

**An Evaluation Methodology for Reliable Simulation Based Studies of Routing
Protocols in VANETs**

Authors: J.M. García-Campos, J. Sánchez-García, D. G. Reina, S. L. Toral, F. Barrero
University of Seville (Spain)

Escuela Superior de Ingenieros. Avda. Camino de los Descubrimientos, s/n
41092 Sevilla (Spain)

email: josgarcam@etsi.us.es, jsanchez73@us.es, dgutierrezreina@us.es, storal@us.es,
fbarrero@us.es

phone: +34 954 48 12 93; fax: +34 954 48 73 73

Corresponding author: S. L. Toral

An Evaluation Methodology for Reliable Simulation Based Studies of Routing Protocols in VANETs

Abstract: Vehicular Ad hoc networks (VANETs) have attracted much attention in the last decade. Many routing protocols have been proposed for VANETs and their performance is usually evaluated and compared using simulation-based studies. However, conducting reliable simulation studies is not a trivial task since many simulation parameters must be configured correctly. The selected parameters configuration can considerably affect the simulation results. This paper presents a methodology for conducting reliable simulations of routing protocols in VANETs urban scenarios. The proposed methodology includes relevant simulation aspects such as measurement period, selection of source-destination pairs for the communication traffic flows, number of simulations, mobility models based on road city maps, performance metrics and different analyses to evaluate routing protocols under different conditions. The proposed methodology is validated by comparing the simulation results obtained for Ad Hoc On-Demand Distance Vector (AODV) routing protocol with and without using the proposed methodology. The obtained results confirm that by using the proposed methodology, we can achieve more reliable simulations of VANETs routing protocols.

Keywords: Simulation studies, Performance evaluation, Routing protocols, VANETs.

1. Introduction

An ad hoc network is an autonomous system of wireless nodes that cooperatively form a network without any specific administration [1]. Each node in an ad hoc network is in charge of routing information among its neighbours. When nodes are free to move randomly and organize themselves arbitrarily, we refer to them as Mobile Ad Hoc Networks (MANETs) [2]. As an evolution of traditional MANETs, VANETs (Vehicular Ad hoc NETWORKs) [3] include communications between vehicles (Vehicle to Vehicle, also known as V2V communications) on the roads and with the road communication infrastructure (Vehicle to Infrastructure, also known as V2I communications). In these dynamic networks in which the topology is continuously changing, it is crucial to analyze how routing protocols react against reach possible changes in the network topology. Many routing protocols have been proposed and evaluated for both MANETs and VANETs [4]. The evaluation of MANETs and VANETs routing protocols by simulation is the most common approach for testing the protocols performance so far. The reason is that real experimentation in multi-hop ad hoc networks is costly in terms of hardware requirements [5][6]. Testing routing protocols with real MANETs and VANETs prototypes require a high number of wireless devices, and consequently, there are only a few available testbeds in the world [7]. Therefore, conducting reliable simulation studies is an important requirement to validate the performance of routing protocols for VANETs. MANET simulations have suffered from credibility problems for the last decades [8], mainly because of bad simulation practices conducted by the research community. This lack of credibility is also observable in VANET simulations since most MANET routing protocols researchers are also working on VANET routing protocols design [9]. The goal of this

paper is to propose an evaluation methodology to improve the reliability of VANET routing protocols simulation results.

The main contributions of this paper are as follows:

- A methodology for simulation-based studies of routing protocols for VANETs based on good simulation practices. It includes many simulation aspects such as Warm Up periods, selection of communication pairs, number of simulations, mobility models, performance metrics and simulation analyses. The selection of these simulation parameters is very important to make a fair and unbiased comparison of routing protocols.
- The validation of the proposed methodology by the comparison of the performance of several well-known reactive routing protocols such as AODV [10], LAR [11], and DYMO [12] with and without using the proposed methodology.

