Vehicular Communications ••• (••••) •••–•••

Contents lists available at ScienceDirect

Vehicular Communications

www.elsevier.com/locate/vehcom

Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions

Rasmeet S Bali [a], Neeraj Kumar [a], Joel J.P.C. Rodrigues [b]

[a] Department of Computer Science and Engineering, Thapar University, Patiala, Punjab, India
[b] Instituto de Telecomunicações, University of Beira Interior, Portugal

A R T I C L E I N F O

Article history:
Received 15 February 2014
Received in revised form 5 May 2014
Accepted 5 May 2014
Available online xxxx

Keywords:
Vehicular ad hoc networks
Clustering
Data dissemination
Routing

A B S T R A C T

Over the last few years, Vehicular Ad Hoc Networks (VANETs) have emerged as a new class of efficient information dissemination technology among communities of users mainly because of their wide range of applications in different domains such as Intelligent Transport Systems (ITS), Safety applications, online entertainment during the mobility of the vehicles etc. Vehicles in VANETs are acting as an intelligent machine, which provides various resources to the end users with/without the aid of the existing infrastructure. But due to the high mobility and sparse distribution of the vehicles on the road, it is a quite challenging task to route the messages to their final destination. To address this issue, clustering has been widely used in various existing proposals in literature. Clustering is a mechanism of grouping of vehicles based upon some predefined metrics such as density, velocity, geographical locations of the vehicles etc. Motivated by these factors, in this paper, we analyzed various challenges and existing solutions used for clustering in VANETs. Our contributions in this paper are summarized as follows: Firstly, a complete taxonomy on clustering in VANETs has been provided based upon various parameters. Based upon this categorization, a detailed discussion is provided for each category of clustering which includes various challenges, existing solutions and future directions. Finally, a comprehensive analysis of all the existing proposals in literature is provided with respect to various parameters such as topology selected, additional infrastructure requirements, road scenario, node mobility, data handled, and relative direction, density of the nodes, relative speed, communication mode, and communication overhead. The analysis provided for various existing proposals allows different users working in this domain to select one of the proposals with respect to its merits over the others.

1. Introduction

Vehicular Ad Hoc Networks (VANETs) consist of Vehicles/Mobile nodes communicating with each other over wireless links with/without existing infrastructure [1]. Vehicles have the capability to communicate directly with other vehicles in Peer-to-Peer (P2P) manner or indirectly using the existing infrastructure alongside the road. Vehicles and roadside infrastructure need to be equipped with dedicated hardware for providing safety and security to the passengers sitting on board. Also standardization of wireless communication technology is required for providing entertainment to the passengers. Therefore research on VANETs has been receiving increasing interest in the last couple of years, both in the algorithmic aspects as well as standardization efforts like IEEE 802.11 p & IEEE 1609 standards. In a clustering structure, the mobile nodes are divided into a number of virtual groups based on certain rules. These virtual groups are called clusters. Under a cluster structure, mobile nodes may be assigned a different status or function, such as cluster-head, cluster-gateway, or cluster-member. A cluster-head normally serves as a local coordinator for its cluster, performing intra-cluster transmission arrangement, data forwarding etc. A cluster-gateway is a non-cluster-head node with inter-cluster links, so it can access neighboring clusters and forward the information between clusters. A cluster-member is usually called as an ordinary node, which is a non-cluster-head node without any inter-cluster links.

The notion of cluster organization has been used for Mobile Ad Hoc Networks (MANETs) in number of issues such as routing, security, Quality of Service (QoS) etc. [1,2]. However due to the characteristics of VANETs such as high speed, variable density of the nodes, clustering schemes which are proposed for conventional MANETs may not be suitable for VANETs. Due to time taken

E-mail addresses: rasmeetsbali@gmail.com (R.S Bali), nehra04@yahoo.co.in (N. Kumar), joeljr@ieee.org (J.J.P.C. Rodrigues).

http://dx.doi.org/10.1016/j.vehcom.2014.05.004
2214-2096/

for cluster formation and maintaining a cluster structure, clustering requires additional control overhead. Thus a good clustering algorithm should not only focus on forming minimum number of clusters but also dynamically maintain the cluster structure without increasing a high communication overhead over the network. Thus clustering allows the formation of a virtual communication backbone that supports efficient data delivery in VANETs and it also improves the consumption of scarce resource such as bandwidth. A low cost clustering method should be able to partition a VANET in a short time with little overhead of control message broadcasting. Hence, VANETs must follow a tight set of constraints as compared with MANETs and therefore require specialized clustering scheme.

The developed clustering algorithm should be distributed, with no central coordinator. The algorithm should also handle the locality properly, i.e., single topology change should have as local impact as possible on the cluster topology and should be able to detect and react to topology changes. Because of the high degree of mobility, a clustering algorithm should converge fast and should have a reduced overhead to minimize the time lost in the clustering process.

1.1. Motivation and challenges

As VANETs have been used in various applications whose ultimate goal is to provide safety and comfort to the passengers sitting in the vehicles, hence there is a requirement of optimized solutions for clustering in VANETs. Also, due to large number of nodes and lack of routers, a flat network scheme may cause serious scalability and hidden terminal problems. A possible solution to above problems is the use of an efficient clustering algorithm. As for efficient communication among the vehicles on the road, Dedicated Short Range Communications (DSRC) is used, so it would be a good idea to divide the vehicles in clusters so that vehicles within the same cluster may communicate using DSRC standards. These facts motivate us to categorize various clustering techniques in VANETs based upon some criteria. But on the other hand, there are number of challenges that need well designed solutions for clustering of vehicles. Some of the challenges are high mobility of the vehicles, sparse connectivity in some regions, security etc.

Due to large and varied nature of parameters that have been considered in different clustering, it was difficult to consider some parameters as standard for evaluation of reviewed protocols. To accommodate this diversity, all the parameters were analyzed and then synthesized into eight standard categories. These eight parameters have been broadly categorized into those that primarily impact vehicular movement; vehicle density and vehicle speed; which characterize the efficiency of clustering protocol; cluster stability, cluster dynamics, cluster convergence and cluster connect time; and that constrain the network performance; transmission efficiency and transmission overhead. This standardization will on one hand help us to provide a comparative analysis of all the reviewed clustering protocols and also assist future researchers in generating a standard set of comparative analysis with existing work.

Vehicle Density designates the average number of vehicles defined in terms of vehicles per kilometer (km) or vehicles per lane. For urban scenarios high value of vehicle density is considered compared to highways. Vehicle speed is the range of speeds considered for simulation by a particular protocol in terms of m/s or km/h. A speed range that varies realistically indicates better adaptability. Transmission efficiency is described as average number of messages or packets that are transmitted or received by a cluster member during a time duration. High transmission efficiency shows that a clustering scheme is more effective in data dissemination. Transmission overhead is the average communication or control overhead required by a clustering scheme for cluster formation and maintenance in terms of number of packets or Mbytes. A clustering scheme that has lower transmission overhead is desired. Cluster stability is the average life-time of a cluster. A high value of cluster stability indicates a better clustering protocol. The parameter cluster dynamics describes the average number of status changes per vehicle defined in terms of average number of cluster changes or cluster head changes in terms of total number of vehicles. A low value of cluster dynamics is more suitable. Cluster connect time refers to percentage time duration that a vehicle stays connected to a single cluster. A high value of cluster connect time indicates higher suitability of a protocol. Cluster Convergence refers to the duration required for all the nodes to join a cluster at the initiation of a clustering scheme. The suitability of a clustering scheme for VANETs is more when it exhibits low clustering convergence.

1.2. Main contributions

Based upon the above discussion, the main contributions of this paper are summarized as follows:

- A complete taxonomy for clustering in VANETs has been provided which categorizes clustering based upon various key parameters;
- A detailed description of the protocols in each category has been provided in the text. Moreover, an analysis of the protocols of each category is provided by careful selection of various parameters;
- Finally, a detailed comparison and discussion of various approaches and protocols have been provided with respect to various parameters. Also, open issues and future directions in this newly emergent area are highlighted in the text, which guides various users to select a particular solution based upon its merits over the others.

1.3. Taxonomy of the clustering in VANETs

For efficient communication among the nodes in the network, stable clustering is required. In this direction, many researchers have used various techniques to form a stable cluster among the nodes. Some of these techniques consist of the use of signal strength received, position of the node from the cluster head, velocity of the nodes, direction and destination of node. Keeping in view of the above issues, the detailed taxonomy of various clustering algorithm is described in Fig. 1.

1.4. Organization

Rest of the paper is organized as follows. Section 2 provides the description about the Predictive clustering. Section 3 describes about the Backbone Clustering. Section 4 describes the MAC based clustering. Section 5 discusses about the Traditional clustering. Section 6 explores on Hybrid clustering. Section 7 describes the Secure Clustering. Section 8 provides the comparative analysis with respect to various parameters. Section 9 explores on the open research issues and challenges. Finally, Section 10 concludes the article and gives the future directions on the topic.

2. Predictive clustering

In predictive clustering, the cluster structure is determined by the current geographic position of vehicles and its future behaviour. This vehicle traffic information helps to associate priorities which then assist in cluster formation. The future position and the intended destinations of vehicles have been used in the literature

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

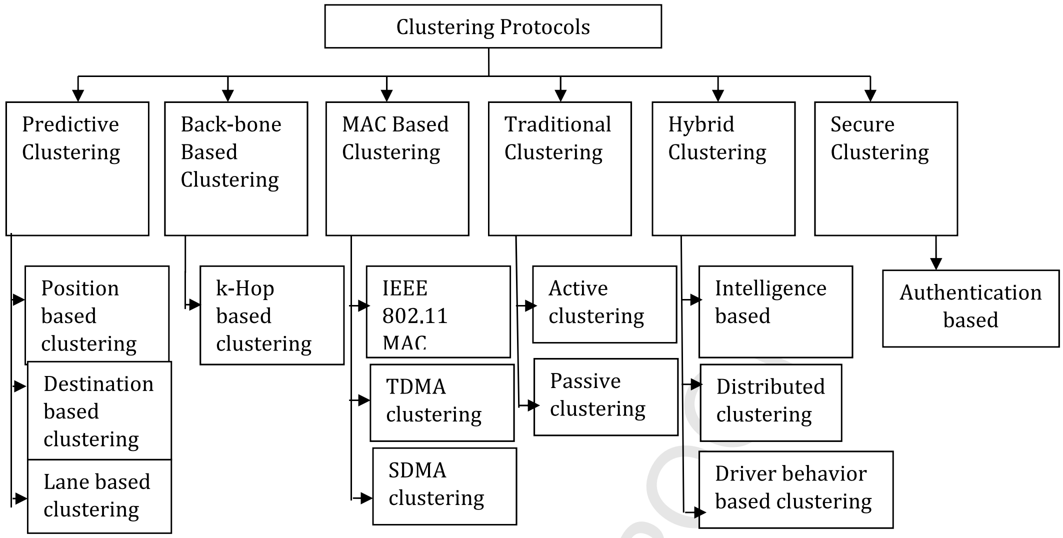

Fig. 1. Taxonomy of existing clustering approaches for VANETs.

to form clusters in VANETs. Some of these protocols are classified as position based and destination based as follows:

2.1. Position based clustering

Position based clustering is a technique of forming the clusters on the basis of geographic position of the vehicle and cluster head. Salhi et al. [2] proposed a new position based clustering algorithm (NEW-ALM) which is an improvement to the existing ALM algorithm. The cluster structure is determined by the geographic position of the vehicle and the cluster-head (CH) is elected based on priorities associated with each vehicle. A hash function based on the estimated travel time is used to generate this priority for the vehicle. The stability of the system is improved by electing the vehicles having a longer trip as the cluster-heads. Though this solution gives a stable cluster structure but its performance is not tested in sparse and jammed traffic conditions which are very frequent in dense urban scenarios.

Wang et al. [3] proposed another position based clustering algorithm. It is a cross layer algorithm based on hierarchical and geographical data collection and dissemination mechanism. The cluster formation in this protocol is based on the division of road segment. However this protocol incurs more overheads for Vehicle-to-Vehicles (V2V) and Vehicles-to-Infrastructure (V2I) communication. Thus its performance is affected based on the availability of an infrastructure.

Fan et al. [4] proposed a clustering scheme where a utility based cluster formation technique is used by extending the definition of Spatial Dependency which was initially proposed in [5]. In the utility function, position and velocity, closest to a predetermined threshold value are used as the input parameters. The threshold is computed based on the previously available traffic statistics. A status message is periodically sent by all the neighbouring vehicles. After receiving this information, each vehicle chooses its CH based on the results produced by the utility function. The node with the highest value is chosen as the CH. This scheme attempts to enhance the classical clustering algorithms by taking into considerations the characteristics particular to VANETs. However it still applies many fixed weights and the parameters like fixed cluster formation interval which implies a synchronous formation of clusters. This scheme fails to adapt to traffic dynamics and is also not applicable for effective cluster re-organization.

