

A Scatter Search Based Heuristic for Reliable Clustering in Vehicular Ad Hoc Networks

Rejab Hajlaoui^{1,2}, Tarek Moulahi³, Hervé Guyennet²

¹Community College, Hail University, Saudi Arabia.

²FEMTO-ST Institute, Univ. Bourgogne Franche-Comté, Besançon, France.

³Faculty of Science and Technology, Sidi Bouzid, Tunisia.

{rejab.hajlaoui, herve.guyennet, tarek.moulahi}@femto-st.fr

Abstract. Achieving a safe, comfort and autonomous driving in vehicular ad hoc networks (VANET) is the great interest of a large number of researchers and car manufacturers. Despite the variety of the proposed approaches and the development of communications technologies, there are no typical solutions. Indeed, several recent studies prove the practical advantages of heuristic method to solve various problems of optimization. Therefore, we used in this paper a Hybrid Scatter Tabu Search (HSTS) based heuristic approach to assign cluster members (CMs) to convenient cluster heads (CHs). We addressed in this work the cluster formation phase in our Weighted K-medoids Clustering Algorithm (WKCA) proposed recently. The main objective is to derive new solutions from the combination of previous one, including the network coverage as a special criterion. To achieve this objective and to locate the global minimum, we integrate the tabu search in the inner process of the scatter search. To the best of our knowledge, there is no study that uses the scatter search to perform clustering in VANET. By simulation, results show that our scheme improves the network stability in term of several metrics compared with prior approaches.

Key words. Vehicular ad hoc network, communication, heuristic algorithm.

1. Introduction

VANET is an extension of the Mobile Ad hoc NETWORK (MANET), where the node movements are restricted by the topology of the street, traffic signals and obstacles [1]. The purpose of the Intelligent Transportation System (ITS) is to collect and to share information in order to improve the road safety and to provide comfort to travelers. Despite the fact that vehicles have high computing power and adequate storage capacity, the problem of delay and poor connectivity persists. On the other hand, due to the high mobility in highway scenario and the multitude of obstacles in the urban environments, the links can be broken frequently. In this case network maintenance has to be triggered automatically to ensure stability and to send messages within the deadlines. Under these constraints, keeping a high stability of network is an NP-Complete combinatorial optimization problem. It means that all exact algorithms require exponential execution time to resolve it. Thus, they are unsuitable in practice even for moderately large instances [2].

For these reasons, several heuristic based researches have been developed to find efficient and adaptable models for routing problems in high mobility networks. Although obtaining an optimal solution is not guaranteed, the use of heuristic methods gives competitive advantages, the most important of which are:

- Approximate the best solutions for the biggest instances.
- Define the right solutions within a reasonable calculation time.

Indeed, among the ways to improve the performance of an algorithm or to fill some of these gaps is to combine it with another method. This cooperation between approaches makes it possible to exploit their advantages in order to improve the overall performance and to minimize the convergence time. It is obvious that population-based methods are superior in identifying different areas in the search space, while trajectory methods are better in exploring a well-defined area in the search space [3]. In this context, we proposed a Hybrid Scatter Tabu Search model to improve the assignment of ordinary nodes to elected cluster heads. Indeed, the power of this contribution is to unify the advantage of population-based methods with the power of trajectory methods. In this way, there is less chance of missing a good solution compared to population-based methods.

In this work, we treated the nodes that exist in common area as shown in “Fig. 1”. These nodes are covered by several CH and their assignments is a challenge in itself. Several simulations are conducted to understand and to analyze the performance of our model by comparing it against competitive protocols designed for the same objective.

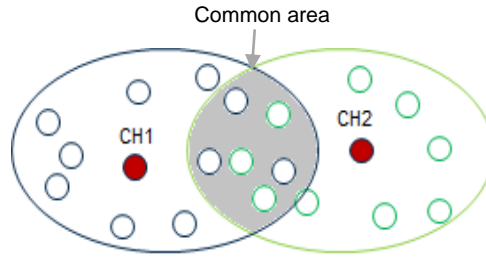


Fig. 1. Overlapping clusters.

