

# A Multiple Classifiers Broadcast Protocol for VANET

Sami S. Alwakeel, Hesham A. Altwaijry  
Department of Computer Engineering  
CCIS – King Saud University  
Riyadh, Kingdom of Saudi Arabia  
{swakeel, twaijry}@ksu.edu.sa

Agung B. Prasetyo  
Department of Computer Engineering  
Faculty of Engineering – Diponegoro University  
Semarang, Indonesia  
agungprasetyo@ce.undip.ac.id

**Abstract**—Many types of artificial intelligent machines have been used for decision making purposes. In VANET broadcast protocols, vehicles must decide the received messages are to be rebroadcast or not. Several attributes such as sender-to-receiver distance, sender-receiver speed difference, number of neighboring vehicles, as well as vehicle's movement direction are important measures to take the broadcast decision. As the relationships of attributes to the broadcast decision cannot be mathematically defined, the use of a classifier-based artificial intelligence may approximately predict the relationships of all the incorporated attributes to such a decision. As the decision is based on prediction, the use of multiple classifiers in decision making may increase accuracy. Therefore, this research employs a combined-classifiers at an abstract level to provide firmer broadcast decisions on VANET. Our research results justify that the performance of our combined multiple-classifiers outperformed a single-classifier scheme. The multi-classifiers scheme contributes to an average increase of 2.5% in reachability compared to that of the efficient counter-based scheme (ECS). The combined multi-classifiers scheme also improves the saving in rebroadcast tries by 38.9%.

**Keywords**—Broadcast-storm, classifier, VANET, vehicular attribute.

## I. INTRODUCTION

An efficient broadcast has always been a hot issue in broadcast protocol area. Several schemes have been available, from heuristic (e.g. probability-based, counter-based broadcast) to topology-based broadcast (e.g. distance-based broadcast). However, most of the solutions have used mostly only a few attributes (whether local or global), such as the use of sender-to-receiver distance, number of message duplicates received, or even only employing probability to reduce the number of nodes/vehicles that rebroadcast messages to mitigate the broadcast-storm problem (the massive message redundancy, contention and collision) [1, 2, 3, 4].

In reality, considering many attributes in the broadcast decision mechanism may lead to a more efficient broadcast scheme. For example, a vehicle having a greater distance from the sender vehicle is more potential to rebroadcast messages than that of having a smaller distance. Likewise, a vehicle that has a higher speed differential to the sender is considered to be a better broadcast candidate as it will go out from the sender's

radio coverage fast. The number of neighboring vehicles can also be used to select the rebroadcast candidates. The denser the neighbors, the smaller the probability for a vehicle to rebroadcast. Therefore, a multiple-attributes scheme are more probable to outperform a single-attribute scheme if such attributes are properly treated.

To properly handle the attributes, a classification algorithm (known as a classifier or an expert) can be used to examine all the possible situations of the attributes dealing with current vehicular network situation. For example, a greater distance threshold should be applied for vehicles deserved rebroadcast in a dense network. However, a smaller threshold is required to maintain high network reachability. A classifier is able to recognize the input conditions of the attributes and to make decisions based on the knowledge obtained from prior training (called as model). Our work employs the following attributes: sender-to-receiver distance, number of message copies, vehicular density, as well as speed differential and movement direction.

Discussion of the research are presented in what follows. Research in broadcast protocols are presented in section 2. Sections 3 and 4 discuss how our experiments are set up and experimented. Results of the study are presented and discussed in section 5 and conclusions can be found in section 6.