Maslekar et al. [6] proposed a new cluster-head election policy for direction based clustering algorithm called as Modified Clustering based on Direction in Vehicular Environment (MC-DRIVE) [7]. The primary functioning of MC-DRIVE is based on the parameter $TH_{DISTANCE}$. This value yields the optimal value of the cluster and is dependent on the speed and the radio range of the vehicles approaching the intersection. $TH_{DISTANCE}$ provides the means for an effective cluster formation and CH election. The proposed clustering algorithm is able to maintain the stability of the cluster in terms of the number of nodes within a cluster. This helps to achieve better accuracy in density estimation. It is also observed that the accuracy can be further improved by reducing the radio range up to a predefined threshold value. However, any further reduction in the radio range leads to an increase in the number of cluster-heads that results an increase in overhead of the system. Wolny [8] optimized the existing DMAC algorithm presented in [9] so that road traffic mobility is represented in an efficient manner. The main idea for modified DMAC was to increase the cluster stability by avoiding re-clustering when groups of vehicles move in different directions. The algorithm is based on periodical transmission of status message and it also forms k-cluster so that nodes can be k-hops away from CH. This is achieved by introducing Time-To-Live parameter in messages sent by the nodes. Modified-DMAC also introduces a method for estimating the connection time (called freshness in DMAC) of two moving nodes. By periodic computation of freshness value, it is possible to avoid re-clustering when two nodes are within the connection range for a short period of time which helps to increase the cluster stability. Although Modified-DMAC increases the algorithm overhead but it reduces the number of cluster changes thereby increasing the stability of cluster formation. Its performance has not been tested in sparse and jammed traffic conditions which are very frequent in dense urban traffic scenarios.

2.1.1. Discussion on position based clustering

Position based clustering solutions are the key for clustering in VANETs. In the last few years, many clustering protocols have been built by considering the various characteristics. Out of these proposals, the protocols based on the vehicles positions are most adequate to VANETs due to their resilience to handling the nodes position variation [10]. Table 1 provides the relative comparison of these protocols with respect to key parameters that influence position based clustering.

Since the above clustering protocols primarily rely on the position of the vehicle, the range of values for vehicular density and vehicle speed exhibits a variation for every protocol as shown in

Table 1
Comparison of position based clustering.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
PPC [3]	LOW	HIGH	LOW	HIGH	LOW	LOW	MEDIUM	HIGH
MODIFIED C-DRIVE [6]	LOW	MEDIUM	LOW	MEDIUM	LOW	LOW	HIGH	HIGH
CGP [2]	HIGH	HIGH	HIGH	LOW	HIGH	LOW	HIGH	HIGH
ALM [4]	LOW	MEDIUM	HIGH	LOW	LOW	LOW	HIGH	HIGH

Table 2
Comparison of destination based clustering schemes.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
LICA [11]	MEDIUM	MEDIUM	MEDIUM	LOW	MEDIUM	MEDIUM	MEDIUM	HIGH
EUCLIDEAN DISTANCE CLUSTERING [12]	HIGH	HGH	MEDIUM	LOW	HIGH	MEDIUM	HIGH	HIGH
AMACAD [13]	LOW	HGH	MEDIUM	LOW	MEDIUM	MEDIUM	MEDIUM	HIGH
CBLR [15]	LOW	MEDIUM	MEDIUM	LOW	MEDIUM	MEDIUM	HIGH	HIGH

Table 1. However the value of cluster convergence rate is low even if vehicle density and cluster dynamics increase which points to better cluster stability for these schemes. The variation in cluster size also affects performance in terms of mean cluster diameter and dismiss threshold for position based clustering. The value of transmission efficiency which ultimately effects packet delivery ratio is also on an average on the lower side. From Table 1, it can be concluded that transmission overhead and cluster connect time needs further analysis for improving the overall efficiency of clustering. The high values of cluster connect time and transmission overhead indicate the need for further analysis and improvement so that position based clustering schemes can be efficiently utilised for VANETs.

2.2. Destination based clustering

Destination based clustering technique takes into account the current location, speed, relative and final destination of vehicle for cluster formation. The destination is known in prior using navigation system. Various proposals in this category are described as follows:

Farhan et al. [11] proposed an algorithm for improving the accuracy of GPS devices called Location Improvement with Cluster Analysis (LICA). Vehicles are able to collect real-time data and relay the information to other vehicles, guiding the drivers to reach the destination safely and efficiently. To measure distance, time-of-arrival and Received Signal Strength (RSS) techniques are used. LICA uses a modified tri-alteration technique in which multiple measurements can be taken for which the average can be used as the final distance measurement resulting in a set of possible refined x–y coordinates on which a cluster analysis is applied, allowing more weight given to accurate data which results an improvement in nodes location estimation. By using accurate distance measurements, the location error is reduced in LICA and thus gives better performance.

Tian et al. [12] presented a clustering method based on a vehicles position and moving direction. The clustering method is based on Euclidean distance, which uses the position information as well as the moving direction to divide the vehicles into clusters. Each vehicle broadcasts beacon message that include its ID latitude, longitude, direction and time to the whole network. The receiving vehicle will first check the beacons hop count value and if the number of hops is larger than the maximum value, it will discard this beacon. Then sender vehicle updates its topology table by calculating the distance between the vehicles. The cluster heads are generated by selecting the vehicle with minimum distance parameter as the cluster-head. The remaining vehicles are then divided into clusters.

Adaptable Mobility-Aware Clustering Algorithm based on Destination (AMACAD) [13] is based on final destination in vehicular networks to enhance the clustering stability. It operates in a distributed way with the final destination, relative destination, speed and current location of a vehicle as parameters to calculate a metric called $F_{v,2}$ by exchanging message with its neighbors. This manages to improve the lifetime of the cluster and thus decreases the number of cluster head changes. It implements an efficient message mechanism to respond in real time and avoid global re-clustering. The algorithm is based on certain assumptions like that the destination of each vehicle is known. It is also assumed that the communication is content based and the routing is geographic based. The minimum value of $F_{v,2}$ is the selection criterion used by a vehicle to join a cluster. Region Group Mobility model proposed in [14] was also modified to make it suitable for VANETs. In AMACAD, the authors evaluated how the variation of the transmission range and speed affects the AMACAD performance. The algorithm works well when average speed of vehicles is almost constant which is most effective in urban areas.

Santos et al. [15] proposed Cluster Based Location Routing (CBLR) algorithm to choose CHs in VANETs. This algorithm is based on the regular transmission of beacons, which are used to distribute the state of the nodes. According to the states of the nodes nearby, a node chooses the appropriate state. To cope with the changes in the topology, each node maintain a neighbor table, in which it lists the nodes with which it can exchange information. The update of this table can also be done according to received beacon messages.

2.2.1. Discussion on destination based clustering

Table 2 compares the destination based clustering schemes. In order to keep the clustering process stable, the frequency of cluster changes is minimized because a vehicle only leaves a cluster when it encounters a CH whose destination is more similar compared to destination of current CH. Thus exploiting the vehicular behaviour by taking into account the final destinations of vehicles enhances the cluster stability and improves the transmission efficiency in message delivery. It also results in higher cluster connect time as the probability of a vehicle leaving a cluster is generally low due to similarity in their destinations. However in case the number of vehicles in a cluster becomes large message broadcast results in high transmission overhead. The impact of vehicle density and vehicle speed on clustering is also not too significant as their values generally are more dependent on characteristics such as existing traffic conditions and road scenario for

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 3
Comparison of lane based clustering schemes.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
Lane based clustering [16]	LOW	HIGH	LOW	MEDIUM	MEDIUM	LOW	MEDIUM	HIGH
BDA [17]	MEDIUM	HIGH	MEDIUM	LOW	MEDIUM	LOW	MEDIUM	HIGH

destination based clustering protocols. Thus these protocols need to be integrated with algorithms that minimize message retransmissions to improve their efficiency.

2.3. Lane based clustering

Lane based clustering forms the cluster structure based estimation of vehicles lane with respect to certain parameters. Some of the proposals in this category are explained as follows:

Fan et al. [16] proposed Broadcasting based Distributed Algorithm (BDA) to stabilize the existing clusters that require only single hop neighbor knowledge and incur minimal overhead. This approach attempts to improve the performance of classical clustering algorithms by making them aware of the vehicle's movement. However, all nodes attempt to re-evaluate their conditions by computing utility values at the same time which may cause traffic overhead and therefore consume more bandwidth. The authors have also theoretically analyzed the message and time complexities of BDA. BDA gives maximum priority to Leadership Duration for cluster formation, which is difficult to compute and may result in large overhead by a node before it joins a cluster.

Almalag et al. [17] presented a lane-based clustering algorithm based on the traffic flow of vehicles. The proposed algorithm is based on the assumption that each vehicle knows its exact lane on the road through some lane detection system and in depth digital street map that includes lane information. It also uses GPS combined with wheel odometer for lane detection of a vehicle. The authors use the same general idea as the utility algorithm in [16], but apply a different set of rules. Each vehicle computes and broadcasts its Cluster Head Level (*CHL*) along with its speed and other parameters. The vehicle with the highest *CHL* will be selected as the CH. *CHL* is determined on the basis of network connectivity level of vehicles and average velocity of traffic flow. The proposed algorithm creates CH that has longer life time as compared to lowest node degree for MANETs.

2.3.1. Discussion on lane based protocols

Lane based clustering algorithms use the availability of lane information to select stable clusters. Table 3 indicates that the above two schemes have low number of CH changes that improves the cluster stability. The transmission overhead of these schemes is also reasonable on account of small number of retransmissions of broadcast messages since re-clustering is performed only at lane intersections. These schemes also display improved transmission efficiency due to better broadcasting reachability and good cluster head lifetime as the vehicles in the same lane move with almost constant relative speed that results in highly stable cluster dynamics. These schemes also exhibit small delay overhead that demonstrates their usefulness for maintaining the cluster even for high mobility vehicular networks. The cluster connect time for these clustering schemes is also reasonable. Thus both the lane based clustering schemes exhibit good clustering characteristics with the constraints that placement of the vehicles should be on the same lane. The observed values of vehicle characteristics such as density and speed is on the lower scale since these protocols are adaptable for urban environment due the constraint of vehicle travelling in the same lane.

3. Backbone based clustering

Backbone based clustering technique is based on forming a backbone for cluster communication. The backbone then performs the communication and assists in CH election among the members of the cluster. Various backbone based clustering techniques in this are classified as follows:

3.1. k-hop clustering

In multi hop or k-hop clustering, cluster structure is controlled by the hop distance. Each cluster has one of the nodes in the cluster as the CH. The distances between a CH and the members of the cluster are within a predetermined maximum number of hops which can be one or more hops. Some of the research proposals in this category are explained as follows:

Zhang et al. [18] proposed a multi-hop clustering scheme based on the mobility metric for representing N-hop mobility. A vehicle is allowed to broadcast beacon message periodically and a vehicle calculates Relative Mobility based upon two consecutive beacon messages received from the same node in N hop distance. Each vehicle node then calculates the aggregate mobility value, which is the sum of relative mobility values into weight value for all the neighboring nodes in N-hops.

The vehicle nodes then broadcast their aggregate mobility value in the N-hop neighborhood and the vehicle with smallest aggregate mobility value is selected as the CH and the other vehicle nodes work as cluster member nodes. The vehicle node joins a cluster if it receives the beacons broadcast from the CH node. When a vehicle node receives multiple beacon messages, it will select the CH which is the closest one in terms of number of hops. If several CHs have the same hops the vehicle node joins the cluster which has the lowest relative mobility.

Zhang et al. [19] proposed a novel k-hop clustering approach that takes into account the highest connectivity, vehicle mobility and host ID to select CH. The proposed clustering approach modifies max–min k-hop heuristic approach defined in [20] for cluster formation by considering highest connectivity in terms of signal strength and vehicle mobility. This scheme is able to dynamically adjust the period of announcing location information according to vehicle velocity in order to suppress transmission overheads. Moreover, the distance-based converge-cast is deployed to collect all memberships within the cluster, including the members located on the cluster border. Another feature of this approach is its ability to enhance cluster stability due to vehicle activation and deactivation by considering the radio link expiration time and the number of vehicles connected to a cluster-head since they are essential to keep vehicle in a cluster. Thus cluster-based topology discovery scheme proposed in this approach utilizes the advantage of k-hop cluster architecture to improve the network topology scalability. It improves the network topology stability with a capability to tolerate false routes and balance traffic loads by considering the inter-cluster link expiration time. By taking into account the factor of vehicle mobility, it reduces the overhead and the latency caused by route path recovery.