The remaining of this paper is organized as follows. In Section 2, we presented some existing solutions. Section 3 and 4 describe our proposal including the theoretical foundation. Section 5 discusses the simulation results. Finally, section 6 concludes the paper and draws some directions for future work.

2. Prior works

Due to the unbounded network size and the dynamic nature of vehicles, the intermittent connectivity is considered as the biggest challenge. Indeed, many researchers used the clustering technique to satisfy the VANET stability and to improve the network life time. For instance, in [4], authors proposed a multi-head clustering algorithm considering a data sharing application in vehicular network.

Each cluster contains a master CH (MCH), and several slave CHs (SCHs). The main goal of using more than one CH is to speed up the data downloading from multiple seed nodes (CHs) based on the bit-torrent downloading mechanism.

In the same context, Little et al. [5] used two CHs to extend the MOBIC (MOBility metrics Clustering)[6] into vehicular environments. The first CH is located at the head and the second is in the tail of a cluster. In [7], authors introduced a new clustering technique which uses velocity and distance to create a stable cluster structure. When more nodes converge to be cluster head, the Cuckoo Search algorithm[8] is triggered to select the super cluster-head which will be at an optimum distance, minimum delay, more network lifetime and high packet delivery ratio. In [9], M. Nasr et al. presented a novel VANET routing algorithm based on clustering technique (CBVRP). This protocol is appropriate for rugged environments, such as deserts scenarios. Indeed, the CBVRP used the vehicle's equipment, velocity and location in the cluster classifications and the CH election process. When the communication is inside a cluster, the CH selected the relay node which leads to the destination. In contrast, when the communication is between clusters, the CH uses the flooding mechanism towards its cluster members in order to find the nearest vehicles able to communicate with the outside. This route is kept until the cluster structure changes. In [10], Hassanabadi et al. presented a novel model called "Affinity PROpagation for Vehicular networks" (APROVE). It uses the Affinity Propagation algorithm in a distributed manner to maximize the similarity $s(i,j)$ between the data point i and its chosen exemplar j . In this algorithm, nodes exchanged two types of messages to make decision on independent clustering:

- Responsibility message $r(i,j)$: sent from i to candidate exemplar j . It indicates how well suited j is to be i 's exemplar.
- Availability message $a(i,j)$: sent from candidate exemplar j back to i . It indicates j 's desire to be an exemplar for i based on supporting feedback from other data points.

In [11], authors developed a new distributed algorithm to build a stable multi-hop cluster suitable for vehicular networks. With minimum number of cluster head, this model contains three techniques that depend only on the positions of both receiver and transmitter nodes. When the GPS signal is lost, the proposed algorithms switched to use the RSS (Received Signal Strength) of the packet received to decide if the packet should be retransmitted. Given the continuing research in this area, several clustering solutions, including mobility-based clustering, cluster-based MAC protocols, topology-based clustering, weight-based clustering and energy-based clustering, have been proposed [11], [12], [13].

3. Theoretical foundation of the proposed model

3.1 Weighted k-medoids clustering approach (WKCA)

The WKCA is a novel clustering algorithm proposed recently [14]. This model based on the k-medoids clustering approach provides an automatic switching from small to large cluster depending on the road conditions. It generates small clusters in dense zone in order to avoid the network congestion, whereas in sparse zone it generates large clusters to ensure large coverage.

A node becomes a cluster head if it has the highest weight based on a several metrics (direction, transmission range, speed and node disconnection frequency). Periodically, each node calculates and sends its weight to the CH. Then, the node having a weight greater than that of the current CH will be announced as the new coordinator. Consequently, the cluster maintenance phase will be triggered immediately. Finally, if a node leaves the coverage area of its CH, it has to join another one. Otherwise, it announces itself as a new cluster head to start its own cluster formation. Indeed, two main phases distinguish this model.