## II. THE BROADCAST-STORM MITIGATION SCHEMES

One simple method to reduce the broadcast-storm is to use a probabilistic approach. The probabilistic based scheme uses probability mechanism for node selection rather than using a threshold mechanism (such as in distance-based threshold) for determining rebroadcast nodes. Basic broadcast techniques in VANETs follow either a  $l$ -persistence or a  $p$ -persistence scheme. The  $l$ -persistence scheme has the advantages of low complexity and high penetration rate, but creates massive redundancy. The  $p$ -persistence scheme may reduce message redundancy but may increase in total latency and degraded penetration rate. For example, literature [5] proposed three schemes: weighted  $p$ -persistence, slotted  $l$ -persistence, and slotted  $p$ -persistence broadcast schemes, whilst literature [6] proposed an adaptive probabilistic based scheme that senses idle channel time to represent the broadcast probability.

An entity or a node attribute (e.g. distance, speed, number of message duplicates) can also be used to determine rebroadcast nodes. Distance/location-based approach uses relative position between sending and receiving nodes in determining broadcast nodes. Literature [1] proposed four schemes dealing with message broadcasting: counter-based, distance-based, location-based and cluster-based schemes. Literature [7] provided a slight modification on the distance based broadcast using the Distance-to-Mean (DTM) approach to calculate additional coverage by calculating the distance to the mean of the sending nodes. Node density based approach is when the decision for a broadcast depended on the number of neighbors available. Variation on the signal coverage, either by changing the power or the antenna to achieve a desired benefit will be included in this approach. With this, the number of neighbors can be adjusted or a desired neighboring nodes topology is obtained. Attempts to modify physical communication media were also carried out as in literature [8, 9].

A mobility based approach is when the speed and movement direction are used as a way for nodes being grouped or segmented. Literature [10] experimented on the use of vehicle's speed to relate to node's rebroadcast probability. They considered that speed can be a representation of vehicle's number of neighbors, in particular in a highway setting. Literature [11, 12] used the speed and direction to evaluate a node for broadcast privilege. Other research as in [13] employed fuzzy logic to select broadcast nodes based on vehicle's speed and distance between sending and receiving nodes. Literature [14] used three vehicular attributes: sender-receiver distance, speed and movement direction in order for selecting the best candidates from its neighboring nodes that are having prospective future moving trends. Simply stated, this scheme looks for the fastest nodes reaching out of the sender's radio coverage for broadcast.

In message duplicate approach, number of message duplicates received is used to select the broadcast nodes. More duplicates heard by a node means lower contribution for a node to have similar message rebroadcasted. Literature [1] proposed a counter-based scheme as a way of selecting broadcast nodes. Every neighboring node receiving a message from a sender sets a waiting time prior to broadcasting. If the duplicates received are less than a predefined value, the node deserves broadcasting the message. Otherwise, the message will be discarded as the additional broadcast will only contribute to an insignificant additional coverage. Literature [2, 4] proposed adaptive approaches to address adaptive functionality in the counter-based broadcast. Literature [15] made the counter-based adaptive by examining all the inter-arrival time of duplicates recorded immediately after the waiting time lasts. Literature [16] took into account the counter-based scheme with several parameters: network size, transmission range and vehicle density to determine broadcast probability. Our study employs multiple vehicular attributes and let a combined classifiers performs the decisions for broadcasting purposes.

#### A. Experiment Setup

Our experiment is set up on a 1,500 x 1,500 meter-square flat topology area with realistic mobility is generated by VanetMobiSim mobility generator [17]. The VanetMobiSim generates realistic vehicular mobility traces that are, in turn, used by the NS-2 network simulator [18] for its wireless network simulation. The NS-2 simulates the MAC 802.11 protocol. The radio range for each vehicle is set to 250 meters CSMA single channel. In DSRC [19], this channel often refers to CCH channel or channel 178.

Messages were generated with Poisson distribution at a rate of 1 message/second throughout the simulation which ends at a sufficient time for the last message generated propagates over the network. The messages were randomly generated over the available vehicles on roads, one per vehicle. It is worth mentioning, the message generation was started 5.0 seconds after the simulation began to facilitate the vehicles sensing the environment to maximize attribute values they can obtain (e.g., maximum number of neighbors). The number of neighbors and speed differential are normalized to its maximum values experienced by individual vehicle over the simulation prior to examining them for their joint probabilistic value, while the sender-receiver distance is normalized to the vehicle's radius of radio coverage. Beaconing is set to 200 millisecond containing current topological position/location, speed vector and the unique identity of the vehicles. Table 1 shows parameters setting used throughout the experiments.