Wei et al. [21] proposed a robust Criticality-based Clustering Algorithm (CCA) for VANETs that employed network criticality to direct the process of building clusters. Network criticality derives

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 4
Comparison of k-hop clustering protocols.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
N-HOP CLUSTERING USING RELATIVE MOBILITY [18]	MEDIUM	MEDIUM	LOW	MEDIUM	LOW	MEDIUM	HIGH	HIGH
K-HOP CLUSTERING [19]	MEDIUM	MEDIUM	LOW	MEDIUM	MEDIUM	MEDIUM	HIGH	HIGH
CCA [21]	LOW	HIGH	LOW	MEDIUM	LOW	HIGH	MEDIUM	HIGH
HCA [22]	LOW	MEDIUM	MEDIUM	MEDIUM	MEDIUM	HIGH	HIGH	HIGH

its roots from the definition of random walk between graphs. It is a global measure on a graph which quantifies the robustness of a network graph to the environmental changes, such as traffic shifts, topology modifications, and changes in the origin and destination for traffic. To evaluate the performance of CCA, it was compared with DBA-MAC algorithm that employs the highest degree as the clustering metric. The CCA algorithm improves the lifetime of clusters, and it reveals a more stable structure for multi-hop mobile wireless networks such as VANETs.

Dror et al. [22] proposed a distributed randomized two hop clustering algorithm and named as Hierarchical Clustering Algorithm (HCA) that was influenced by the work presented in [23]. HCA forms TDMA like synchronized clusters. In order to reduce the number of collisions by simultaneous transmissions in the same cluster, transmissions are only allowed on assigned slots by the CH. The algorithm differs from other clustering algorithms for VANETs as it is capable of creating clusters with a larger span from the CH. It also, does not require the knowledge of the vehicles' locations which contributes to its robustness. The algorithm handles the channel access and does not assume any lower layer connectivity. Even though HCA forms few redundant clusters, the formed clusters are much more stable and robust to topological changes caused by vehicular movement. However the mobility pattern influences the algorithm's behaviour and has a great impact on the cluster stability. Nevertheless, HCA also suffers from some difficulties in terms of inter cluster interferences which cause redundant cluster changes and message loss due to message collisions.

3.1.1. Discussion on K-hop clustering protocols

Multi-hop clustering algorithms shown in Table 4 utilize the advantages of k-hop cluster architecture to improve clustering efficiency. It is evident from Table 4 that K-hop clustering schemes have better cluster stability as well as low cluster dynamics. This can be attributed to the reduced variation in CH and cluster-member lifetime. Thus k-hop clustering schemes can provide improved and reliable performance for VANETs, especially for large multi-hop wireless networks when number of vehicles increase in the network. However the impact of vehicle speed and behaviour of vehicle density also needs further analysis as its affect has not been investigated in detail in these protocols. Although these protocols have low cluster convergence time but they suffer from inter cluster interference which needs to be analyze for improving transmission efficiency and reducing transmission overhead. The increase transmission efficiency and hierarchical clustering structure results in larger span as compared to single-hop cluster spans that result in good cluster convergence and large cluster connect time. The protocol also needs to be further investigated by considering different vehicular characteristics.

4. MAC based clustering

Several Medium Access Control (MAC) based clustering techniques have been proposed for cluster formation in VANETs. These techniques use IEEE 802.11 MAC protocol to generate clusters. Some of popular MAC based protocols are discussed as follows:

4.1. IEEE 802.11 MAC based clustering

Su et al. [24] proposed a cluster based Multichannel communication scheme that integrates Clustering with MAC protocols (CB-MMAC). The proposed scheme mainly consists of three core protocols called Cluster Configuration Protocol that groups all vehicles in the same direction into clusters. The Inter-cluster Communication Protocol which dictates the transmissions of real-time safety messages and non-real-time traffics among clusters over two separate IEEE 802.11 MAC-based channels respectively and the Intra-cluster Coordination and Communication Protocol that employs Multichannel MAC algorithms for each CH vehicle for collecting/delivering safety messages from/to cluster-member vehicles using the upstream Time-Division Multiple-Access (TDMA)/downstream-broadcast method and allocating available data channels to cluster member vehicles for non-real-time traffics.

The proposed scheme requires the use of two transceivers—one used for delay sensitive communication within the cluster, while the other is used for inter-cluster data transfer.

Bonini et al. [25] proposed a cross-layered clustering scheme for fast propagation of broadcast messages which is called as Dynamic Backbone Assisted MAC (DBA-MAC) scheme that may be considered an extension of the MAC scheme described in [26]. A dynamic virtual backbone infrastructure is established through a distributed proactive technique. The backbone formation process considers the current distance among candidate backbone vehicles and the estimated lifetime of the wireless connection among neighboring backbone member. The authors have compared the proposed solution with three similar proposals, a simple 802.11 MAC flooding scheme where each vehicle receiving an alert message and then broadcasted it by using the standard IEEE 802.11 back-off scheme, a Fast Broadcast protocol proposed in [24] and the static backbone such as roadside infrastructure system whose nodes are placed at the maximum distance preserving the connectivity. DBA-MAC has been shown to be compliant with IEEE 802.11 DCF systems, and the performance of it shows its advantages in performance, reliability, and overhead reduction.

4.1.1. Discussion on IEEE 802.11MAC based protocols

MAC based protocols have increased percentage collisions and average message delivery delay that results in lower transmission efficiency and high transmission overhead due to increased contention when number of vehicles or speed of the vehicle increases. Message delivery delay is mainly caused by mobility and sparse distribution of vehicles. It directly impacts the application design and deployment for VANETs. Liu et al. [27] identified message delivery distance and density of vehicles as two main factors for such behavior based on a bidirectional vehicle traffic model. The consideration of bi-directional traffic also affects cluster connect time and results in lower cluster convergence. However in these protocols the percentage collisions and incurred overhead for delivering safety messages is decreased by reducing channel contentions for achieving timely and reliable delivery of safety messages. Due to low average relative speed among cluster heads, the overall impact of variation of vehicle speed on these clustering schemes is also low. Table 5 indicates the variation in cluster dynamics and cluster

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 5
Comparison of MAC-based clustering protocols.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
CB-MMAC [24]	MEDIUM	HIGH	HIGH	MEDIUM	LOW	LOW	MEDIUM	HIGH
DBA-MAC [25]	MEDIUM	LOW	LOW	LOW	LOW	LOW	HIGH	HIGH

Table 6
Comparison of TDMA based clustering protocols.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
CBMAC [30]	LOW	MEDIUM	MEDIUM	LOW	MEDIUM	MEDIUM	MEDIUM	HIGH
VeMAC [29]	LOW	LOW	MEDIUM	MEDIUM	MEDIUM	HIGH	MEDIUM	MEDIUM
TC-MAC [31]	LOW	LOW	MEDIUM	MEDIUM	HIGH	MEDIUM	MEDIUM	HIGH

stability that can be attributed to fluctuations in cluster lifetime and cluster size in these protocols.

4.2. TDMA based clustering

The process of assigning time slots can be scheduled using TDMA technique in which slots are assigned for data transmission. Some of the proposals in this category are described as follows:

Biswas et al. [28] proposed Vehicular Self-Organized MAC (VeSOMAC) protocol based on a self-configuring TDMA slot reservation protocol which is capable of inter-vehicle message delivery with short and deterministic delay bounds. To achieve the shortest delay, vehicles determine their TDMA time slot based on their location and movement on the road. Also, the TDMA slot assignment is designed to be in the same sequential order with respect to the vehicles physical location. The process of assigning time slots is performed without using infrastructure or virtual schedulers such as a leader vehicle. However, this assumption of forwarding messages without processing time or propagation delay is unrealistic. In reality, if the message needs to be delivered from the tail to the head of the platoon, it will need a time frame for each hop.

Omar et al. [29] proposed a multichannel MAC protocol for VANETs, called VeMAC, to reduce interference between vehicles and reduce transmission collisions caused by vehicle mobility. VeMAC is based on a TDMA scheme for inter-vehicle communication. Vehicles in both directions and RSUs are assigned time slots in the same TDMA time frame. Also VeMAC is designed based on one control channel and multiple service channels in the network (as with DSRC/WAVE standards). VeMAC assumes that there are two transceivers on each vehicle and that all vehicles are time synchronized using GPS. The first transceiver is assigned to the control channel, while the second transceiver is assigned to the service channels.

Gunter et al. [27] proposed cluster based medium access control protocol (CBMAC), where the CH takes on a managerial role and facilitates intra-cluster communication by providing a TDMA schedule to its cluster members. The CBMAC protocol uses an adoption of CBLR protocol proposed in [13] for cluster formation. Unlike CBLR which is based on regular transmission of status messages, the frequency for sending these messages depends on the state of node in CBMAC. In this scheme the CH takes the responsibility to assign bandwidth to the member of the cluster which reduces the packet collisions due to IEEE 802.11 and also improves QoS. Although CBMAC minimizes the hidden station problem and provides better scalability, it depends on the CH every time a new TDMA frame starts, which will lead to increased communication overhead. CBMAC also demonstrated that the probability of a node will being CH for a short period of time is higher than the probability when the period is long. Traffic Density also amplifies this affect as increase in traffic density increases the number of neighbouring CHs that are in transmission range. However in CBMAC neighbouring CHs are shown to operate together for a certain amount of time rather than immediately changing their state. However, this solution leads to some scaling issues due to CHs exchanging their local schema with conflicting CHs.

Almalag et al. [31] proposed a new TDMA Cluster-based MAC (TC-MAC) that can be used for intra-cluster communications in VANETs. This protocol integrates the centralized approach of cluster management and a new scheme for TDMA slot reservation. The main objective of this work was to allow vehicles to send and receive non-safety messages without any impact on the reliability of sending and receiving safety messages, even if the traffic density is high. The authors also changed the concept of having two intervals by having vehicles listening to the control channel and the service channels during the same cycle. TC-MAC also aims to decrease collisions and packet drops in the channel, as well as provide fairness in sharing the wireless medium and minimizing the effect of hidden terminals. TC-MAC is able to deliver non-safety messages within reasonable time constraints, as well as meeting the requirements of minimum latency in case of safety messages.

4.2.1. Discussion on TDMA based clustering

The access to the medium within a cluster is based on TDMA which is primarily used for optimizing communication. These clustering protocols reduce intra cluster collisions as well as packet loss compared to traditional clustering protocols and thus provide fairness in sharing the wireless medium for VANETs. Table 6 shows that TDMA algorithms have relatively smaller delay of multi-hop safety messages as compared to other clustering schemes. Thus they provide better transmission efficiency for cluster maintenance which improves the overall throughput of both inter-cluster and intra-cluster communication.

These protocols also exhibit high transmission overhead due to extra cost of channel assignment for TDMA slots. However by using some optimization techniques the performance of these protocols can be further enhanced. Calafate al. [32] proposed a scheme that minimizes content delivery time by seeking optimal packet size for content delivery. Thus TDMA based clustering schemes have capable transmission characteristics, but their behavior needs further analysis on traditional vehicle characteristics such as vehicle speed and density. Cluster stability is also low for these protocols. Although cluster connect time is comparatively reasonable, but high clustering convergence due to TDMA time slot is a serious bottleneck in implementing these protocols in VANTE's.

4.3. SDMA based clustering

In SDMA based protocols, the road is subdivided into fixed length segments, and a segment is divided into a fixed number of blocks. Each block is assigned a timeslot representing the allowed time for a vehicle to transmit data. SDMA is known to have

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 7
Comparison of SDMA-based clustering protocols.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
CDGP [36]	MEDIUM	LOW	LOW	LOW	LOW	MEDIUM	HIGH	HIGH
Traffic Gather [34]	HIGH	LOW	LOW	LOW	MEDIUM	MEDIUM	HIGH	LOW

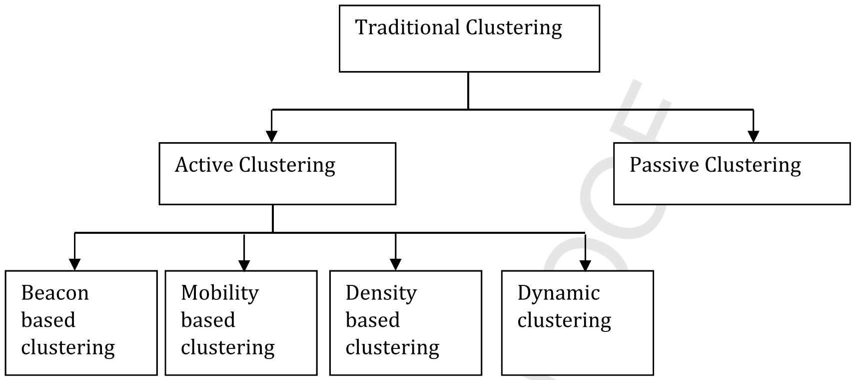

Fig. 2. Classification of traditional clustering.

better performance in a dense network where practically all slots are used. But, the performance decreases proportionally with the density. Hence, in sparse networks, SDMA gives poor performance.