Phase 1 : cluster formation

A vehicle V is assigned to cluster C according to its similarity value (SV).

$$\text{Where } SV = (w_1 * D) - (w_2 * \Delta S) + (w_3 * \Delta P) \quad (1)$$

- D : Boolean variable. It indicates the direction of the node compared to the CH.
- ΔS : difference in speed (compared to the CH).
- ΔP : proximity to CH.

In which w_1 , w_2 and w_3 reflect the relative importance of D , ΔS and ΔP .

Phase 2 : cluster head switching

A vehicle V is elected as a new CH if its weight (W) is greater than the weight of the current CH, provided that its behavior in the past was not wobbling or suspicious. Indeed, the weight (W) is calculated based on four metrics (direction, transmission range, speed and node disconnection frequency).

3.2 Scatter Search (SS)

The Scatter Search is a population-based metaheuristics. It is suitable to solve a wide range of optimization problems including Routing Protocol, Traveling Salesman Problem (TSP), and clustering [15], [16]. Recent studies demonstrate the practical advantages of this approach compared to similar heuristics. The most important is that the SS used an adaptive and an associated memory to be adapted to particular contexts [17][18]. Basically, the SS starts with a set of feasible solutions. At the next step, some of these solutions are extracted and combined. The resulted offspring solutions will be enhanced according to an improving procedure. By the end, these new feasible solutions are evaluated according to some criteria to be included or not in the collections. These steps are presented in the Basic Scatter Search Algorithm below [19].

Basic Scatter Search Algorithm

Input: population of the problem.
Output: the best of solutions.

- (1) Initialize the population Pop using a Diversification Generation Method.
- (2) Apply the Improvement Method to the population.
- (3) Reference Set Update Method (Good solutions for RefSet1 and Diversity solutions for RefSet2).
- (4) **While** ($itr < MaxItr$) **do**
- (5) **While** (Reference set is changed) **do**
- (6) Subset Generation Method
- (7) **While** (subset-counter $< > 0$) **do**
- (8) Solution Combination Method.
- (9) Improvement Method.
- (10) Reference Set Update Method;
- (11) **End while**
- (12) **End while**
- (13) **End while**
- (14) Return the best of solutions

3.3 Tabu Search (TS)

Proposed by Glover in 1986, the tabu search is a local search heuristic used to solve complex and large problems. It is an iterative approach starting with an initial feasible solution. Then, from a given position the procedure moves step by step to explore the neighborhood and to select the one that minimizes the objective function. At each iteration, the algorithm chooses the best neighbor not tabu, even if it degrades the cost function. For this reason, the tabu research is known as an aggressive method. The process continues until stopping criteria is met. Unlike other methods, the TS overcomes the problem of local optima using adaptive memory. The basic idea is to save the solutions that are temporarily forbidden to avoid cyclic movements [20].

4. Proposed model

The proposed solution that we have developed is based on the clustering technique. The main challenge is to find the most appropriate cluster for all node in the common area. Indeed, at an instant “t”, a vehicle V has to be assigned only to one cluster head. In this work, each cluster is described as a graph $G(V, E)$, as shown in “Fig. 2”.

- V: represents the set of vehicles defined by an index i , $i \in \{1, 2, 3, \dots, N\}$.
- C: represents the set of cluster heads defined by an index j , $j \in \{1, 2, 3, \dots, M\}$.
- E: represents the set of edges. Each edge E_{ij} designs the similarity value (SV) of the node i to the cluster head j .

The goal is to maximize the similarity within clusters. It is calculated by the summation of the costs of all edges. In our case, for all pairs of nodes $\{i, j\}$, the cost's SV_{ji} and SV_{ij} are equal, then the problem is said to be symmetric.