TABLE I. SIMULATION PARAMETERS SETTING

Parameter	Value
MAC type	802.11
Routing protocol	Message passing (DumbAgent)
Bandwidth	10 Mbps
Interface queue type	Queue/DropTail/PriQueue
Radio propagation	Two ray ground
Antenna model	Omnidirectional
Transmission range	250 meters
Network density (N)	50 – 175 vehicles
Simulation time	N+30 seconds

To assign more chances to potential vehicles for message rebroadcast, probability differentiation is applied to their attributes. Table 2 shows the mentioned vehicle's attribute probability assignment. These probability values are chosen after several trials have been made. Note that all the applicable attribute settings are normalized to its maximum values experienced by individual vehicle throughout simulation. Even though the distance and differential speed are continuous

attributes, we prefer to discretize and treat them as discrete variables.

TABLE II. PROBABILITY ASSIGNMENT TO VARIOUS ATTRIBUTES

Attribute	Range	Probability
Sender-receiver distance	0.0 – 0.44	0.3
	0.45 – 0.64	0.5
	0.65 – 1.00	0.8
Neighboring density	0.0 – 0.14	0.8
	0.15 – 0.44	0.5
	0.45 – 1.00	0.3
Speed differential	0.0 – 0.54	0.3
	0.55 – 0.74	0.5
	0.75 – 1.00	0.8
Message copies heard	1 – 4	0.8
	5 – 6	0.5
	>= 7	0.1
Directional conformity	true	0.4
	false	0.6

The probability settings above are used to train classifiers/models in Weka [20]. Such settings shown in Table 2 has been transformed and written to the Weka’s .arff file. The .arff file is the format used by Weka for stating the training inputs.

### B. Performance Measures

The information to acquire is the reachability (RE) and the saved-rebroadcast (SRB). Reachability here is defined as the fraction of vehicles in the network that receive the broadcast message. That is, the number of vehicles receiving the broadcast divided by the total number of vehicles that are reachable, directly or indirectly, from the source node. The packet reachability probability is the representation of network connectivity.

### C. The Five Attributes

There are five vehicle's attributes that are considered in this research, and they are:

1. Distance between sending and receiving vehicles – Greater distances contribute to a wider coverage. Therefore, vehicles having greater distances must have a higher chance for message rebroadcast.
2. Number of neighboring vehicles (vehicle density) – Having more neighboring vehicles means the vehicles should have smaller chances for rebroadcast to suppress message redundancy. Therefore, a higher number of neighboring vehicles means a lower probability for rebroadcast.
3. Number of similar message heard during waiting time – More duplicates heard by the vehicles may reveal a small need for the vehicles to rebroadcast. Therefore, having a large number of duplicates means lower probability for rebroadcast.

4. Speed differential between sending and receiving vehicles – Without considering vehicle’s direction, a large difference in speed may contribute to the rapid dissemination of information. The larger the speed differential is, the more chance the vehicle has to rebroadcast.

5. Directional conformity between sending and receiving vehicles – Specific applications demand directional conformity, such as alerts from an ambulance must go forward. However, other applications might need the message to go backward, such as for accident alerts. In normal message dissemination in certain topology scenarios, a specific direction may contribute more to network performance.