Salhi et al. [33] proposed a protocol for hybrid vehicular architecture, called Clustered Gathering Protocol (CGP). The protocol is designed to provide real-time data (e.g. average speed of vehicle) related to speed of vehicle etc. to base station. The choice of the closest node when it is at the end of the segment can increase the delay in cluster formation by repeatedly running of the election procedure. This also results in increased delay in message propagation in CGP. The use of IEEE 802.11 DCF can also prevent the establishment of possible communications between vehicles in neighbouring segments even after receiving a Clear to Send (CTS) packet sent by the neighbouring CH. Another drawback of CGP is that it doesn't define any retransmission mechanism to deal with the reception of erroneous data.

Chang et al. [34,35] proposed a different dynamic cluster based vehicle to vehicle protocol using SDMA. The protocol was called Traffic Gather. This protocol inherits all the drawbacks of the use of a static medium access technique in wired or sensors networks. Thus, in the case of sparse density, many allowed slots will not be utilized. Although the reliability of SDMA increases in the case of dense network, use of flooding technique may cause a broadcast storm problem even without using a mechanism of retransmission.

Brik et al. [36] proposed a new data collection protocol for vehicular environments called Clustered Data Gathering Protocol (CDGP). The use of a clustering technique in hybrid architecture, Dynamic SDMA in the data collection phase and retransmission mechanism to deal with erroneous data is the major characteristics of CDGP. It avoids collision problems by implementing a centralized, dynamic medium access technique, and enhancing the reliability by the integration of retransmission mechanism.

4.3.1. Discussion on SDMA-based clustering protocols

Table 7 shows that SDMA based clustering protocols show average clustering convergence and transmission efficiency. This is due to the data collection time and number of time slots increasing linearly at approximately constant rate for the discussed protocols. The SDMA mechanism also effects clustering overhead in terms of packet delivery ratio. The vehicle density also influences the message transmission time and results in lower cluster stability. SDMA based schemes also have larger cluster connect time due to re-clustering frequency being high. The throughput and number of transmissions also have an impact on cluster dynamics. These protocols also need to be investigated to improve vehicular parameters such as cluster stability and cluster dynamics. The cluster connect time and vehicle density affects the performance of SDMA-based clustering schemes. The vehicle density will also degrade the message transmission time for these protocols. However SDMA based clustering protocols can be used for providing V2I wireless communication such as Intelligent Transportation System (ITS) in the near future. Initial investment costs could discourage the deployment of a ubiquitous roadside infrastructure to support on-the-road networks and their absence implies discontinuous coverage and short-lived connectivity [37]. The scheme proposed by Salhi et al. [33] has not been discussed in Table 7 as it does not specify any simulation results.

5. Traditional clustering

This section discusses the Traditional Clustering techniques used in VANETs. These techniques are subdivided in to active and passive clustering based upon the role of nodes in VANET. Fig. 2 shows the subcategory of each of these techniques in VANETs.

5.1. Active clustering

In case of active clustering protocols, there are continuous updates of the clustering information and routing table for route discovery after a fixed interval of time. They generally initiate clustering process through flooding which generates a sustained routing overhead. The various Active Clustering protocols are described as follows:

5.1.1. Beacon based clustering

In Beacon based clustering, clusters are formed based on some vehicular or network parameter detected by beacons of hello messages by the receiving vehicle. Little et al. [38] proposed a beacon based clustering model, which is an extension of the algorithm proposed in [2,5]. In this approach, the clusters are formed based on mobility metric and the signal power detected at the receiving vehicles on the same directed pathway. RSS value is used as a criteria to assign weights to the nodes and based on this weights the CH is elected. Using this method, the proposed protocol helps

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 8
Comparison of beacon-based clustering protocols.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
DPP [38]	LOW	MEDIUM	LOW	LOW	LOW	MEDIUM	MEDIUM	LOW
ER-AC [39]	HIGH	LOW	LOW	LOW	MEDIUM	MEDIUM	HIGH	LOW
LORA-CBF [41]	LOW	LOW	LOW	LOW	LOW	MEDIUM	MEDIUM	MEDIUM

in forming stable clusters. However, it does not consider the occurrence of losses in the wireless channel. In practical scenario effects of multipath fading are bound to affect the cluster formation method and thus the stability.

Teshima et al. [39] proposed an active clustering scheme that combines the traditional Epidemic routing with autonomous clustering Scheme proposed in [40]. They considered a complete cluster as a single virtual node, and only the CH stores data packets. Whenever a cluster, encounters a new neighboring cluster, the CH forwards data packet to the CH of the neighboring cluster. The data packets are forwarded to the destination node based on the hierarchical tree with the CH at the root. The CH constructs the CH-based tree that contains the list of cluster members to manage its cluster. The proposed scheme is more efficient in terms of data storage since it store data packets only in the cluster head and all the nodes do not have to store data packets, which result in reduced consumed packet buffer.

Santos et al. [41] presented a reactive location based routing algorithm that uses cluster-based flooding for VANETs called LORA-CBF. This clustering approach is based on regular beacon transmissions which advertise the state of the node. Each node can be a CH, gateway or cluster member. The CH maintains information about its member and gateways packets. Based on the state of the neighbouring nodes, a node can select its own state. A CH will only consider a change of its state if it receives a message from another CH. A CH receiving a hello message from another CH will remain in the same state if it has more Mobile Nodes on its cluster than the sender. This simple criterion favors larger clusters and does not take into account the mobility of the cluster members, how cohesive the smaller cluster is, or if the clusters are moving in opposite directions. Also, with large neighbourhoods, the clusters will have the tendency to grow uncontrollably, thus potentially overloading their CHs.

5.1.1.1. Discussion on beacon-based clustering protocols Beacon based clustering protocols provide increased transmission overhead especially due to increase in number of vehicles and hop-counts. The effect of transmission efficiency in terms of volume of consumed packet buffer and end-to-end delay is high but decreases gradually with an increase in size of the network in terms of vehicle density. This indicates an efficient message delivery at low traffic but the packet delivery ratio starts degrading as hop count or number of vehicles increases.

The periodic transmission of beacons helps in cluster convergence but it also affects the throughput of vehicular network, especially at higher traffic density and ultimately the transmission overhead as shown in Table 8. The protocols can be modified by using some quota based protocols like TTL Based Routing as described in [42] that restricts the maximum number of copies of a message in the network as well as enhances the chance of message delivery. This will assist in improving cluster dynamics that has relatively lower values for all the three schemes discussed above. The impact of vehicle speed also needs to be improved for these protocols.

5.1.2. Mobility based clustering

Maglaras et al. [43] proposed a distributed clustering algorithm which forms stable clusters based on force directed algorithms. Every node applies to its neighbours a force according to their distance and their velocities. Vehicles that move to the same direction or towards each other apply positive forces while vehicles moving away apply negative forces. According to the current state of the node and the relation of its F to neighbour's F, every node takes decisions about clustering formation, cluster maintenance and role assignment. This work also proposed mobility metric based on forces applied between nodes according to their current and their future position and their relative mobility.

A new stability-based clustering algorithm (SBCA) [44], that aims to reduce the communication overhead that is caused by the cluster formation and maintenance, as well as to increase the lifetime of the cluster. SBCA makes use of mobility, number of neighbours, and leadership or CH duration in order to provide a more stable architecture. The nodes remain associated with a given cluster and not with any CH as is the case with most existing clustering approaches. When one CH is no longer in the cluster, another CH takes over; the cluster structure does not change but only the node playing the role of CH. This allows for stable cluster architecture, with low overhead and better performance. SBCA protocol significantly improves the cluster residence time, for each node, reducing the overhead and thus improving the performance/reliability also.

Souza et al. [45] presented a beacon-based clustering algorithm for prolonging the cluster lifetime in VANETs by using a new aggregate local mobility criterion to decide upon cluster re-organization. A node's ALM is the variance of the relative mobility over all neighbours. Two nodes moving closer together result in a negative mobility. A lower variance means less mobility of the node in relation to its neighbours. The intuition behind this scheme is that a node with less variance relative to its surroundings is a better and more stable choice for CH. The proposed clustering algorithm displays a better performance in terms of stability. However since the nodes are highly dynamic in nature the position of the nodes change very fast and hence may induce a computational overhead in calculating the weight associated with the nodes.

The algorithm called, Affinity Propagation for Vehicular networks (APROVE), [46] a distributed mobility-based clustering scheme that forms cluster with low relative velocity between Cluster Members also gives stable clusters by improving cluster head duration, cluster member duration, and reducing rate of CH change. The proposed clustering technique uses the fundamental idea of Affinity Propagation proposed by [47] which has been shown to produce clusters in much less time, and with much less error than previous techniques [48]. The proposed algorithm is validated by comparing it to the mobility-based ad-hoc clustering scheme, MOBIC [41] and simulation results confirm the superior performance of APROVE, when compared to other accepted mobility-based clustering techniques.

Kayis et al. [49] addressed mobility by first classifying nodes into speed groups such that nodes will only join a CH of similar velocity. Code Division Multiple Access scheme is used to assign orthogonal codes to the previously identified vehicular nodes. Every vehicle knows the speed group to which it belongs prior to cluster formation phase. The data packet header transmitted by the sending node is modified by adding speed group information to it. If a major speed change occurs and the node moves at a speed out of its clustering group interval during a threshold value

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 9
Comparison of mobility-based clustering protocols.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
APROVE [46]	LOW	HIGH	MEDIUM	MEDIUM	MEDIUM	LOW	HIGH	MEDIUM
SBCA [44]	MEDIUM	MEDIUM	MEDIUM	LOW	LOW	LOW	LOW	HIGH
SP-CL [43]	LOW	HIGH	MEDIUM	HIGH	LOW	LOW	MEDIUM	HIGH

Table 10
Comparison of density-based clustering protocols.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
DBC [51]	HIGH	HIGH	LOW	MEDIUM	LOW	MEDIUM	MEDIUM	MEDIUM
D-CUT [50]	LOW	LOW	LOW	MEDIUM	LOW	MEDIUM	LOW	LOW

speed T, it updates its speed and clustering group information and the node seeks another cluster to join. In this way, the time until a cluster member leaves the communication range of its CH is extended which increases the life span of the cluster. Consequently, node transition rate between clusters is also decreased.

5.1.2.1. Discussion on mobility based clustering protocols Mobility based clustering protocols minimize relative mobility as well as distance of each CH to its cluster members and thereby attempts to improve the cluster convergence and cluster dynamics. Table 9 indicates that the discussed protocols also display better cluster stability that can be attributed to reduced values of clustering overhead and average number of cluster head changes thereby creating lesser and more stable clusters. This improved stability helps in improving the performance as well as reliability of VANETs and thus make mobility based clustering protocols for those environments which have dynamic behavior and where mobility can be represented efficiently. However vehicle density and vehicle speed which are the predominant factors affecting mobility need to be investigated in more detail for all these protocols. It can also be concluded from Table 9 that the packet delivery ratio for the discussed protocols has a lower value which indicates a decreased transmission efficiency and increased transmission overhead.

5.1.3. Density based clustering

Yairet al. [50] proposed an iterative algorithm named as Distributed Construct Underlying Topology (D-CUT) in which each node discovers and maintains a geographically optimal clustering for the current network configuration. D-CUT algorithm partitions the network into geographically optimized clusters. The protocol is applied in two phases. In first phase, beacons in the same cluster are aggregated by a CH in a synchronized manner. In the second phase, the CH disseminates a compressed aggregated beacon of its own cluster to its adjacent clusters. The vehicles produce a snapshot of the surrounding vehicle map, and update the clustering solution according to the changes in the network configuration. All neighboring vehicles share matching partitioned vehicle map producing the same new partitioned map for each vehicle in the network. The algorithm updates the partitioning according to the most recent topological changes thus maintaining the geographically optimized clusters.

Kuklinshi et al. [51] proposed a multi-level cluster algorithm called the Density Based Clustering (DBC) based on several factors like connectivity level, link quality, relative node position prediction of a nodes position in future and node reputation. This algorithm has three phases. In the first phase, a node estimates its connectivity level defined as number of connection, which is used to discover density of local neighborhood of a node. Every node counts the number of received acknowledgments to find the number of active links. This information determines whether a node belongs to the dense or sparse parts of networks by comparing the connectivity level against a threshold value. The aim of the second phase is to select stable links from all the current links. This selection is made on some prediction about future, but it also takes into account the past knowledge of speed and direction of vehicle. This is the basis for estimating the links quality. In this evaluation a vehicle also uses signal-to-noise ratio of the link. In the last phase communication history is used to determine nodes reputation before it becomes a cluster member. The effects of multipath fading are also taken into account in this density based clustering algorithm.

