed under different vehicle traffic densities (vehicles/m) and data-sending rates (packets/sec). Other techniques used to present the comparative analysis of the results are DRDCDA, KBPDA, CASCADE, and IVG (Ibrahim and Weigle, 2008; Kumar and Dave, 2013, 2016; Benslimane and Bachir, 2003).

The graphical downward trends indicated in Figure 4 clearly show that PEPA requires less number of retransmissions in comparison to the existing mechanisms and hence, saves bandwidth. This means we are able to create a perfect transfer of

**Table 1. Simulation parameters.**

Parameter	Value or type
Simulation area	2500m × 2500m
Simulation time	600s
Radio model	Two-ray ground
MAC layer	IEEE802.11
Number of vehicles	50–400
Vehicle speed	40–120km/h
Constant bit rate (CBR rate)	1–4 packet/s
Vehicle communication range	500m
Bandwidth	20Mbps
Data packet size	500–2,500 byte
Cache replacement policy	FIFO
Mobility model	Manhattan
Traffic environment	Urban
Number of lanes	4–8
Traffic type	Multimodal
Routing protocol	AODV
Number of simulation runs	10

aggregated packets from source to destination and also most of the packets are getting delivered. This is more so because propagation of the packets here is unidirectional and we are not using another side of the lane for one-side communication.

The proposed aggregation scheme controls the number of retransmissions through clustering and the Tx parameter. The performance of aggregation is less significant when traffic density is low, since the vehicle may not find a nearby cluster head to share the information, and hence the aggregated packet may stay with the cluster head for long as indicated in Figure 5. It is said that as the number of vehicles (vehicle density) are increased, the process of aggregated packet dissemination is not hampered or the aggregated packet does not have to wait for the nearby vehicle in order to complete the data transfer and hence this helps the communication process with more packets reaching the destination.

Figure 6 indicates that the proposed aggregation scheme achieves a good packet delivery ratio (PDR) even though the data-sending rate is increased. As the number of data packets per second increases, there is a slight decrease in PDR but its performance is

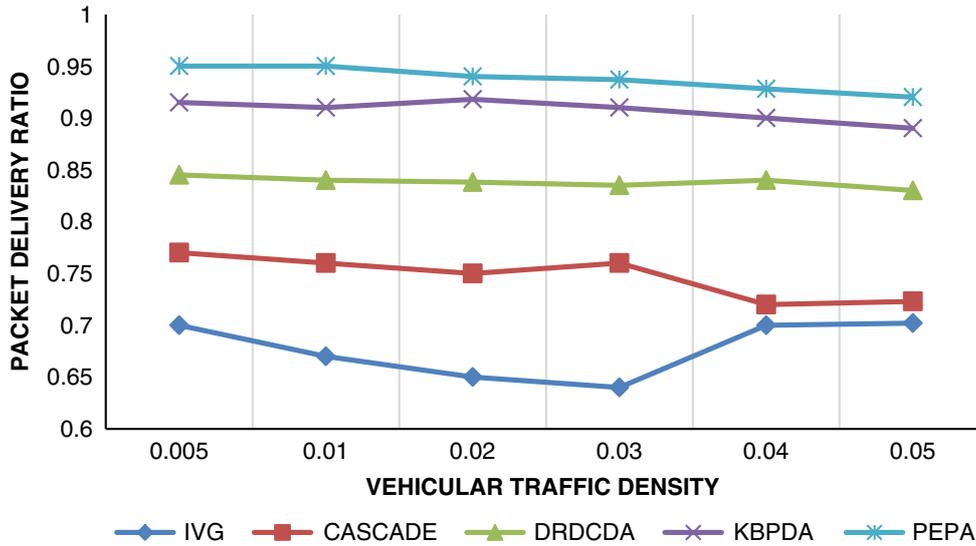


Figure 4: Packet delivery ratio of PEPA, KBPDA, DRDCDA, CASCADE, and IVG under different densities of vehicles.