### D. Handling Vehicle’s Attributes

When a new message is received, the vehicle of interest records its distance from the sending vehicle, speed differential, its relative movement direction, and its neighboring vehicle density. The vehicle schedules the rebroadcast of the message with a random wait time (Random Assessment Delay–RAD) say,  $T_{max}$ . If a duplicate is received during the waiting, the vehicle recalculates the distance and its differential speed  $s$  against the current sender. The vehicle will record the smallest distance and speed experienced during the RAD timer. Number of message copies that arrive will be recorded. When the RAD expires (i.e.  $T_{max} = 0$ ), the vehicle assigns the probability values according to Table 2. Note that the distance, speed, and neighboring density values are normalized to the respective maximum values experienced by the vehicle throughout simulation (see Figure 1).

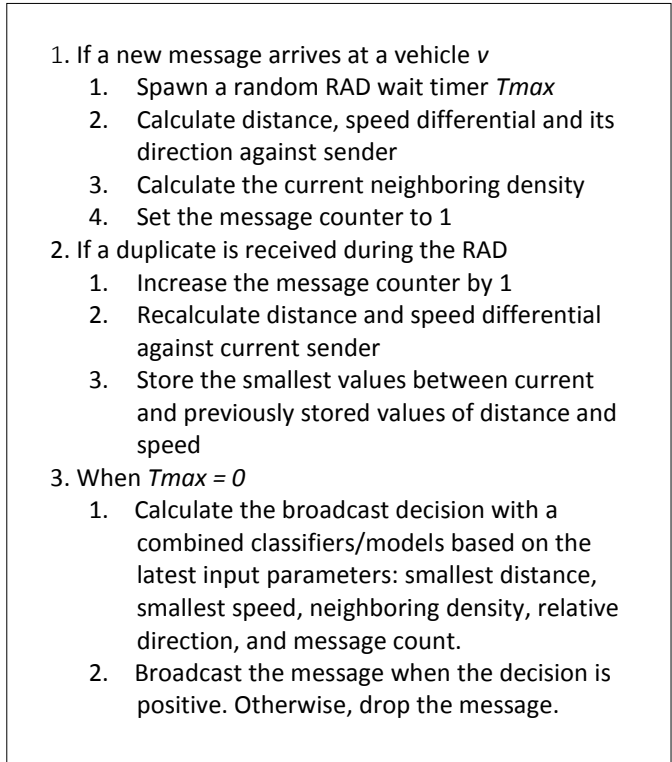


Fig. 1. The multiple classifiers broadcast algorithm

The decision whether to rebroadcast the message is performed by a combination of classifiers that assess the five attributes probabilities. The output of the classifiers are commonly a binary decision class: 1 and 0, or 'yes' and 'no'.

#### IV. EXPERIMENT ON MULTIPLE-CLASSIFIERS BROADCAST PROTOCOL

##### A. Selection of Classifiers for Broadcast

Several models have been selected and experimented with the scenario mentioned above at the density of 75, 125 and 175 vehicles and different random number generator seed for generating random length back-off timer. Those models are: naive-Bayes, J48, Kstar, Multi-layer Perceptron, and RBF-Network. The Weka's models/classifiers above have been trained and used to independently conduct broadcast decisions. The reachability of each model is presented in Figure 2. It is shown that the reachability (RE) of naive-Bayes, Multi-layer Perceptron and RBF-Network models are comparable to that of the simple-flooding scheme. The rest two, J48 and KStar models fail in providing sufficient reachability.

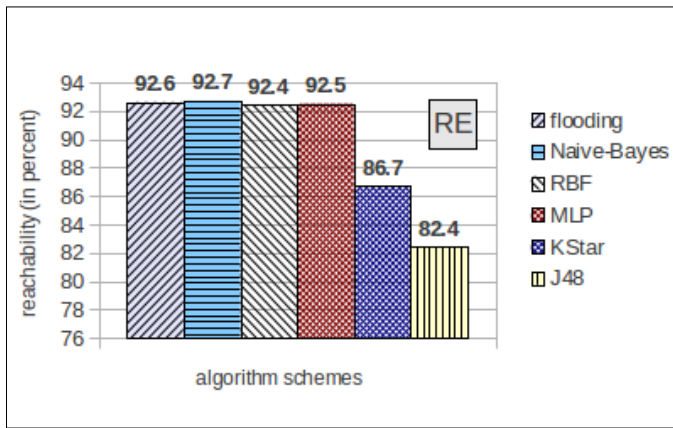


Fig. 2. Average reachability for various network density and different random number generator seed over various models

As suggested by Figure 2, potential models for being used for the next experiments are naive-Bayes, Multi-layer Perceptron and RBF-Network. We incorporate those three models for further experiments.