5.1.3.1. Discussion on density based clustering protocols Density based clustering protocols allow strong connections between cluster members and low variation in number of cluster head changes that results in improved cluster stability. The data in Table 10 indicate that cluster stability is high for density based protocols irrespective of number of vehicles in a cluster. The density information helps in improved awareness in each vehicle about the composition of its cluster and provides strong connections between cluster members for creating a more reliable clustering topology that provides comparable cluster convergence. However further study needs to be undertaken for improving transmission efficiency and decreasing transmission overheads. Dependence only on density limits the cluster connect time and cluster dynamics. The vehicle speed also needs to be further investigated as it has a direct impact on density of the vehicles.

5.1.4. Dynamic clustering

VANETs have relatively more dynamic nature as compared to MANETs resulting in fast change in the network topology. The design and implementation of an efficient and scalable algorithm for information dissemination in VANETs is a major issue that should be tackled. Indeed, in this dynamic environment, an increasing number of redundant broadcast messages will increase resource utilization, which would indirectly affect the network performance [52]. Dynamic clustering technique forms cluster structure based on node dynamics like mobility patterns, velocity and density.

Kakkasageri et al. [53] developed a multi agent based dynamic clustering scheme for VANETs. The scheme comprises of heavy-weight static and light-weight mobile agents and forms a moving dynamic cluster on a lane between two intersections by considering parameters such as vehicle speed, direction, connectivity degree to other vehicles and mobility pattern. Initially, cluster members are identified based on vehicle's relative speed and direction for dynamic clustering. CH is selected among the cluster members based on stability metric derived from connectivity degree, average speed and time to leave the road intersection. It consists of a set of static and mobile agents. The relative speed differ-

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 11
Comparison of dynamic clustering protocols.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
MULTI AGENT DRIVEN DYNAMIC CLUSTERING [53]	LOW	MEDIUM	LOW	LOW	MEDIUM	LOW	MEDIUM	LOW
VWCA [54]	MEDIUM	LOW	LOW	MEDIUM	MEDIUM	LOW	MEDIUM	LOW

Table 12
Evaluation of passive clustering protocols.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
PASSCAR [56]	MEDIUM	LOW	HIGH	MEDIUM	MEDIUM	LOW	LOW	LOW

ence among neighbouring vehicles is the main parameter used for cluster formation. The neighbour vehicles traveling in the same direction on a lane are only considered. The cluster member with the highest stability metric is considered as the CH. This scheme also has certain limitations like assumption that all vehicles need to have relatively strong computational resources, capability of authenticating and validation of vehicles during dynamic clustering process, which is not practical for current vehicular networks.

Daeinabi et al. [54] proposed a novel clustering algorithm, vehicular clustering based on the Vehicular Weighted Clustering Algorithm (VWCA) that takes into consideration the number of neighbours based on dynamic transmission range, the direction of vehicles, the entropy model proposed in [55], and the distrust value parameters. These parameters can increase stability and connectivity and can reduce overhead in network. VWCA works with an Adaptive Allocation of Transmission Range technique, where hello messages and density of traffic around vehicles are used to adaptively adjust the transmission range among them. VWCA uses distrust value in the weighted sum operation. The distrust value has been obtained from proposed Monitoring Malicious Vehicle algorithm. Using distrust value, vehicles that have lower distrust value than their neighbours are elected as cluster-heads. Therefore, cluster-heads are more trusty vehicles than other vehicles in the network. The VWCA technique mainly focuses on improving the CH duration, membership duration and security. Using VWCA, communication overheads required for joining to a new cluster in network decreases because the membership duration for each vehicle has increased. In addition, using the entropy term in the weighted sum operation, VWCA can reduce the number of overheads created by high speed vehicles. Furthermore, VWCA can increase network connectivity when electing cluster-heads.

5.1.4.1. Discussion on dynamic clustering based protocols Table 11 shows the comparison of Dynamic clustering protocols. It shows that cluster dynamics only has an average value that can be attributed to cluster lifetime decreasing by variable rates. The vehicle density also has a negative effect on cluster stability. Initially, the stability is relatively high but decrease with network load. The transmission efficiency has a comparable value with other clustering techniques that can be attributed to moderate range of values of control overhead and percentage connectivity. Since these schemes also have permissible overhead and connectivity, these provide more flexibility and adaptability and can be considered as a good add-on to existing clustering schemes. However the impact of vehicle speed needs to be investigated for realistic vehicle scenario and impact of these clustering protocols on cluster connect time also needs further analysis.

5.2. Passive clustering

Passive clustering is a clustering mechanism that passively constructs a cluster structure [21,56]. At any time, a node in a cluster possesses an external or internal state. In passive clustering each vehicle can lower the control overhead in packet flooding by the use of on-going data packets instead of extra explicit control packets to construct and maintain the clusters. When a node receives data packets, it may change its cluster state based on the state information piggybacked in on-going data packets. This reduces the number of explicit control packets. Thus Passive clustering mechanism generates significantly less overhead for cluster maintenance than the traditional cluster-based technique because its nodes do not maintain cluster information all the time.

Wang et al. [56] proposed a Passive Clustering Aided Routing protocol for VANETs (PassCAR) that refines the passive clustering mechanism proposed in [57] whose main goal was to construct a reliable and stable cluster structure for enhancing the routing performance in VANETs. The proposed mechanism also includes the route discovery, route establishment, and data transmission phases. The main idea behind PassCAR was to select suitable nodes to become cluster-heads or gateways, which then forward route request packets during the route discovery phase. PassCAR assesses the suitability of nodes using a multi-metric election strategy.

This strategy considers link reliability, link stability, and link sustainability as the main factors and quantifies them using the metrics of node degree, expected transmission count, and link lifetime, respectively. Each CH or gateway candidate self-evaluates its qualification for CH or gateway based on a priority derived from a weighted combination of the proposed metrics. PassCAR designs an efficient passive clustering based mechanism that operates at the logical link control sub-layer, and the proposed mechanism can easily be associated with any routing protocol to support stable, reliable, and permanent data delivery.

5.2.1. Discussion on passive clustering

The number of clusters constructed using Passive Clustering as shown in Table 12 remains steady and low for varying vehicular concentration that indicates medium cluster stability. This can be attributed to the consideration of node degree as a key parameter in this protocol. The achieved transmission efficiency is also comparable with other clustering protocols. However due to consideration of link quality, passive clustering has high overhead for cluster formation and maintenance especially in urban environment where obstacles have an effect on link quality. This degrades the cluster connect time. Passive clustering also displays lower cluster convergence due to the use of information being piggy back in ongoing data packets.

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 13
Comparison of intelligence-based clustering protocol.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
FUZZY BASED CH ALGO [58]	MEDIUM	HIGH	LOW	HIGH	LOW	LOW	HIGH	LOW
ALCA [59]	HIGH	HIGH	LOW	HIGH	LOW	LOW	MEDIUM	LOW
MULTI AGENT DRIVEN DYNAMIC CLUSTERING [53]	MEDIUM	HIGH	MEDIUM	MEDIUM	LOW	LOW	HIGH	LOW

Table 14
Comparison of cooperative de-centralized clustering.

	Vehicle density	*Cluster stability*	*Vehicle speed*	*Cluster dynamics*	*Transmission efficiency*	*Clustering convergence*	*Transmission overhead*	*Cluster connect time*
QuickSilver [61]	LOW	LOW	HIGH	LOW	MEDIUM	HIGH	HIGH	MEDIUM

6. Hybrid clustering

Hybrid clustering techniques combine two or more existing techniques such as use of artificial intelligence, fuzzy logic etc. Following are the schemes in this category of clustering.

6.1. Intelligence based clustering

Hafeez et al. [58] proposed a distributed and dynamic Cluster head selection criteria to organize the network into clusters. CH is elected based on stability criteria which reflect the relative movement between adjacent vehicles. The vehicle's acceleration is also used in this work to predict its speed and position in future. However the decision to accelerate, to retardation or to stay on the same speed depends on many factors such as the distance between the vehicle and its front neighbour, the relative speed between them, the road conditions and the driver's behavior. Since the drivers' behaviors and how they estimate the inter distance and other factors are subjective, so triangular fuzzier is used to deal with this uncertainty using the fuzzy logic inference system. The proposed scheme can achieve a highly stable cluster topology which makes it more suitable for implementation in VANETs. However the distributed processing overhead results in decrease message transmission efficiency.

Kumar et al. [59] proposed an Agent Learning-based Clustering Algorithm (ALCA). Agents are able to learn from the environment in which they are operating and perform the task of CH selection. The proposed approach consists of selection of CH keeping in view of the direction of mobility and density of the nodes. The direction of the mobility of the nodes is calculated by the agent in an interactive manner. Agent learns from the direction of motion of the vehicle and traffic flow across different junctions of the road. Agents are deployed at different road junctions for monitoring the activities of the vehicles. Agents perform their action, and accordingly, their action is rewarded or penalized in unit steps. The density of the vehicles and average speed are used for dividing the time into different zones. These zones are then used for collecting the information about the vehicles used as input to the agents for clustering. Learning rate is also defined for the agents to take the adaptive decisions. For each action performed by the agents, the corresponding action is rewarded or penalized, and value of the learning parameter is incremented or decremented. This process continues until the maximum value is reached. The performance of the proposed scheme is evaluated by varying the number of agents with various parameters. The results obtained show that the proposed scheme can be used in the future applications in VANETs.

6.1.1. Discussion on intelligence based clustering protocols

Table 13 shows that all the intelligence based clustering protocols have good cluster stability which is due to large CH duration and cluster member duration. The value of these parameters also improves as the vehicle density increases. These protocols generate reasonably stable clusters but they also cause large transmission overhead which reduces the packet delivery ratio resulting in reduced transmission efficiency for the discussed protocols. Since hybrid techniques or heuristics are employed for cluster formation the additional overhead results in high cluster dynamics. Thus these protocols can be a good alternative for use in future vehicular networks or for those networks that implement a specific application like security, multimedia applications etc.

6.2. Cooperative de-centralized clustering

Cooperative vehicular systems are currently being investigated to design innovative *ITS* solutions for road traffic management and safety. Through various wireless technologies, cooperative systems can support novel decentralized strategies for ubiquitous and cost effective traffic monitoring system [60].

QuickSilver [61] is a light weight distributed clustering protocol that integrates a traditional source routing protocol for intra cluster node centric communication and the construction of a multichannel link for contention free inter cluster data centric communication. It is a system architecture that provides efficient use of available resources to guarantee that no harmful competition takes place for the channel bandwidth. QuickSilver employs lightweight-clustering where clusters form and behave in an uncoordinated manner without requiring a cluster ID and there are no CHs. QuickSilver utilizes two radio interfaces that allows vehicle to maintain their intra cluster connectivity and at the same time look for inter cluster contact opportunities. Cluster formation and maintenance is done by building a cluster formation and management list of neighbors at each node. It focuses on creation of stable links.

6.2.1. Discussion on cooperative de-centralized clustering

Table 14 shows that these clustering protocols have low cluster stability and average cluster connect time. This is due to the fact that average number of inter-cluster links that are active when vehicles are in contact initially increases as the overlapping region for a vehicle increases and then it shows a corresponding decrease as vehicles move away from each other. Transmission efficiency also has comparable value for these protocols. The effect of channel assignment on different node densities is represented in the form of number of links that shows an increase with the number of channels. This indicates the effectiveness of the protocols for inter-cluster communication as shown in Table 14. However high transmission overhead and lower vehicle density results in reduced effectiveness for intra-cluster communication in spite of decentralized clustering schemes considering realistic vehicular speed conditions.

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

Table 15
Comparison of driver behavior based hybrid clustering protocols.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
Cellular Automata Clustering [64]	LOW	HIGH	LOW	HIGH	HIGH	LOW	MEDIUM	LOW
COIN [63]	LOW	MEDIUM	LOW	MEDIUM	MEDIUM	LOW	LOW	LOW

6.3. Driver behavior based hybrid clustering

Vehicles are nowadays provided with a variety of sensors capable of gathering information from their surroundings. In near future, these vehicles will also be capable of sharing all the harvested information, with the surrounding environment and among nearby vehicles over smart wireless links. They will also be able to connect with emergency services in case of accidents [62].

Blum et al. [63] proposed a Clustering for Open Inter vehicle communication Networks (COIN) algorithm. In COIN, CH election is based on vehicular dynamics and driver intentions as input for clustering instead of any conventional parameter like vehicle ID, relative mobility or some other parameters that are used in a classical clustering method. Further COIN attempts to preserve CH for a longer duration and uses mobility information for clustering. Cheng et al. [64] proposed an innovative car-society clustered network based on an imaginative classification scheme. The proposed scheme forms clusters by including vehicles that have the same interest and operating in the same communication range. The aim of the proposed approach is to increase the lifetime of the interest group, and to increase throughput in V2V environments. The proposed scheme develops the interesting ontology of Cellular Automata clustering by using Zone of Interest for mobi-cast communications in VANET environments.