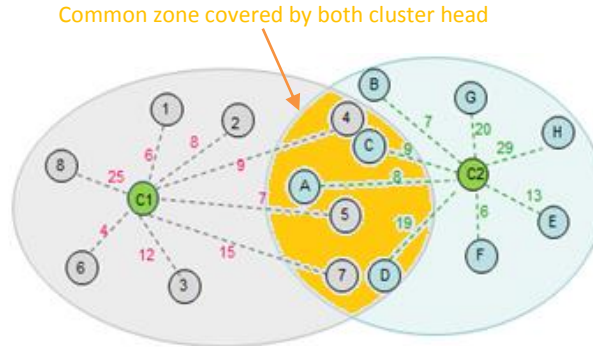


Fig. 2. Clusters representation

4.1 Initialization (by user)

At the beginning, the user has to define values of the starting parameters

- NI: Number of Iterations.
- NII: Number of Iterations without Improvement.
- NIS: Number of Initial Solutions.

4.2 Reference set generation

This method is the first step in the SS process. Indeed, the set of starting solutions are not randomly generated, but obtained from the WKCA model detailed in section 3.a.

4.3 Evaluation

To know the quality of any potential solution, we used the fitness function (F). It indicates how similar the nodes of a cluster are compared to their cluster head.

$$F = \sum_{j=1}^M \sum_{i=1}^N SV_{ij} \quad (2)$$

4.4 Subset generation

Several methods are used in the literature to generate the subsets. In our solution, we treated the nodes that can be assigned to more than one cluster as shown in “Fig. 2”.

- Let B1 the best solution obtained by the WKCA algorithm.

C1	1	2	3	4	5	6	7	8
C2	A	B	C	D	E	F	G	H

- Let B2 the feasible solution very different from the solution B1. (B2 is obtained by switching the nodes of cluster 1 of common area to the cluster 2 and vice versa).

C1	1	2	3	6	8	A	C	D
C2	4	5	7	B	E	F	G	H

- Set diff: the difference between the two solutions. It represents the average of nodes switched from C1 to C2, and vice versa.

$$Diff = \text{number of permuted nodes} / \text{total number of nodes} = 6/16 = 38\%$$

4.5 Combination method

This method is used to combine the elements of the subsets in order to form new solutions. In our case, we used the Path-Relinking combination method (PR) [18] to generate new trial solutions. By applying the PR method, each node in the common area is evaluated comparing to all reachable cluster heads. Then the best assignment with high similarity value is selected. To maintain a fair distribution, the density of the two clusters has to be very close and the difference should not exceed a given value “ μ ”. In other words, we avoid having dense clusters compared to other one, resulting in poor bandwidth exploitation and frequent collisions.

For instance, based on the similarity values of border nodes as shown in “Fig. 3”, it is more suitable to assign node 4 to cluster C2 instead of cluster C1. Likewise, it is better to move the node D from cluster C2 to cluster C1.

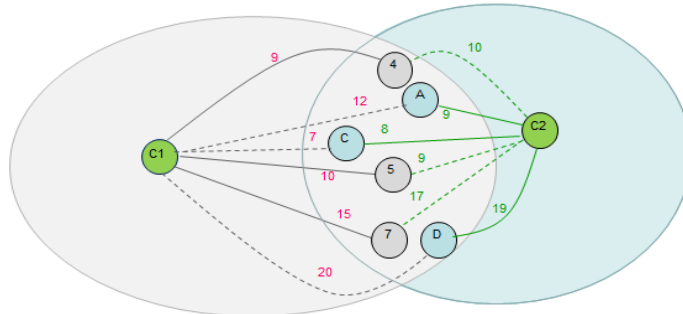


Fig. 3. Similarity values of border nodes

After applying the PR combination method to all nodes, we get $C1^*$ and $C2^*$ from $C1$ and $C2$, as shown in “Fig. 4”.