This paper is organized as follows: section 2 presents relevant related works on the evaluation methodology of routing protocols for MANET and VANET scenarios. Section 3 describes the proposed methodology. Section 4 includes the validation of the proposed methodology using AODV, LAR and DYMO routing protocols. Finally, section 5 contains the main conclusions of this work.

2. Related work

One of the pioneering works highlighting the importance of conducting reliable simulations in wireless multi-hop ad hoc networks can be found in [13]. The authors report the bad practices followed by researchers in a high number of simulation studies for MANETs. The main findings of this study are, i) there is a high dissimilarity among

the simulation scenarios in terms of density, size of the scenarios and mobility, and all of them impact significantly on the obtained simulation results; ii) the mobility of nodes do not model real scenarios since most studies use the random waypoint mobility model, which has been reported to be harmful for simulation studies [14], and iii) the execution of the simulations and data collection are not properly carried out for a statistical analysis. As a conclusion, the authors in [13] provide some guidelines to run reliable simulations. As a solution for the dissimilarity in the simulation scenarios, the same authors provide in [15] a mechanism to generate standard scenarios for MANETs based on the number of nodes and the size of the scenarios. However, this mechanism is only valid for random waypoint mobility model. In [16], the authors also review the main issues in MANET simulations studies. Additionally, the authors put in evidence the discrepancies in the simulation results obtained from different network simulators. Moreover, they indicate an issue related to the traffic pattern generation and its impact in the obtained simulation results. This issue is specifically addressed by the evaluation methodology proposed in this paper. On this line, in [8] the authors evaluate the performance of several well-known routing protocols with different traffic patterns. They highlight the impact of the repetition of the source nodes in the communication pairs. In this paper, we also address this issue, but we take a step forward and we evaluate the impact of the repetition of the destination nodes in the selected communication source-destination pairs. Actually, we demonstrate that the repetition of destination nodes has a higher impact on the simulation results than the repetition of the source nodes. In a more recent work presented in [17], the authors state that still a high number of simulation studies about routing protocols for wireless multi-hop ad hoc networks do not follow good simulation practices. Among the bad practices described in

[17], the authors underline the random selection of source-destination pairs in the traffic patterns as an issue for the evaluation of routing protocols. This paper addresses many of the issues indicated in the previous works on the evaluation of routing protocols and proposes an evaluation methodology to solve many of them. It includes simulation aspects such as Warm Up period, selection of source-destination pairs, mobility models, number of simulations, and selection of simulation analyses. All the aforementioned simulation parameters must be selected carefully in order to obtain reliable simulation results.

3. The proposed methodology for reliable simulations in VANETs

This section describes the proposed methodology that evaluates how the simulation parameters must be selected for obtaining reliable simulation results. Figure 1 shows the main points of the proposed methodology such as the communication pair selection, the measurement period, the selection of number of simulations, the selection of a mobility model, the performance metrics and the analyses. The main objectives of the proposed methodology are, 1) firstly, to highlight relevant simulation parameters that affect the simulation results, and 2) to provide guidelines on the selection of these parameters to obtain a low disperse set of samples in order to obtain reliable statistics. In the next subsection, we describe in more detail each point of the proposed methodology.