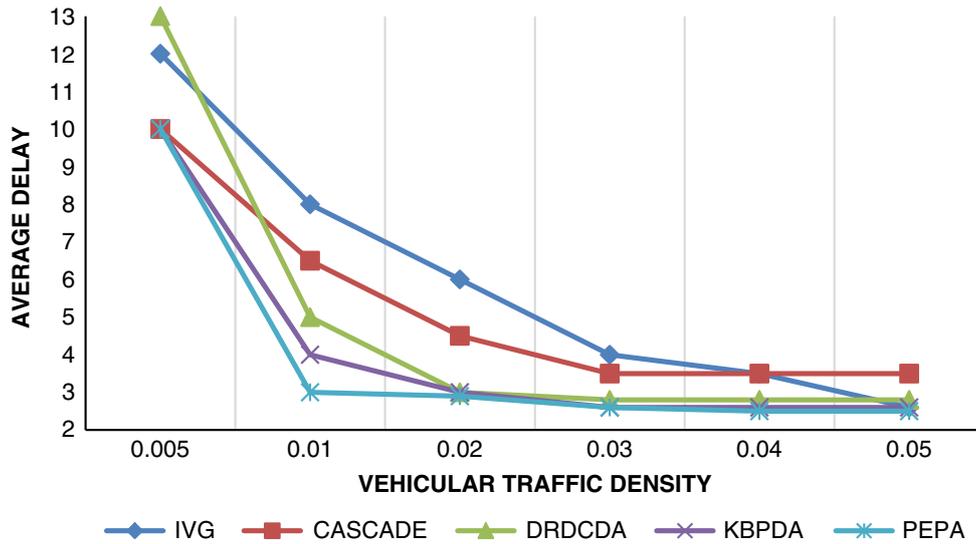


Figure 5: Average delay of PEPA, KBPDA, DRDCDA, CASCADE, and IVG under different densities of vehicles.

superior to the existing state-of-the-art methods. This reduction in PDR is more so because as the sending data rate increases, the number of packets taking part in collisions also increases, and hence some packets go waste and need further retransmission in order to complete the communication. With the variety of data-sending rate, the average delay is calculated and is shown in the plot of Figure 7.

It is shown that CH-to-CH propagation of aggregated information takes less time for a packet to reach

as long as the nearby RSU or the immediate next CH is available. This suits more to the city scenarios. There is a slight increase in average delay with an increase in data-sending rate and is due to collisions taking place in the dissemination path. The larger the data rate, more are the collisions, still, PEPA performs better in comparison to the other state-of-the-art techniques.

Aggregation accuracy indicates the number of packets received from cluster members with correct signatures. The larger the number of packets received