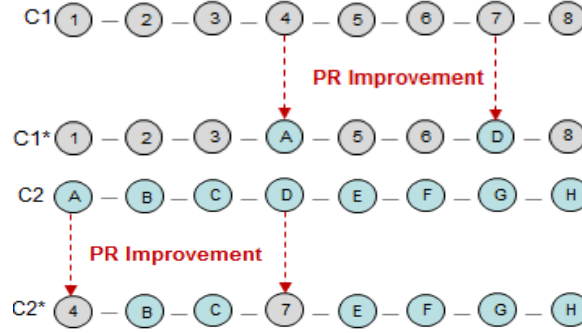


Fig. 4. Solutions generated by Combination method

4.6 Improvement method

In this step, we used the tabu search to transform the solutions obtained by the combination method into more efficient solutions. In our implementation, the TS treated the border nodes. If the improved solution increases the global similarity within the cluster, then it will be included in the reference set, otherwise, it will be ignored. To avoid the loss of time in case of large number of neighborhood, the selected number of iterations does not exceed 10. Indeed, the “Fig. 5” resumes the steps of the proposed Scatter Tabu Search model as well as the methods associated with it (Path Relinking, WKCA, TS, SV).

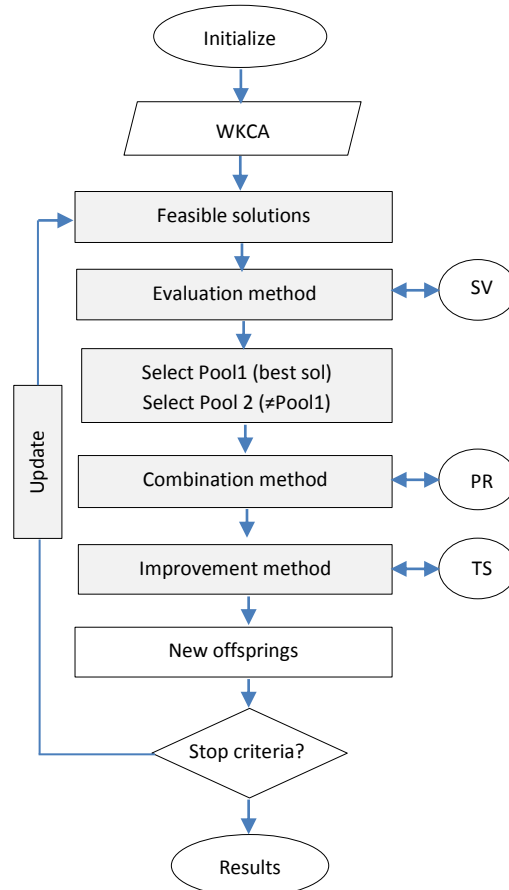


Fig. 5. Flow chart of the proposed model.

5. Experimental results and analysis

5.1 Simulation setup

To evaluate the performance of our algorithm, simulations have been carried out using NS3 [21]. We have set experimentations in highway scenario by varying the density from 40 to 240 vehicles per km and the speed from 20 to 120 kmph. The table below provides the used simulation parameters.

Table 1. Simulation setup.

Parameters	Values
Simulator	NS3
Scenario	Highway.
Nb of lanes	Two-lanes.
Number of vehicles	40 .. 240.
Speed	20-120 Kmph.
Vehicle distribution	Randomly
Transmission range	200m, 300m and 500 m.
Simulation Time	1500 s.
MAC Protocol	DSRC.
Data Type	Constant Bit Rate (CBR).