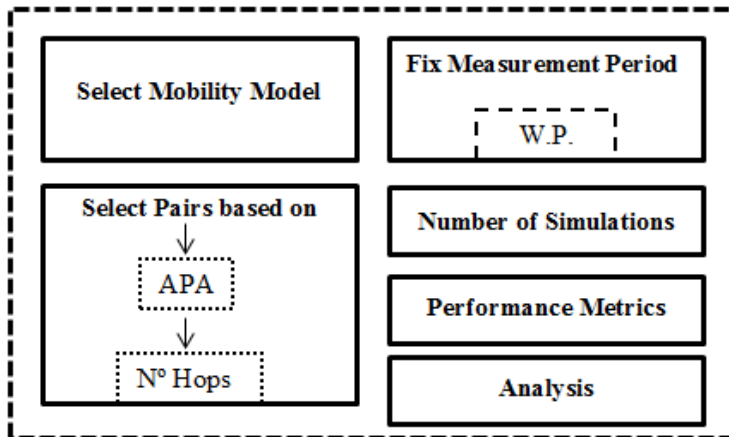


Figure 1 Main aspects of the proposed methodology

1) Measurement period

In this subsection we introduce an important concept to carry out good simulations for VANETs. The *Warm Up period* (W.P. in Figure 2) is the time frame which ensures the stability of three relevant simulation aspects such as:

- All communication pairs have started transmitting application packets.
- The mobility model has achieved a stable state.
- The buffers of the nodes have stabilized.

Notice that the establishment of communications among a source-destination pair can start at different times. Normally, the starting times are selected randomly, so some pairs could have more time to transmit data packets than others. This fact can influence the simulation results if the selected pairs do not have the same properties in terms of average number of hops and path availability between the source and the destination nodes. By using a Warm Up period, we avoid discrepancies among the measurement period of the performance metrics during the simulation time. To obtain reliable and

non-dispersed simulation results, performance metrics must be measured from W.P. value to the end of the simulation period, which is named after *Measurement Simulation Period* (M.S.P. in Figure 1). In order to select the Warm Up period, we have to consider two aspects, the first one is the period during which the communication flows are established, and the second aspect is the mobility model, since we need to guarantee that the mobility of nodes is stable.

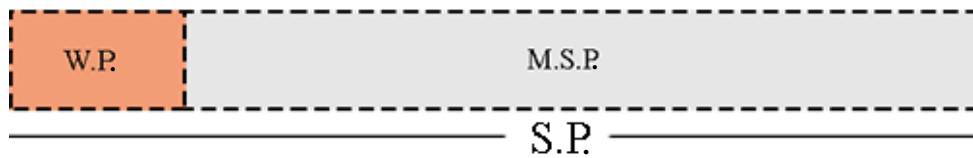


Figure 2 Status time bar

In order to show the importance of the Warm Up period values on the simulation results, we depict in Figure 3 the throughput performance metric for different Warm up values. We consider an urban scenario (4000x4000 m²) corresponding to a fragment of the city of Washington D.C with 125 nodes moving with the IDM (Intelligent Driver Model) [18] mobility model. From now on, we will refer to this scenario as the scenario under test that will be used to test several features of the proposed methodology. We use the throughput performance metric as standard metric. It is defined as the number of application packets delivered within the simulation time. In the simulations, all source nodes start generating application packets from the period between 0 and 50 seconds until the end of the simulation. The results shown in Figure 3 have been obtained using 25 different pairs selected randomly in the scenario under test. We include the 95% confidence intervals in Figure 3. From now on, similar confidence intervals are shown in this paper for the rest of figures.

We can observe in Figure 3 that when the Warm Up value increases, the throughput also increases and this is because the number of source nodes that have started to transmit application packets is higher. When the Warm Up value is higher than 50 (See 75 and 100 in Figure 3) the throughput values are very similar (See Table 1 for more details). Regarding the confidence intervals for each Warm Up value, they are high because the source and destination pairs are selected randomly. We highlight the importance of the communication pair selection in the next subsection. In consequence, the Warm Up value selection only affects the mean while the dispersion does not depend on this parameter (the dispersion is the descriptive statistic used by researchers in order to compare the performance of routing protocols). From now on, we consider 50s as the Warm Up value recommended for the simulation of the scenario under test.