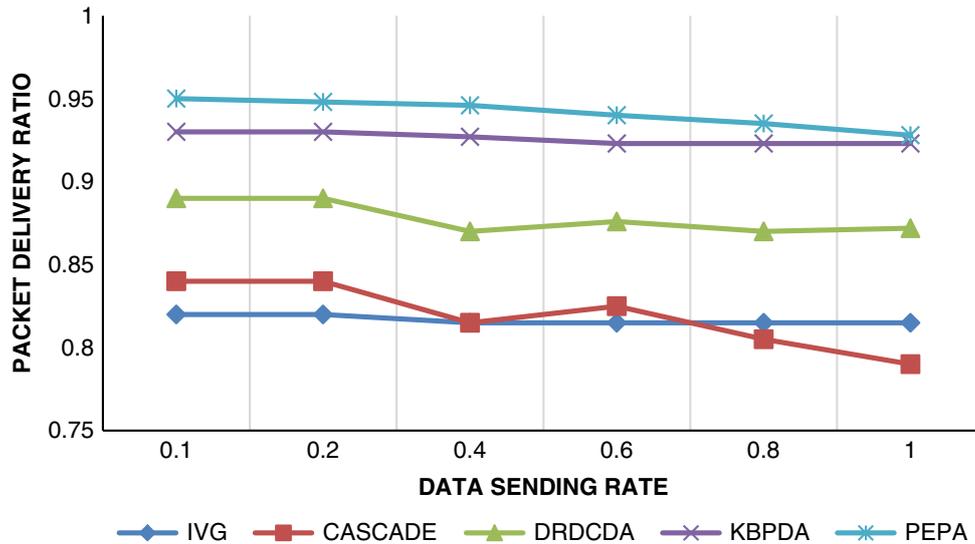


Figure 6: Packet delivery ratio of PEPA, KBPDA, DRDCDA, CASCADE, and IVG under varying data-sending rates.

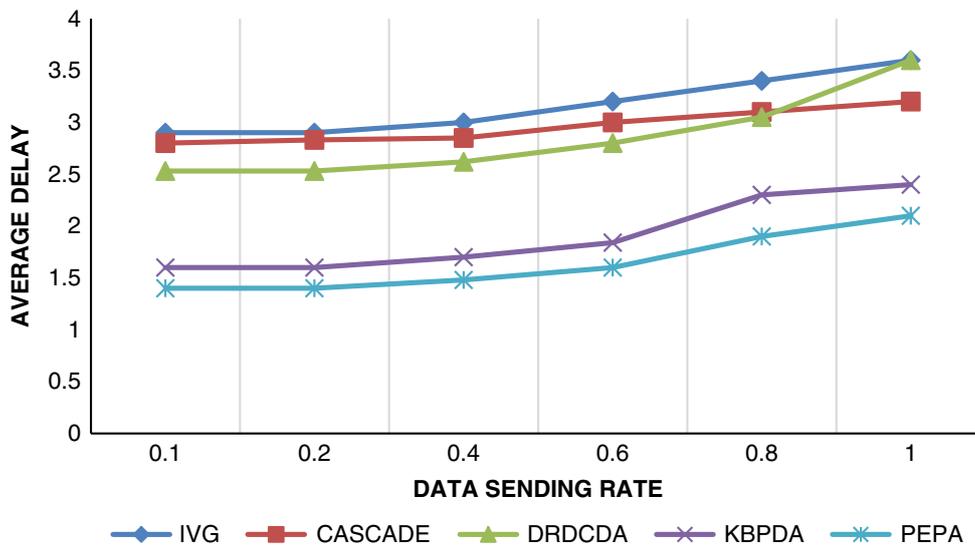


Figure 7: Average delay of PEPA, KBPDA, DRDCDA, CASCADE, and IVG under varying data-sending rates.

from cluster members with correct signatures, the larger is aggregation accuracy. This parameter helps in better identification of the packet that has a higher degree of correctness. Figure 8 shows the better aggregation performance of PEPA.

## Conclusion and future scope

Not only using structured aggregation of data on the road but through structure-free aggregation also, we

can achieve correct data aggregation in VANETs. The proposed scheme provides the probable correctness parameter of aggregated information. This makes the decision-making process easier for the user, such as whether to keep the aggregated packet or to ignore it while taking a decision. The result indicates the robust performance of PEPA under variable traffic densities and data rates in comparison to the state-of-the-art schemes. This work needs to be tested thoroughly for highway scenarios where the density