5.2 Simulation Results

In order to evaluate the improvement given by the Hybrid Scatter Tabu Search in vehicular networks (HSTS), we have performed a careful analysis according to several metrics [22]. Our contribution is compared with the basic Weighted K-medoid Clustering Algorithm (WKCA). Indeed, as we have explained before, the output of the WKCA is used as the initial Reference Set of the Scatter Search. This comparison will show us if the additional work that we have added will improve the results obtained, especially in terms of delay.

a. Packet delivery ratio (PDR): is defined as the ratio of packets successfully delivered to the destination compared to the number of packets sent out by the source

$$PDR = \frac{\sum received\ packets}{\sum sent\ packets} * 100 \quad (3)$$

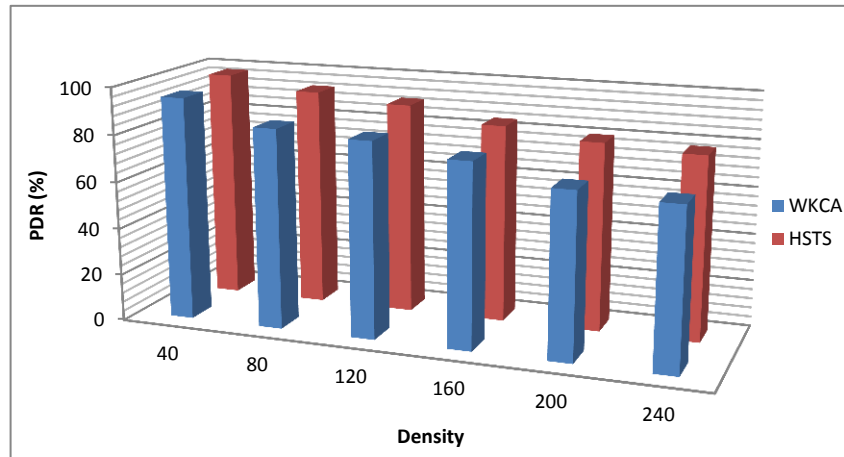


Fig. 6. Density VS PDR

As observed in “Fig. 6”, the HSTS outperforms the WKCA in term of PDR for different densities. The reason behind this superiority is that the obtained clusters are highly similar. Therefore, there is less chance of route failure and the data reaches the destination without any shortage. Indeed, for both models, the high density will create a flood of data to find the appropriate route from source to destination. This broadcast storm will lead to congestion and some packets will be dropped by the master node. For instance, the abandoned packets by the WKCA reached 35% in dense zone. Whereas, the HSTS maintains its performance regardless of the network status and the average of dropped packet does not exceed 22%. Therefore, about 80% of packets reach the destination even in bad traffic condition.

b. End-to-end Delay: is defined as the needed time to send packet from source to destination.

$$Delay = \frac{\sum(PRT - PTT)}{\sum RP} * 1000 \quad (4)$$

- PRT: Packet Reception Time.
- PTT: Packet Transmission Time.
- RP: Received Packet.