##### B. Abstract Level Combiner

As Weka's models provide only a predicted class number and its confidence probability value of the class prediction for being correct, the combination of classifiers is conducted in the abstract level outputs as follows:

1. Unanimity function. This function combiner combines the class output from the three classifiers and treats them equally in such a way that the positive decision on rebroadcast is made when all the classifiers output the similar positive decision. The unanimity combiner can be imagined as an AND gate having three inputs. This unanimity can be viewed as to providing more suppression on the naive-Bayes classifier with the use of the other two classifiers.

2. Simple-majority function. The decision will be based on the three classifiers give the most of the class outputs. Here, at least two positive decisions are required to make a vehicle rebroadcast a message.

3. Distributivity function. This function tries to reduce the broadcast suppression on the best among the three classifiers (i.e. naive-Bayes) by providing more probability to the combiner. The positive decision to rebroadcast is made when either the best classifier provides a positive class or when the rest two classifiers output the same positive class regardless of the former classifier's decision.

4. Greatest-average-confidence function. The decision will be selected from the highest averaged confidence probability value for each class when the three classifiers output different class values. For example, when there are two classifiers that output the same class value, the confidence values of both classifiers will be averaged and compared to the confidence probability value of the other class. The function selects the class having the greatest confidence (average) probability value.

#### V. RESULTS AND DISCUSSIONS

As mentioned, this experiment employs the best three models as suggested by the experiment previously conducted (see Figure 2) combined in parallel and assigned similar inputs. Those are naive-Bayes, Multi-layer Perceptron and RBF-Network. Figure 3 depicts the reachability performance of various schemes.

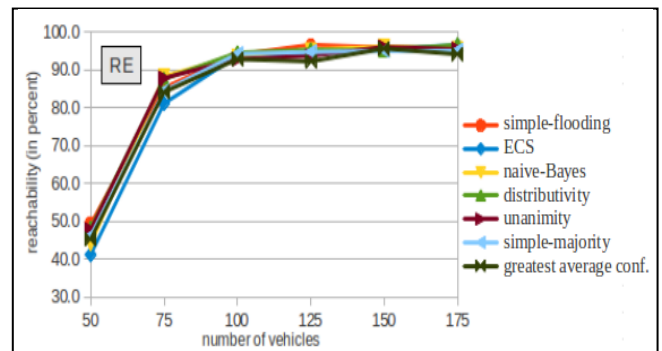


Fig. 3. The reachability of the schemes over various densities

It reveals that our schemes with any combiners (i.e. unanimity, distributivity, simple-majority and greatest-average-confidence) are comparable in reachability performance. Reachability differences are not clearly shown when the reachability performance presented in this fashion. Therefore, the reachability performance differences between schemes will be assessed based on their average reachability over the simulated densities. Figure 4 shows the average reachability comparisons of the four classifiers combination functions used in our schemes over simple-flooding, ECS scheme and naive-Bayes classifier at various network densities (i.e. from 50 – 175 vehicles).