6.3.1. Discussion on driver behavior based hybrid clustering protocols

Table 15 shows that, Driver Intention based Hybrid clustering protocols improve the effectiveness of clustering in terms of cluster lifetime of same interest groups of users that shows good cluster stability. The transmission efficiency for these protocols is also comparable with other clustering schemes but these may have generated for lower values of vehicular speed and vehicle densities. This also results in satisfactory transmission overhead and cluster dynamics which may however not happen if realistic vehicular conditions are considered. Clustering Convergence and Cluster Connect Time also need further investigations. Thus although these protocols provide adequate stability in terms of lifetime of same interest groups but they also need to be further analyzed on the basis of several other parameters for considering their suitability in vehicular environments.

7. Secure clustering

VANETs can support applications and services for safety and comfort for the passengers on the road and assist in improving the efficiency of the road transportation network. However, several serious challenges remain to be solved before efficient and secure VANET technology becomes available. One of these challenges is an efficient authentication of messages using cryptographic techniques [65]. Solutions for secure clustering in VANETs require efficient clustering algorithms in terms of complexity, scalability, availability and reach ability. Several algorithms have been proposed in the literature based on Public Key Infrastructure (PKI) for enabling communications security in vehicular environments. These are based on a trusted third party called as Certification Authority (CA) which is responsible for certifying the public keys of vehicles. Several research schemes have been proposed for distributing the responsibility of the CAs among a set of nodes in the network, using mobility as metric to elect the vehicles that will assume the role of CA.

Blum et al. [66] used a PKI with virtual infrastructure where a set of elected CHs are responsible for disseminating messages after digitally signing them. This scheme is intended only for the attack called intelligent collisions. However, a PKI in VANETs must cope with different attacks.

Raya et al. [67] proposed a distributed PKI for VANETs managed by many certification authorities, each corresponding to a particular region. The different CAs has to be cross-certified so that vehicles from different regions can authenticate each other's CA. This requires that each vehicle store the public keys of all CAs whose certificates are needed to be verified. A location-based approach to form a cluster has been used where the area is divided into small zones or cells that form clusters. A vehicle automatically knows to which group it belongs by comparing its GPS position to a preloaded dissection of the area map into cells. The CH is dynamically determined as the vehicle closest to the center of the cell. The disadvantage of this proposal is the non-availability of the CA in case of a break in the connectivity.

Sivagurunathan et al. [68] proposed a self-organized key management system based on clustering. In their model, the network is divided into number of clusters based on the concept that any user can sign any other public key. The set of signatures forms the network of trusted relationships. However, the drawback of this self-organized approach stems from the assumption that trust is transitive and therefore the system becomes more vulnerable to the intrusion of malicious vehicles.

Gazdar et al. [69] proposed an efficient dynamic architecture of PKI for VANETs based on a trust model. Each elected vehicle will be the CA in its cluster. The proposed clustering algorithm is based on a Trust Metric (Tm) which defines the trust level of a vehicle and is a continuous value in the interval [0–1] and mobility metric which is the relative velocity of a vehicle related to its vicinity. Vehicles start with $Tm = 0.1$ and must prove a good behavior and a good cooperation to increase their Tm through a hierarchical monitoring process that supervises the behavior of nodes at the MAC and network layers as described in [53]. A vehicle with $Tm = 1$ is defined as a confident vehicle. Each node, with a high trust level Tm, monitors its neighbours with lower trust levels. Whenever a vehicle becomes a member in a given cluster, it auto generates a short term pair of keys and then requests certification from its CA. These locally generated pairs of short term keys and their correspondent certificates are assumed to stay valid as long the signing CA keeps serving its cluster. The authors also used a new approach called the *VANET Dynamic Demilitarized Zone* (*VDDZ*). The role of the VDDZ is to prevent unknown vehicles from directly communicating with CA vehicles, thus shielding CAs from malicious nodes. Whenever a vehicle V sends a JOIN request to a CA in a given cluster, the VDDZ must intercept and authenticate this request.

7.1. Discussion on secure clustering protocols

Secure clustering is an approach to increase the channel transmission efficiency and to decrease delay through authentication of a vehicle or a message. These protocols also provide higher level of security which provides data correctness and higher number of message broadcasts that increases the transmission overhead.

Table 16
Comparison of authentication based clustering protocols.

	Vehicle density	Cluster stability	Vehicle speed	Cluster dynamics	Transmission efficiency	Clustering convergence	Transmission overhead	Cluster connect time
ESA [67]	LOW	MEDIUM	MEDIUM	LOW	MEDIUM	MEDIUM	HIGH	MEDIUM
SAV [69]	LOW	MEDIUM	LOW	MEDIUM	MEDIUM	MEDIUM	HIGH	MEDIUM

Table 16 shows that secure clustering approach improves parameters like number of vehicles in a cluster, and cluster lifetime which indicates comparative performance of these protocols in terms of cluster connect time and clustering convergence. However the impact of vehicle density and speed in cluster stability when authentication is also incorporated requires further analysis and can be a viable area for future research work. This has not been addressed in any secure clustering protocols.

8. Comparative analysis and discussion

This section provides a detailed comparison and analysis of various techniques used for clustering in VANETs. Table 17 provides a comparative analysis of various protocols discussed above by considering number of parameters used for clustering in VANETs. These parameters cover most of the key areas that need to be considered for designing a clustering protocol. The detailed analysis of all the main clustering protocols with respect to various parameters is listed in Table 17. The parameters selected for discussion of clustering are: Type of topology handled, Additional Infrastructure requirement, Road scenario handled, Relative mobility of vehicles during clustering, types of messages handled, relative direction of vehicles in a cluster, density of vehicles, different velocity of vehicles in a cluster, communication mode employed for clustering and additional overhead incurred by a clustering protocol. VANETs face many research challenges in multiple areas, from privacy and anonymity to the detection and eviction of misbehaving nodes and many others in between. Multiple solutions have been proposed to address those issues [70].

With respect to the basic parameter used for clustering, we have identified distance between vehicles, direction of vehicle, final destination, vehicle speed, mobility pattern, and duration of a vehicle within a segment, position within a segment, medium contention and number of hops for message transmission as some of the other parameters. Most of the existing clustering protocols use a combination of a parameter that depicts vehicle behavior and another parameter for optimizing message delivery for generating a cluster. Topology indicates the structure of vehicles after the completion of cluster formation. The topology of a cluster can be divided into hierarchical or flat structure. Due to their inherent structure and domain requirements, a large majority of reviewed clustering protocols have tree like hierarchical topology.

The vehicles in a network may require some additional infrastructure like a GPS, transceivers, Lane Detection System, Digital maps, RSU's, odometer etc. for cluster formation and maintenance. This requirement results in some extra hardware modification to be performed on the network which may result in more efficient clustering but may not be possible to implement depending upon the local environment. Thus only a few protocols have shown any need of additional infrastructure and most of the protocols rely on available communication network infrastructure for clustering. The road scenario handled by a protocol is also an important issue and it is used in simulations for estimating parameters like density and speed of vehicles. The clustering protocols reviewed have been classified to be simulated under either highway or urban road traffic conditions and very few clustering protocols have considered both the road scenarios. Mobility is another aspect which indicates whether the vehicles mobility during cluster formation has been considered or not. Various proposals have been considered based upon different mobility models for vehicles.

Type of data handled by a cluster is an issue that is used to identify the application area for a clustering protocol. Since safety message broadcast with very small latency has been considered as one the main objective of VANET, most of the proposals have attempted to optimize them. Some recent protocols have considered other messages like non-real time message transmissions also. Relative direction of a vehicle specifies the direction of vehicles that participate in cluster formation. Since the movement of vehicles in a cluster could be constrained by road condition, so the direction of vehicles can be in the same direction called uni-directional movement or in opposite direction known as bidirectional traffic. Majority of the reviewed clustering protocols have considered uni-directional vehicular movement for cluster formation since it increases the stability of clusters. This is due to the reason that vehicles in a cluster can be in contact for a larger time interval in case of unidirectional clustering with each other as compared to relatively small contact time for bi-directional clustering.

Density is another important aspect which is used to estimate the size of the cluster in VANETs. Most of the exiting protocols have considered this aspect in which density varies from low, medium to high values. However there are no standard values for these levels of densities and different protocols have considered different ranges varying from a minimum value of 2 vehicles/km/lane to a high value of approximately 60 vehicles/km/lane. The issue of vehicle speed is also important and it defines the speed range of vehicles in a cluster. Different clustering protocols have used various values of speeds and it is an important parameter for verifying the performance of the protocols.

Communication mode is also an important issue in which vehicles use the available communication infrastructure. The communication mode between vehicles can be broadcast or its variant like unicast or multicast, store and carry or any other. Most of the existing protocols use broadcasting as the communication mode because of it being supported in most of the underlying channel access protocols. The use of the wireless media for transferring the data between a Cluster head and cluster members also results in some communication cost during clustering. However, due to the lack of communication reliability in VANETs, it is necessary to implement fault-tolerant techniques during the discovery of service providers in VANETs [71–75]. Communication overhead is used to estimate the coverage and connectivity of the nodes which is used for cluster maintenance by the Cluster head. This is an important factor to be investigated because some of the cluster members may be located at the corners of the cluster and have poor connectivity with the CHs so these nodes may not be able to communicate with their respective CHs. The communication overhead has been characterized into three levels identified as low, medium and high [76,77]. Since vehicular networks do not have any problem regarding battery life unlike MANETs, the communication overhead in most of the protocols has been found to be medium. In most of the clustering protocols, metrics such as delay, cluster lifetime, cluster head duration, number of packets transferred in unit time, throughput etc. are considered to compute the effectiveness of the any proposed scheme. Table 17 provides a detailed categorization of all the existing proposals based upon the above mentioned issues.

Table 17
Relative comparison of various clustering protocols.

Protocol (year)	*Clustering parameter*	*Topology*	*Additional infrastructure*	*Road-side scenario*	*Node mobility*	*Type of data handled*	*Relative direction of vehicles*	*Absolute vehicle density*	*Relative speed of vehicles*	*Communication mode*	*Average clustering overhead*
CBLR (2003)	–	Hierarchical	GPS and local coordinate system	Circular scenario	–	–	–	–	–	Broadcast	Medium
COIN (2003)	Vehicular dynamics	–	–	–	–	–	–	–	–	–	Low
LORA-CBF (2005)	Location improvement	–	–	–	–	–	–	–	–	Broadcast	–
Information propagation scheme for VANETS (2005)	Signal strength	–	–	–	Yes	–	–	–	–	–	–
CB-MMAC (2007)	Multichannel MAC communication scheme	Hierarchical	Transceivers	Highway	–	Real time and non-real time data	–	12, 24, 40 (V/km/lane)	20–50 m/s	TDMA-broadcast	Low
CBMAC (2007)	TDMA	–	–	Urban	–	–	–	2, 4, 7 (V/km/lane)	–	TDMA-broadcast	High
DBA-MAC (2007)	Distributed, dynamic clustering	Hierarchical	802.11 Devices, GPS	Highway	–	Critical safety message	Bidirectional	25, 50, 7 per 3 lanes	20–30 m/s	Broadcast	–
VeSOMAC (2007)	*TDMA*	–	–	Highway	–	–	–	–	–	–	–
Clustering for inter-vehicle comm. (2007)	Distance	–	–	Urban	–	–	Bidirectional	–	–	Broadcast	High
M-DMAC (2008)	Mobility	K-clusters	–	Urban	Yes	Safety	Bidirectional	–	11-31m/s	Broadcast	Medium
TrafficGather (2008)	SDMA	–	–	–	–	–	–	–	–	–	–
DISCA (2007)	Direction, leadership duration	Hierarchical	–	Highway & city	–	–	–	–	–	Broadcast	–
Position based technique (2009)	Position of a vehicle within segment	Hierarchical	–	–	–	Safety	–	–	–	–	Medium
APROVE (2009)	Mobility, affinity propogation	–	GPS	Highway	Yes	–	UniDirectional	–	15, 25, 35, 40, and 50 m/s	Broadcast	High
CGP (2009)	SDMA	–	–	–	–	–	–	–	–	Broadcast	–
DBC (2009)	Density	Hierarchical	–	–	–	–	–	–	–	Broadcast	Density
ALM (2010)	Aggregate local mobility	–		Urban	Yes	–	Bidirectional	Yes	10, 15, 20, 25, and 30 m/s	–	Medium
Euclidean distance clustering (2010)	Euclidean distance	Hierarchical	GPS	Highway	Yes	Safety	UniDirectional	10–60 vehicles/km	15, 25 m/s	Broadcast	Medium
NEW ALM (2010)	Position of a vehicle within segment	–	–	–	–	–	–	–	–	–	Medium
Lane based clustering (2010)	Direction of traffic flow	–	Lane detection system, odometer, LIDAR, etc.	Urban	–	–	UniDirectional	60	40–120 km/h (1–33 m/s)	Broadcast	Low