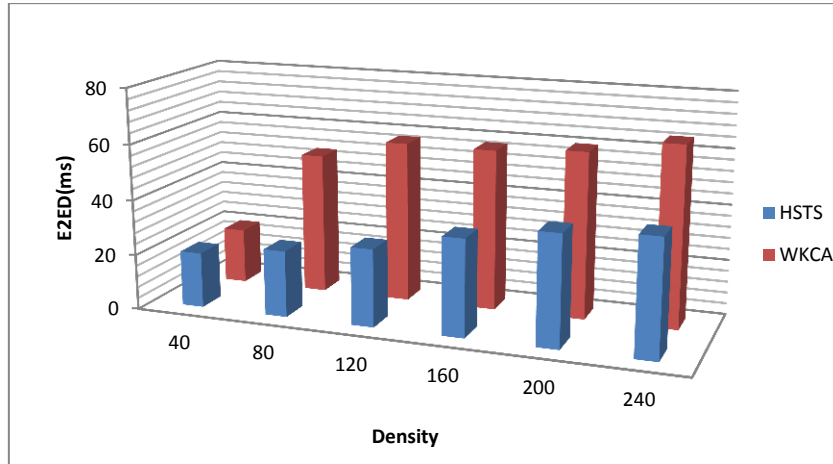


Fig. 7. Density VS E2ED

“Fig. 7” presents a comparison between the HSTS and the WKCA protocols in term of E2ED. The result shows that our enhanced model provides less time amount of propagation delay whether for high or low density. Indeed, when the similarity of nodes within clusters is not optimized, the linking path between the source and the destination will contain more relays. This excess number of nodes causes more delay with high bandwidth usage. For example, with low density (40), when the speed reaches 120 kmph, the delay decreases significantly to reach 20 ms for all approaches. However, when the speed decreases and the density becomes the main challenge (240), the delay in the WKCA increases dramatically to reach 67 ms. While the HSTS maintains a high performance and the delay does not exceed 42 ms in the worst case.

c. Throughput: is defined as the rate of successful data delivered over a communication channel.

$$\text{Throughput} = \frac{\sum \text{Packet successfully received by destination}}{\text{Unit time}} \quad (5)$$

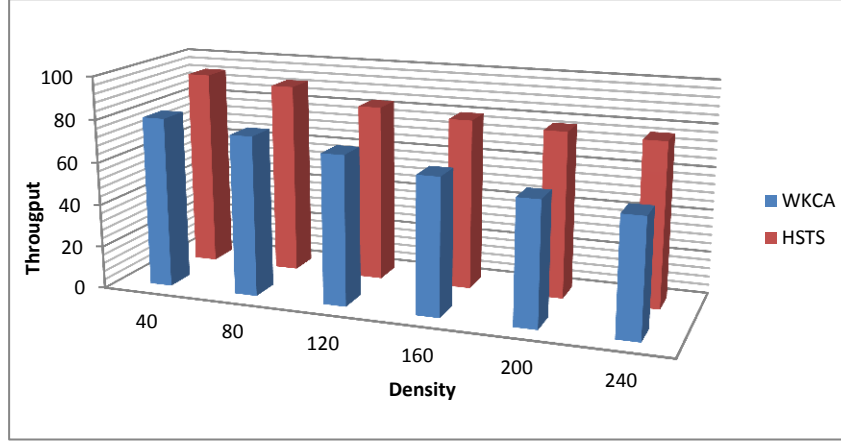


Fig. 8. Density VS Throughput

“Fig. 8” compares the impact of the density on the throughput for all schemes. Indeed, the WKCA protocol has the lowest throughputs especially in dense situation due to the inability of the available bandwidth to cope with the large volume of control packets. To deal with this challenge of density, the HSTS formed clusters with high similarity. This procedure will create direct paths with minimum number of relays. Therefore, the throughput has been improved with the scatter search. Mostly, the traffic crowding is a nonlinear function. Thus, any reduction in traffic stream reduces the collisions and allows a significant rise in throughput. With the best road condition, the situation changes rapidly and the rate of successful delivered messages over a communication channel increases gradually. For instance, by reducing the density from 120 to 80 vehicles, the throughput of both HSTS and WKCA is improved by respectively 10% and 7%.

6. Conclusion

Due to the diversity of routing environments (residential, urban, highway,...) and the high mobility of vehicles, constructing stable networks seems to be the basic challenge. To tackle this problem, we improved the weighted k-medoid clustering algorithm (WKCA) proposed recently. We used a hybrid scatter tabu search to maximize the intra-cluster similarities. We treated the overlapping clustering where a node can simultaneously belong to more than one cluster. As shown in section 5, the proposed model (HSTS) outperforms the basic WKCA in term of PDR, E2ED and throughput in all situations. Finally, we can conclude that the greater the similarity within a cluster, the better the network stability. As a future work, we plan to improve

the behavior of our model based on the Cuckoo Search heuristic especially in urban scenario with the existence of intersections and obstacles.