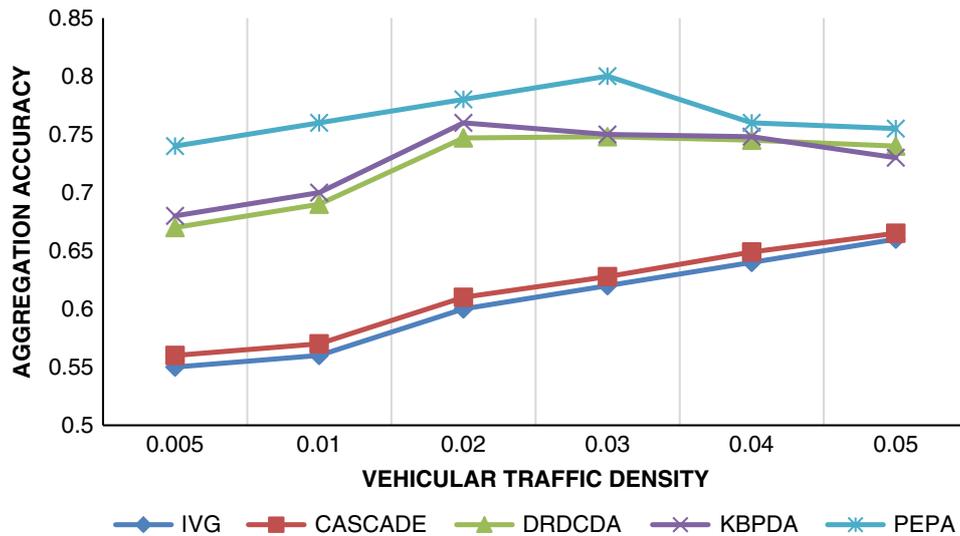


Figure 8: Aggregation accuracy of PEPA, KBPDA, DRDCDA, CASCADE, and IVG under varying vehicular traffic densities.

of vehicles is very less, the cluster head may not find the nearby RSU or CH to propagate the aggregated packet, and hence average delay may arise. Hence, a linear programming model can be used to enhance dynamic aggregation further.

## Literature Cited

Allani, S., Yeferny, T. and Chbeir, R. 2018. "A scalable data dissemination protocol based on vehicle trajectories analysis". *Ad Hoc Networks* 71: 31–44.

Allani, S., Chbeir, R., Yeferny, T. and Yahia, S. B. 2018. Smart directional data aggregation in VANETs. *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, IEEE, pp. 63–70.

Allani, S., Yeferny, T., Chbeir, R. and Yahia, S. B. 2020. "Towards a smarter directional data aggregation in VANETs". *World Wide Web* 1–20.

Bali, R. S., Kumar, N. and Rodrigues, J. J. P. C. 2014. 'Clustering in vehicular ad hoc networks: taxonomy, challenges and solutions'. *Vehicular Communications* 1(3): 134–152.

Bilal, S., Muhammad, S., Khan, U., Madani, S. A., Nazir, B. and Othman, M. 2014. 'Road oriented traffic information system for vehicular ad hoc networks'. *Wireless Personal Communications* 77(4): 2497–2515.

Caliskan, M., Graupner, D. and Mauve, M. 2006. 'Decentralized discovery of free parking places'. *3rd ACM International Workshop on Vehicular ad hoc Networks* Los Angeles, CA.

Dietzel, S., Petit, J., Kargl, F. and Scheuermann, B. 2014. 'In-network aggregation for vehicular ad hoc networks'. *IEEE Communications Surveys & Tutorials* 16: 1909–1932.

Flajolet, P. and Martin, G. N. 1985. 'Probabilistic counting algorithms for database applications'. *Journal of Computer and System Sciences* 31(2): 182–209.

Ibrahim, K. and Weigle, M. C. 2008. CASCADE: Cluster-based accurate syntactic compression of aggregated data in VANETs *IEEE Globecom Workshops IEEE*, pp. 1–10.

Ji, Y., Wu, C. and Yoshinaga, T. 2016. "Context-aware unified routing for VANETs based on virtual clustering", In *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, IEEE, pp. 1–6.

Kumar, D., Sharma, A., Kumar, R. and Sharma, N. 2020. "A holistic survey on disaster and disruption in optical communication network", *Recent Advances in Electrical & Electronic Engineering* 13: 130.

Kumar, R. and Dave, M. 2013. 'A framework for handling local broadcast storm using probabilistic data aggregation in VANET'. *Wireless Personal Communications*. 72(1): 315–341.