(*continued on next page*)

Table 17 (*continued*)

Protocol (year)	*Clustering parameter*	*Topology*	*Additional infrastructure*	*Road-side scenario*	*Node mobility*	*Type of data handled*	*Relative direction of vehicles*	*Absolute vehicle density*	*Relative speed of vehicles*	*Communication mode*	*Average clustering overhead*
ER-AC (2011)	Epidemic routing and autonomous clustering	Hierarchical	–	Urban	–	–	–	–	10–15 m/s	Store and forward	High
MC-DRIVE (2011)	Direction of traffic flow	–	GPS, digital maps	–	Yes	–	UniDirectional	Low, high	8.5–14 m/s	Broadcast	Medium
HCA (2013)	Randomized, k-hop clustering	Hierarchical	–	Urban	–	Safety and warning		Low, medium, high	15–30 m/s	–	Low
DCA (2011)	Spatial dependency	Hierarchical	GPS	–	–	–	UniDirectional	–	10–20 m/s	Broadcast	–
Multi-hop algo. using Rel. Mob. (2011)	Aggregate mobility	Hierarchical	–	Both	Yes	–	–	–	10–35 m/s	Broadcast	High
D-CUT (2011)	Geographical location, density	Hierarchical	GPS	–	–	–	–	–	–	–	–
LICA (2011)	Location improvement	–	GPS	–	Yes	–	–	–	–	Broadcast	–
VeMAC (2012)	*TDMA*	–	GPS, transceivers	–	–	–	Bidirectional	–	–	–	Medium
VWCA (2011)	Direction based algorithm	Hierarchical	–	Highway	–	–	–	–	70, 120 km/h	Broadcast	low
AMACAD (2011)	Final destination	–	GPS or navigation system	Urban	Yes	–	–	0 to 5 per 100 m^2	5–20 m/s	Combination of broadcast, store and forward	Medium
FUZZY BASED C-H ALGO. (2012)	Average speed difference	–	*GPS*	Highway		Safety messages	UniDirectional	–	20–30 m/s	–	–
CCA (2012)	Relative velocity	–	–	Urban	Yes	–	Bidirectional	Constant	20–35 m/s	Broadcast	Medium
SBCA (2012)	Mobility	–	–	Highway	–	–	UniDirectional	Low, medium, high	25–35 m/s	–	Low overhead
SP-CL (2012)	Relative force, beacon based	–	–	Highway	–	–	–	–	20–45 m/s	–	Low overhead
k-hop clustering approach (2012)	highest connectivity	Hierarchical	GPS	–	–	–	–	–	–	Broadcast	Low overhead
Multi agent driven dynamic clustering (2012)	Relative speed and direction	–	–	Urban	Yes		UniDirectional		10, 40, 60, 80 km/h	Broadcast	–
QuickSilver (2012)	Sequence number of messages	Flat	–	Urban	–	–	–	–	–	Unicacst/multi-cast	–
CDGP (2012)	SDMA	–	GPS, digital maps, interconnected RSU's	Highway	–	–	UniDirectional	–	50–170 km/h	Broadcast	Low
TC-MAC (2012)	TDMA	–	–	Highway	–	Both safety and non-safety	–	5, 12, 21, 50 v/km	–	Multicast	Medium
ALCA (2012)	Direction of traffic flow, mobility	–	–	–	–	–	–	–	30–50 miles/h	–	High

(*continued on next page*)

Table 17 *(continued)*

Protocol (year)	*Clustering parameter*	*Topology*	*Additional infrastructure*	*Road-side scenario*	*Node mobility*	*Type of data handled*	*Relative direction of vehicles*	*Absolute vehicle density*	*Relative speed of vehicles*	*Communication mode*	*Average clustering overhead*
PassCAR (2013)	Passive clustering	Hierarchical	–	Urban and highway	Yes	–	–	10–20 v/km	80–20 km/h	Broadcast	–
ESA (2006)	Secure message aggregation and group communication	Hierarchical	GPS	Urban	–	Safety related	Bidirectional	2–5 v/km	50 km/h	Broadcast	High
SCA (2009)	Distributed PKI based on trust management	–	–	Highway	Yes	Safety and non-safety	Bidirectional	50 v/km	30–140 km/h	Multicast	High

9. Open research issues/challenges

As VANET is a new emerging technology which can be used in wide range of applications, so lot of research proposals exist in literature using VANETs. Although lot of issues have been explored in this paper and many solutions have also been proposed for the same, but still there are many open research issues and challenges which need further investigations. Some of the open research issues are as follows:

- To the best of our knowledge, there exist no proposals in the literature which have considered the clustering with respect to the bidirectional traffic model in which vehicles are assumed to be approaching the RSUs from both directions.
- The applicability of supervised/unsupervised learning techniques for various operations such as Data Dissemination, Routing, QoS guarantee, Security etc. require further investigation. As vehicles are highly mobile so how these techniques can be applicable to the above defined problems with respect to the mobility of the vehicles is an open research problem.
- The use of Symmetric/Asymmetric key management is also an open issue. The important issue in this case is the placement of key, i.e., key can be placed either at the RSUs, or on Gateway or with vehicles. What are the major benefits/drawbacks of placing the key at different locations in VANETs. If keys are shared among the vehicles in Peer-to-Peer (P2P) manner, then the standard Distributed Hash Table (DHT) can be implemented for the same or we require some other mechanism which takes care of key management with respect to the mobility of the vehicles [78–80].
- During the process of clustering, how the horizontal and vertical handoff mechanisms are controlled with respect to the varying mobility and density of the vehicles is another research issue and challenge which needs further investigation from the research community. Apart from the various standards provided by the Mobile IPv6 (MIPv6), users require more enhancements in the existing proposals which make his travel an unforgettable experience.
- Caching has been widely used in wireless networks in which most frequently accessed contents are kept in cache for enhancing the performance of the networks. But as the vehicles are highly mobile in nature, so one of the key research issues is placement of cache, i.e., either in the vehicles, or at nearest RSUs, or at Gateways. The performance of various cache management techniques need to be exploited further so that contents can be accessed without any performance degradation.
- As RSUs are the central part of the VANETs architecture for providing various resources to the moving vehicles, so optimal deployment of these RSUs can also be investigated further. For optimal deployment of RSUs which preserve the coverage and connectivity, various techniques such as Support Vector Machine (SVM), Genetic Algorithm (GA), Fuzzy Logic, Game Theory, Supervised/Unsupervised Learning can be applied.
- Most of the research proposals have considered the dense Urban regions or highways (High Connectivity) scenarios for proposing new solutions, but certain areas may also exist where we have sparse distribution and poor connectivity of the vehicles. For those regions, special investigation is required and solutions should be designed keeping in view of the coverage and connectivity issues.
- To the best of our knowledge, multi-level clustering in VANETs has not been exploited to its full potential. To perform multi-level clustering, various attributes such as direction of the vehicles movement, size and shape of the cluster, radius of the cluster, and distance from RSUs may be considered.

- The applicability of agent technology can also be considered for clustering in VANETs. The agents may be static or mobile and can learn from the environment where they are operating.

10. Conclusion and future directions

Vehicular Ad Hoc Networks (VANETs) are used in wide areas of applications in recent times. Clustering of vehicles has been investigated by the research community from different perspective in many of the applications used in VANETs. This paper provides a complete taxonomy on clustering in VANETs based upon various parameters. Also, a detailed discussion with comparative analysis is provided for each categorization of clustering which includes various challenges, existing solutions and future directions. Each section is described with various clustering techniques and their advantages/disadvantages over the others. The analysis provided for various existing proposals allow various users working in this domain to select one of the proposals with respect to its merits over the others. In the future, we would like to propose a new clustering technique based upon the discussion and analysis provided in this paper. The clustering technique would be adaptive which considers various parameters used for clustering in VANETs.

Acknowledgements

This work has been partially supported by the *Instituto de Telecomunicações*, Next Generation Networks and Applications Group (NetGNA), Portugal, by National Funding from the FCT – *Fundação para a Ciência e a Tecnologia* through the PEst-OE/EEI/LA0008/2013 Project.

References

[1] S. Al-Sultan, M. Moath, Al-Doori, H. Al-Bayatti Ali, H. Zedan, A comprehensive survey on vehicular ad hoc network, J. Netw. Comput. Appl. (2014), in press.
[2] I. Salhi, M. Cherif, S. Senouci, Data collection in vehicular networks, in: Autonomous and Spontaneous Networks Symposium, 2008, pp. 20–21.
[3] Z. Wang, L. Liu, M. Zhou, N. Ansari, A position based clustering technique for ad hoc intervehicle communication, IEEE Trans. Syst. Man Cybern., Part C, Appl. Rev. 38 (2) (2008) 201–208.
[4] W. Fan, Y. Shi, S. Chen, L. Zou, A mobility metric based dynamic clustering algorithm (DCA) for VANETs, in: International Conference on Communication Technology and Application, Beijing, 2011, pp. 752–756.
[5] F. Bai, N. Sadagopan, A. Helm, The IMPORTANT framework for analyzing the impact of mobility on the performance of routing protocols for ad hoc networks, Ad Hoc Netw. (2003) 383–403.
[6] N. Maslekar, J. Mouzna, H. Labiod, M. Devisetty, M. Pai, Modified C-DRIVE: clustering based on direction in vehicular environment, in: IEEE Intelligent Vehicles Symposium, vol. 4, 2011, pp. 845–850.
[7] N. Maslekar, M. Boussedjra, J. Mouzna, H. Labiod, A stable clustering algorithm for efficiency applications in VANETs, in: 7th International Conference on Wireless Communications and Mobile Computing, Istanbul, 2011, pp. 1188–1193.
[8] G. Wolny, Modified DMAC clustering algorithm for VANETs, in: Third International Conference on Systems and Networks Communications, Sliema, 2008, pp. 268–273.
[9] S. Basagni, Distributed clustering for ad hoc networks, in: Fourth International Symposium on Parallel Architectures, Algorithms, and Networks, Perth/Fremantle, WA, 1999, pp. 310–315.
[10] A. Fonseca, T. Vazão, Applicability of position-based routing for VANET in highways and urban environment, J. Netw. Comput. Appl. 36 (3) (2013) 961–973.
[11] F. Ahammed, J. Taheri, A. Zomaya, LICA: robust localization using cluster analysis to improve GPS coordinates, in: First ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications, New York, USA, 2011, pp. 39–46.
[12] D. Tian, Y. Wang, G. Lu, G. Yu, A VANETs routing algorithm based on Euclidean distance clustering, in: 2nd IEEE International Conference on Future Computer and Communication, Wuhan, 2010, pp. V1-183–V1-187.
[13] M.M.C. Morales, C.S. Hong, Y. Bang, An adaptable mobility aware clustering algorithm in vehicular networks, in: 13th Asia–Pacific Network Operations and Management Symposium, Taipei, 2011, pp. 1–6.
[14] Y. Zhang, J. Mee Ng, C. Ping Low, A distributed group mobility adaptive clustering algorithm for mobile ad hoc networks, J. Comput. Commun. 32 (1) (2009) 189–202.
[15] R.A. Santos, R.M. Edwards, N.L. Seed, Inter vehicular data exchange between fast moving road traffic using an ad-hoc cluster-based location routing algorithm and 802.11b direct sequence spread spectrum radio, in: Post Graduate Networking Conference, 2003.
[16] M.S. Almalag, M.C. Weigle, Using traffic flow for cluster formation in vehicular ad-hoc networks, in: 35th IEEE Conference on Local Computer Networks, Denver, CO, 2010, pp. 631–636.
[17] P. Fan, Improving broadcasting performance by clustering with stability for inter-vehicle communication, in: 65th IEEE Conference on Vehicular Technology, Dublin, 2007, pp. 2491–2495.
[18] Z. Zhang, A. Boukerche, R.W. Pazzi, A novel multi-hop clustering scheme for vehicular ad-hoc networks, in: MobiWac'11, 9th ACM International Symposium on Mobility Management and Wireless Access, 2011, pp. 19–26.
[19] L. Zhang, E. El-Sayed, A novel cluster-based protocol for topology discovery in vehicular ad hoc network, in: Third International Conference on Ambient Systems, Networks and Technologies, in: Procedia Computer Science, vol. 10, ACM, New York, USA, 2012, pp. 525–534.
[20] T. Camp, J. Boleng, L. Wilcox, Location information services in mobile ad hoc networks, in: Proceedings of the IEEE International Conference on Communications (ICC), vol. 5, 2002, pp. 3318–3324.
[21] L. Weiwei, T. Ali, L. Alberto, Robust clustering for connected vehicles using local network criticality, in: IEEE Conference on Communications, ICC, Ottawa, 2012, pp. 7157–7161.
[22] E. Dror, C. Avin, Z. Lotker, Fast randomized algorithm for hierarchical clustering in vehicular ad-hoc networks, Ad Hoc Netw. 11 (7) (2013) 2002–2015.
[23] P. Fan, J. Haran, J. Dillenburg, P. Nelson, Cluster based framework in vehicular ad-hoc networks, in: Lecture Notes in Computer Science, vol. 3738, 2005, pp. 32–42.
[24] H. Su, X. Zhang, Clustering based multichannel MAC protocols for QoS provisions over vehicular ad hoc networks, IEEE Trans. Veh. Technol. 56 (2007) 3309–3323.
[25] L. Bononi, M. Di Felice, A cross layered MAC and clustering scheme for efficient broadcast in VANETs, in: IEEE International Conference on Mobile Ad Hoc and Sensor Systems, Pisa, 2007, pp. 1–8.
[26] L. Bononi, M. Felice, L. Donatiello, D. Blasi, V. Cacace, L. Casone, S. Rotolo, Design and performance evaluation of cross layered MAC and clustering solutions for wireless ad hoc networks, Perform. Eval., Int. J. 63 (2006) 1051–1073.
[27] Y. Liu, J. Niu, J. Ma, L. Shu, T. Hara, W. Wang, The insights of message delivery delay in VANETs with a bidirectional traffic model, J. Netw. Comput. Appl. 36 (5) (2012) 1287–1294.
[28] S. Biswas, J. Misic, V. Misic, ID-based safety message authentication for security and trust in vehicular networks, in: 31st International Conference on Distributed Computing Systems Workshops, ICDCSW, Minneapolis, MN, 2011, pp. 323–331.
[29] H.A. Omar, W. Zhuang, L. Li, VeMAC: a TDMA-based MAC protocol for reliable broadcast in VANETs, IEEE Trans. Mob. Comput. 12 (9) (2012) 1724–1736.
[30] Y. Gunter, B. Wiegel, H.P. Grossman, Medium access concept for VANETs based on clustering, in: IEEE 66th International Conference on Vehicular Technology, Baltimore, MD, 2007, pp. 2189–2193.
[31] M.S. Almalag, S. Olariu, M. Weigle, TDMA cluster-based MAC for VANETs, in: IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks, WoWMoM, San Francisco, CA, 2012, pp. 1–6.
[32] C.T. Calafate, G. Fortino, S. Fritsch, J. Monteiro, J. Cano, P. Manzoni, An efficient and robust content delivery solution for IEEE 802.11p vehicular environments, J. Netw. Comput. Appl. 35 (2) (2012) 753–762.
[33] I. Salhi, M.O. Cherif, S.M. Senouci, A new architecture for data collection in vehicular networks, in: IEEE International Conference on Communications, Dresden, 2009, pp. 1–6.
[34] W. Chang, H. Lin, B. Chen, TrafficGather: an efficient and scalable data collection protocol for vehicular ad hoc networks, in: 5th IEEE Conference on Consumer Communications and Networking, Las Vegas, NV, 2008, pp. 365–369.
[35] S.V. Bana, P. Varaiya, Space division multiple access (SDMA) for robust ad hoc vehicle communication networks, intelligent transportation systems, in: IEEE Intelligent Transportation System, Oakland, CA, 2001, pp. 962–967.
[36] B. Brik, N. Lagraa, M.B. Yagoubi, A. Lakas, An efficient and robust clustered data gathering protocol (CDGP) for vehicular networks, in: Second ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications, NY, USA, 2012, pp. 65–74.
[37] C. Campolo, H.A. Cozzetti, R. Scopigno, Augmenting vehicle to roadside connectivity in multi-channel vehicular ad hoc networks, J. Netw. Comput. Appl. (2012).
[38] T. Little, A. Agarwal, An information propagation scheme for VANETS, in: IEEE 8th International Conference on Intelligent Transportation Systems, 2005, pp. 155–160.
[39] S. Teshima, T. Ohta, E. Kohno, Y. Kakuda, A data transfer scheme using autonomous clustering in VANETs environment, in: 10th IEEE International Symposium on Autonomous Decentralized Systems, Tokyo & Hiroshima, 2011, pp. 477–482.
[40] T. Ohta, S. Inoue, Y. Kakuda, K. Ishida, An adaptive multi hop clustering scheme for ad hoc networks with high mobility, IEICE Trans. Fundam. E86-A (7) (2003) 1689–1697.