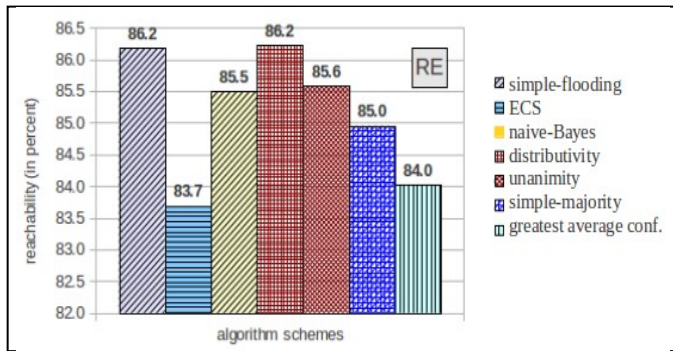


Fig. 4. The average reachability comparisons among the schemes

It is shown that overall, all the multiple-classifiers schemes with any of the combination functions experimented outperform the ECS scheme for their average reachability. A notable reachability performance has been shown by our scheme using 'distributivity' combination function that reaches the same value of the simple-flooding's average reachability. Compared to the ECS scheme, the 'distributivity' combiner contributes to an additional increase of 2.5% in average reachability. This also exceeds the reachability value given by the single classifier (naive-Bayes scheme).

Figure 5 reveals that at denser networks, the ECS's broadcast saving (SRB measure) decreases. This means the exponential function used in the ECS scheme puts more broadcast probability along with the increase of the network density. In contrast, all other combiners used in our schemes show a gradual increase along with the increase of the density. The 'unanimity' combiner outperforms all other schemes in terms of saved-rebroadcast. It contributes to an additional increase of 38.9% in rebroadcast saving compared to that of ECS measured at 175 vehicle density.

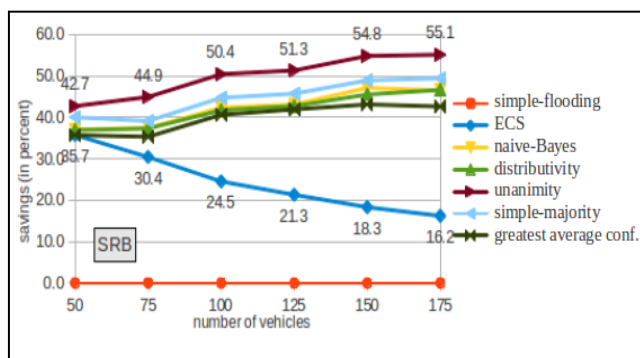


Fig. 5. The saved-rebroadcast comparisons among the schemes

Our multiple-classifiers schemes, so far, are superior to the ECS scheme in terms of reachability and saved-rebroadcast performance. Usually the saved-rebroadcast is inversely related to the reachability. This means, for achieving a higher reachability, a lower saved-rebroadcast value must be paid. The ECS scheme cannot maintain the broadcast saving when the network density increases. Simply stated, the ECS scheme

could not correctly adapt the density changes or the employed exponential function should be reconsidered. Conversely, our schemes have shown a growing broadcast saving along with the increase of the network density.

## VI. CONCLUSIONS

As Weka's classifiers only outputs limited information for every unlabeled instance inputted (i.e. the predicted class and the confidence value of the instance being predicted to be in such a class) the combination of several classifiers can only be done in the abstract level. The classifier-based schemes are adaptive and consistently provide a remarkable rebroadcast saving over the experimented densities.

A notable reachability performance has been shown by the combined-classifiers using the distributivity function that reaches the same value of flooding's average reachability. Compared to the ECS scheme, the multiple-classifier scheme with 'distributivity' function contributes to a 2.5% increase in reachability (in average). The unanimity combiner outperforms all other schemes in terms of saved-rebroadcast. It contributes to an increase of 38.9% savings on rebroadcast compared to that of ECS measured at 175 vehicle density.

The combined-classifiers are proven to have firmer broadcast decisions than any single classifier. The unanimity combiner, for example, offers significant improvements over all performance parameters than that of the naive-Bayes classifier. To improve the performance of the combined classifiers, more potential classifiers/models could also be employed. It is expected that having many models might make the decision making more accurate. However, employing many classifiers must be accompanied by a powerful processor to reduce processing delay.