Kumar, R. and Dave, M. 2016. 'Data relationship degree-based clustering data aggregation for VANET'. *International Journal of Electronics* 103(3): 485–503.

Kumar, R., Kumar, D. and Kumar, D. 2019. 'Performance enhancement of traffic information gathering (PEntInG) algorithm for vehicular ad-hoc networks'. *International Journal of Communication Systems* 32(15).

Lee, B. K., Jeong, E. H. and Jeong, Y. 2014. 'Message propagation based on three types of density

classification for smooth and secure vehicular traffic flow'. *International Journal of Multimedia and Ubiquitous Engineering* 9(12): 383–404.

Lochert, C., Scheuermann, B. and Mauve, M. 2007. 'Probabilistic aggregation for data dissemination in VANETs'. 4th ACM International Workshop on Vehicular ad hoc Networks, Montreal.

Lochert, C., Scheuermann, B. and Mauve, M. 2010. 'A probabilistic method for cooperative hierarchical aggregation of data in VANETs'. *Ad Hoc Networks* 8(5): 518–530.

Lochert, C., Scheuermann, B., Wewetzer, C., Luebke, A. and Mauve, M. 2008. 'Data aggregation and roadside unit placement for a VANET traffic information system' *Proceedings of the 5th ACM International Workshop VANET*.

Milojevic, M. and Rakocevic, V. 2015. 'Location aware data aggregation for efficient message dissemination in vehicular ad hoc networks'. *IEEE Transactions in Vehicular Technology* 64(12): 5575–5583.

Molina-Gil, J., Caballero-Gil, P. and Caballero-Gil, C. 2014. 'Aggregation and probabilistic verification for data authentication in VANETs'. *Information Sciences* 262: 172–189.

Nadeem, T., Dashtinezhad, S., Liao, C. and Iftode, L. 2004. TrafficView: a scalable traffic monitoring system. *IEEE International Conference on Mobile Data Management*.

Poongodi, M., Hamdi, M., Sharma, A., Ma, M. and Singh, P. K. 2019. "DDoS detection mechanism using trust-based evaluation system in VANET. *IEEE Access* 7: 183532–183544.

Rathee, G., Sharma, A., Kumar, R., Ahmad, F. and Iqbal, R. 2020. "A trust management scheme to secure mobile information centric networks". *Computer Communications* 151: 66–75.

S Al-Sultan, M., Moath, Al-Doori, H., Al-Bayatti Ali and Zedan, H. 2014. 'A comprehensive survey on vehicular ad hoc network'. *Journal of Network and Computer Applications* 37(1): 380–392.

Schoch, E., Kargl, F., Weber, M. and Leinmuller, T. 2008. 'Communication patterns in VANETs'. *IEEE Communications Magazine* 46(11): 119–125.

Sleet, H., Basir, O., Langar, R. and Boutaba, R. 2010. 'Region-based location service-management protocol for VANETs'. *IEEE Transactions on Vehicular Technology* 59(2): 917–931.

Ucar, S., Ergen, S. C. and Ozkasap, O. 2014. VeSCA: vehicular stable cluster-based data aggregation. *International Conference on Connected Vehicles and Expo (ICCVEx)*, Vienna.

Ucar, S., Ergen, S. C. and Ozkasap, O. 2016. 'Multi-Hop cluster based IEEE 802.11p and LTE hybrid architecture for VANET safety message dissemination'. *IEEE Transactions on Vehicular Technology* 65(4).

Wisshhof, L., Ebner, A., Rohling, H., Lott, M. and Halfmann, R. 2003. SOTIS – a self-organizing traffic information system *Proceedings of 57th IEEE Semi-annual, Vehicular Technology Conference*.

Zekri, D., Defude, B. and Delot, T. 2012. 'A cooperative scheme to aggregate Spatio-temporal events in VANETs' 16th International Database Engineering and Applications Symposium (IDEAS), pp. 100–109.