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004

[41] R.A. Santos, A. Edwards, R. Edwards, L. Seed, Performance evaluation of routing protocols in vehicular ad hoc networks, Int. J. Ad Hoc Ubiq. Comput. 1 (2005) 80–91.

[42] A. Tahsin, R. Prodhan Dasm, H. Kabir, G.C. Shoja, TTL based routing in opportunistic networks, J. Netw. Comput. Appl. 34 (2011) 1660–1670.

[43] L.A. Maglaras, D. Katsaros, Distributed clustering in vehicular networks, in: IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications, Barcelona, 2012, pp. 593–599.

[44] A. Ahizoune, A. Hafid, A new stability based clustering algorithm for VANETs, in: IEEE Local Computer Networks Workshops, Clearwater, FL, 2012, pp. 843–847.

[45] E. Souza, I. Nikolaidis, P. Gburzynski, A new aggregate local mobility (ALM) clustering algorithm for VANETs, in: IEEE International Conference on Communications, Cape Town, South Africa, 2010, pp. 1–5.

[46] C. Shea, B. Hassanabadi, S. Valaee, Mobility based clustering in VANETs using affinity propagation, in: IEEE Conference on Global Telecommunications, Honolulu, 2009, pp. 1–6.

[47] P. Basu, N. Khan, T. Little, A mobility based metric for clustering in mobile ad hoc networks, in: International Conference on Distributed Computing Systems, Mesa, AZ, 2001, pp. 413–418.

[48] R. Ghosh, S. Basagni, Mitigating the impact of node mobility on ad hoc clustering, Wirel. Commun. Mob. Comput. 8 (2008) 295–308.

[49] O. Kayis, T. Acarman, Clustering formation for inter-vehicle communication, in: Intelligent Transportation Systems Conference, Seattle, WA, 2007, pp. 636–641.

[50] A. Yair, M. Segal, Near-optimal, reliable and self-organizing hierarchical topology in VANET, in: Eighth ACM International Workshop on Vehicular Inter-Networking, 2011, pp. 79–80.

[51] S. Kuklinski, G. Wolny, Density based clustering algorithm for VANETs, Int. J. IP Technol. 4 (2009) 149–157.

[52] M. Bakhoya, J. Gaber, P. Lorenz, An adaptive approach for information dissemination in vehicular ad hoc networks, J. Netw. Comput. Appl. 34 (2011) 1971–1978.

[53] M.S. Kakkasageri, S.S. Manvi, Multiagent driven dynamic clustering of vehicles in VANETs, J. Netw. Comput. Appl. 35 (2012) 1771–1780.

[54] A. Daeinabi, A.G.P. Rahbar, A. Khademzadeh, VWCA: an efficient clustering algorithm in vehicular ad hoc networks, J. Netw. Comput. Appl. 34 (2011) 207–222.

[55] X. Liu, Z. Fan, L. Shi, Securing vehicular ad hoc networks, in: Second Conference on International Pervasive Computing and Applications, Amsterdam, The Netherlands, 2007, pp. 424–432.

[56] S. Wang, Y. Lin, PassCAR: a passive clustering aided routing protocol for vehicular ad hoc networks, Comput. Commun. 36 (2013) 170–179.

[57] T.J. Kwon, V.K. Varma, T.R. Hsing, Efficient flooding with passive clustering – an overhead-free selective forward mechanism for ad hoc/sensor networks, Proc. IEEE 91 (8) (2003) 1210–1220.

[58] K.A. Hafeez, L. Zhao, Z. Liao, B.M. Ma, A fuzzy-logic-based cluster head selection algorithm in VANETs, in: IEEE International Conference on Communications, Ottawa, ON, 2012, pp. 203–207.

[59] N. Kumar, N. Chilamkurti, J.H. Park, ALCA: agent learning-based clustering algorithm in vehicular ad hoc networks, Pers. Ubiquitous Comput. 17 (8) (2012) 1683–1692.

[60] R. Bauza, J. Gozalvez, Traffic congestion detection in large-scale scenarios using vehicle-to-vehicle communications, J. Netw. Comput. Appl. 36 (5) (2012) 1295–1307.

[61] R. Crepaldi, M. Bakht, R. Kravets, QuickSilver: application-driven inter-and intra-cluster communication in VANETs, in: Third International Workshop on Mobile Opportunistic Networks, Zürich, Switzerland, 2012, pp. 69–76.

[62] J. Barrachina, P. Garrido, M. Fogue, F.J. Martinez, J.C. Cano, C.T. Calafate, P. Manzoni, VEACON: a vehicular accident ontology designed to improve safety on the roads, J. Netw. Comput. Appl. 35 (2012) 1891–1900.

[63] J. Blum, A. Eskandarian, L. Hoffman, Mobility management in IVC networks, in: Proceedings of the IEEE Intelligent Vehicles Symposium, 2003, pp. 150–155.

[64] S.T. Cheng, G.J. Horng, C.L. Chou, Using cellular automata to form car society in vehicular ad hoc networks, IEEE Trans. Intell. Transp. Syst. 12 (2011) 1374–1384.

[65] B. Ying, D. Makrakis, T. Hussein Mouftah, Privacy preserving broadcast message authentication protocol for VANETs, J. Netw. Comput. Appl. (2012).

[66] J. Blum, A. Eskandarian, The threat of intelligent collisions, IT Prof. 6 (2004) 24–29.

[67] M. Raya, A. Aziz, J.P. Hubaux, Efficient secure aggregation in VANETs, in: Third International Workshop on Vehicular Ad Hoc Networks, 2006, pp. 67–75.

[68] S. Sivagurunathan, P. Subathra, V. Mohan, N. Ramaraj, Authentic vehicular environment using a cluster based key management, Eur. J. Sci. Res. 36 (2009) 299–307.

[69] T. Gazdar, A. Benslimane, Belghith, A secure clustering scheme based keys management in VANETs, in: 73rd IEEE International Conference on Vehicular Technology, Yokohama, 2011, pp. 1–5.

[70] D.A. Rivas, J.M. Barceló-Ordinas, M.G. Zapata, J.D. Morillo-Pozo, Security on VANETs: privacy, misbehaving nodes, false information and secure data aggregation, J. Netw. Comput. Appl. 34 (2011) 1942–1955.

[71] K. Abrougui, A. Boukerche, H. Ramadan, Performance evaluation of an efficient fault tolerant service discovery protocol for vehicular networks, J. Netw. Comput. Appl. 35 (2012) 1424–1435.

[72] P. Fan, P. Sistla, P. Nelson, Theoretical analysis of directional stability-based clustering algorithm for VANET, in: ACMVANET, 2008, pp. 80–81.

[73] Y. Lu, B. Zhou, F. Jia, M. Gerla, Group based secure source authentication protocol for VANETs, in: IEEE Globecom Workshop on Heterogeneous, Multi Hop Wireless and Mobile Networks, Miami, FL, 2010, pp. 202–206.

[74] Z. Rawashdeh, S. Mahmud, A novel algorithm to form stable clusters in vehicular ad hoc networks on highway, EURASIP J. Wirel. Commun. Netw. (2012), 13 pp.

[75] J.T. Isaac, S. Zeadally, J.S. Camara, Security attacks and solutions for vehicular ad hoc networks, J. Veh. Ad Hoc Sens. Netw. 4 (2010) 894–903.

[76] J. Sun, C. Zhang, Y. Zhang, Y. Fang, An identity based security systems for user privacy in vehicular adhoc networks, IEEE Trans. Parallel Distrib. Syst. 21 (2010) 1227–1239.

[77] H. Guo, L. Ngoh, Y. Wu, L. Liow, C. Kwek, F. Tao, J. Ang, Embedded info-security for vehicular networks, in: International Conference on Communications and Networking, Hangzhou, China, 2008, pp. 29–33.

[78] B.F. Ferreira, Joel J.P.C. Rodrigues, J.A. Dias, J.N. Isento, Man4VDTN – a network management solution for vehicular delay-tolerant networks, Comput. Commun. 39 (February 2014) 3–10.

[79] N. Kumar, N. Chilamkurti, Joel J.P.C. Rodrigues, Learning automata-based opportunistic data aggregation and forwarding scheme for alert generation in vehicular ad hoc networks, Comput. Commun. 39 (February 2014) 22–32.

[80] N.G.J.S. Vasco, Joel J.P.C. Rodrigues, F. Farahmand, GeoSpray: a geographic routing protocol for vehicular delay-tolerant networks, Inf. Fusion 15 (2014) 102–113.

Please cite this article in press as: R.S Bali et al., Clustering in vehicular ad hoc networks: Taxonomy, challenges and solutions, Veh. Commun. (2014), http://dx.doi.org/10.1016/j.vehcom.2014.05.004