## REFERENCES

- [1] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," in: Proceedings of ACM/IEEE Mobicom '99, 1999, pp. 151-162.
- [2] Y.-C. Tseng, S.-Y. Ni, E.-Y. Shih, "Adaptive approaches to relieving broadcast storms in a wireless multihop mobile ad hoc network," in: Proceedings of the 21st International Conference on Distributed Computing Systems, 2001, pp. 481-488.
- [3] S.-Y. Ni, Y.-C. Tseng, Y.-S. Chen, J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," *Wireless Networks* (2002) 153-167.
- [4] Y.-C. Tseng, S.-Y. Ni, E.-Y. Shih, "Adaptive approaches to relieving broadcast storm in a wireless multihop mobile ad hoc network," *IEEE Transactions on Computers* (2003) 545-557.
- [5] N. Wisitpongphan, O. K. Tonguz, J. S. Parikh, P. Mudalige, F. Bai, and V. Sadekar, "Broadcast storm mitigation techniques in vehicular ad hoc networks," *IEEE Wireless Communications*, pages -, Dec 2007.
- [6] V. Gau and J.-N. Hwang, "Adaptive probabilistic broadcasting over dense wireless ad hoc networks," *International Journal of Digital Multimedia Broadcasting*, 2010:-, 2010.
- [7] M. Slavik, I. Mahgoub, and M.M. Alwakeel, "Analysis and evaluation of distance-to-mean broadcast method for vanet," *Journal of King Saud University Computer and Information Sciences*, 26:153-160, 2014.

- [8] C. Hu, Y. Hong, and J. Hou. "On mitigating the broadcast storm problem with directional antennas," In IEEE International Conference on Communications (ICC'03), pages 104–110, 1999.
- [9] M. Torrent-Moreno, J. Mittag, P. Santi, and H. Hartenstein, "Vehicle-to-vehicle communication: Fair Transmit Power Control for Safety-Critical Information," In IEEE Transactions on Vehicular Technology, pages 3684 – 3703, Sept. 2009
- [10] Y. Mylonas, M. Lestas, and A. Pitsillides, "Speed adaptive probabilistic flooding in cooperative emergency warning," In Proceedings of the 4th Annual International Conference on Wireless Internet (WICON'08), pages –, Nov 2008.
- [11] C. Wu, K. Kumekawa, and T. Kato, "A novel multi-hop broadcast protocol for vehicular safety applications," Journal of Information Processing, 18:110–124, 2010.
- [12] A. Qayyum, L. Viennot, and A. Alouiti, "Multipoint relaying for flooding broadcast messages in mobile wireless networks," In Proc. 35th Annual Hawaii Intl. Conf. System Sciences, Big Island, Hawaii, pages 3866–3875, 2002.
- [13] I-H. Bae, "Design and evaluation of a hybrid intelligent broadcast algorithm for alert message dissemination in vanets," International Journal of Grid and Distributed Computing, 4(4):1–9, Dec 2011.
- [14] J. Yang and Z. Fei, "Broadcasting with prediction and selective forwarding in vehicular networks," International Journal of Distributed Sensor Networks, 2013.
- [15] M. Bakhouya, J. Gaber, and P. Lorenz, "An adaptive approach for information dissemination," in Journal of Network and Computer Applications, volume 34(6):1971–1978, 2011.
- [16] Mohammed, A., Ould-Khaoua, M., Mackenzie, L., "An improved rebroadcast probability function for an efficient counter-based broadcast scheme in MANETs," In: 25th Annual UK Performance Engineering Workshop (UKPEW'09), Jul 2009
- [17] VanetMobiSim. <vanetmobisim.sourceforge.net/>.
- [18] The Network Simulator ns-2. <www.isi.edu/nsnam/ns>.
- [19] DSRC.<www.etsi.org/technologies-clusters/technologies/intelligent-transport/dsrc>
- [20] Weka 3: Data Mining Software in Java. <www.cs.waikato.ac.nz/ml/weka/>