References

1. R. Hajlaoui, H. Guyennet and T. Moulahi, "A Survey on Heuristic-Based Routing Methods in Vehicular Ad-Hoc Network: Technical Challenges and Future Trends", *IEEE Sensors Journal*, Vol 16 Issue 17, 2016.
2. F. Aadil, K.B. Bajwa, S. Khan, N.M. Chaudary, A. Akram: CACONET: Ant Colony Optimization (ACO) Based Clustering Algorithm for VANET. *PLoS ONE* 11(5), 2016.
3. EG. Talbi: A Taxonomy of Hybrid Metaheuristics. *Journal of Heuristics*. 2002.
4. L. Shou-Chih, Y-J. Lin, J-S. Gao: A Multi-Head Clustering Algorithm in Vehicular Ad Hoc Networks. *International Journal of Computer Theory, Engineering*, April 2013.
5. T. D. C. Little, A. Agarwal: An information propagation scheme for VANETs. *IEEE Intelligent Transportation Systems Conf.*, pp. 155-160, September 2005.
6. P. Basu, N. Khan, T. D. C. Little: A mobility based metric for clustering in mobile ad hoc networks. *Int. Conf. Distributed Computing Systems Workshop*, pp. 413-418, April 2001.
7. A. Malathi, N. Sreenath: An Efficient Clustering Algorithm for Vanet. *International Journal of Applied Engineering Research* Vol 12, Num 9, pp. 2000-2005, 2017.
8. XS. Yang, S. Deb: Cuckoo search via Le'vy flights. *IEEE Proceedings of world congress on nature & biologically Inspire Computing*, India, pp. 210-214, 2009.
9. M. M. M Nasr, A. M. S. Abdelgader, Z-G. Wang, L-F. Shen: VANET Clustering Based Routing Protocol Suitable for Deserts. *Sensors journal*, 2016.
10. B. Hassanabadi, C. Shea, L. Zhang, S. Valaee: Clustering in vehicular ad hoc networks using affinity propagation. *Ad Hoc Network journal*, 2014.
11. C. Shea, B. Hassanabad, S. Valaee: Mobility-based clustering in VANETs using affinity propagation. *The 28th IEEE Conference on Global Telecommunications*, 2009.
12. A.S.K. Mammu, U. Hernandez, N. Sainz: Cluster based MAC in VANETs for safety applications. *International Conference on Advances in Computing, Communications, Informatics (ICACCI)*, pp.1424-1429, 2013.
13. A. B. M. Moniruzzaman, Md. S. Rahmann: Analysis of Topology Based Routing Protocols for Vehicular Ad-Hoc Network (VANET). *International Journal of Computer Applications*, 2014.
14. R. Hajlaoui, T. Moulahi, H. Guyennet : A Weighted k-Medoids Clustering algorithm for effective stability in Vehicular Ad Hoc Networks. *19th International Conference on Communication Systems, Applications, Venice, Italy*, 2017.
15. S. Scheuerer, R. Wendolsky: A scatter search heuristic for the capacitated clustering problem. *European Journal of Operational Research*, pp. 533-547, 2006.
16. D. S. Yamashita, V. A. Armentano, M. Laguna: Scatter search for project scheduling with resource availability cost. *European Journal of Operational Research*, pp. 623-637, 2004.
17. S. Scheuerer, R. Wendolsky: A scatter search heuristic for the capacitated clustering problem. *European Journal of Operational Research*, pp. 533-547, 2006.
18. F. Glover, M. Laguna, R. Marti: Fundamentals of scatter search and path relinking. *Control and Cybernetics Journal*. Vol 39 pp. 653- 684, 2000.
19. M. Laguna, R. Martí: Scatter Search: Methodology, Implementations. *Kluwer Academic Press*, 2003.
20. F. Glover, E. Taillard, M. Laguna, D. de Werra: Tabu Search. *Annals of Operations Research*, 1993.
21. The Network Simulator - <https://www.nsnam.org/>.
22. J. Agrawal, A. Singhal, R.N Yadav: Multipath routing in mobile Ad-hoc network using meta-heuristic approach. *IEEE International Conference on Advances in Computing, Communications, Informatics (ICACCI)*, India, 2017.