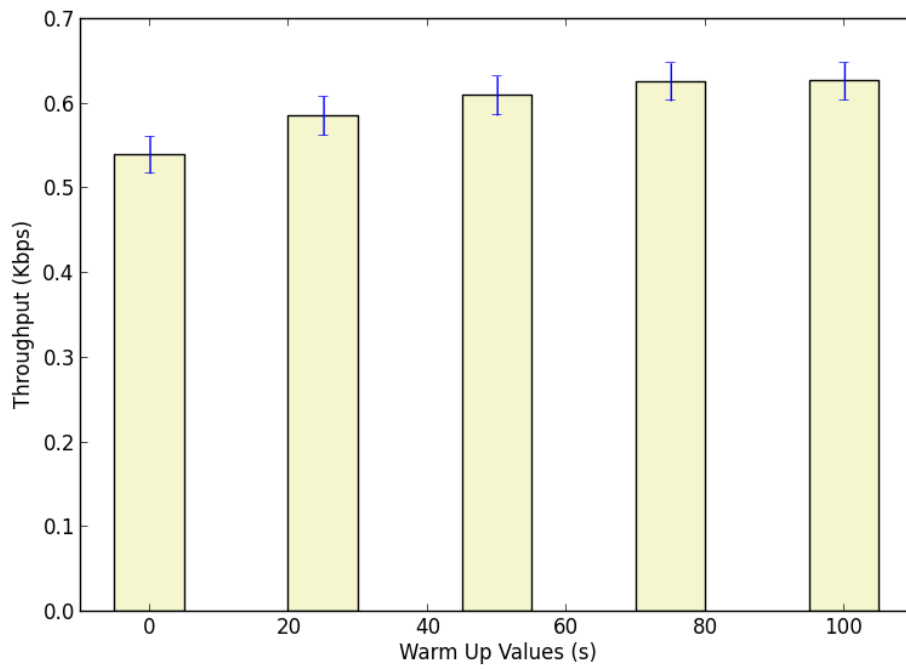


Figure 3 Throughput vs Warm Up values

Table 1 Throughput vs Warm up values in the scenario under test

Warm up Values (s)	0	25	50	75	100
Throughput (Kbps)					
Mean	0.5395	0.5859	0.6093	0.6256	0.6265
Confidence interval	0.0220	0.0229	0.0231	0.0223	0.0222

2) *Source-destination selection*

In this subsection, we present a communication pair selection based on four features that strongly affect the simulation results such as the path availability, the separation in terms of the number of hops between the source and destination nodes, and the repetition of source and destination nodes.

We define the communication pair selection as the mechanism by which the source and destination nodes of a communication flow are selected. The source node is responsible for generating the data packets, and the destination node is the target node in the network for those generated packets. Consequently, intermediate nodes will route the generated data packets towards the destination node using routing information. In simulation analyses, the communication pairs are normally selected at the beginning of the simulations. In most simulation-based studies of routing protocols for VANETs [19], the source and destination nodes are selected randomly among all nodes of the network. Although the original aim of this practice is achieving a fair selection of pairs, this can impact negatively on the dispersion of the obtained simulation results for several reasons. First, by using a random selection we cannot guarantee that all source-destination pairs have similar properties in terms of number of hops and path availability. Consequently, the simulation results may vary drastically from one pair to another. It is expected that routing protocols will obtain worse results when the number

of hops increases and the path availability is lower. This situation is even more aggravated if the source-destination pair selected cannot be established. This means that it is not possible to establish a communication path from the source node to the destination node during the simulation time. Furthermore, the performance of the routing protocols can also be biased if the number of hops is very low. Second, outliers are prone to appear when random selection is applied because of the great variability of the results. This affects to the mean of the simulation results. To solve this problem, we propose to use the Average Path Availability (APA) [20] and the number of hops between the source and destination nodes as the key metrics to select source-destination pairs. The APA metric is defined as the fraction of time during which a path is available between two nodes. We select source-destination pairs which have similar APA values because pairs with different APA values produce very dissimilar results. As an example, Figure 4 shows the distribution of APA values in the scenario under test. We can observe in Figure 4 that high values of APA are more probable than low values in the scenario under test but also that there are some APA values which are zero. This situation corresponds to source-destination pairs that cannot be established.

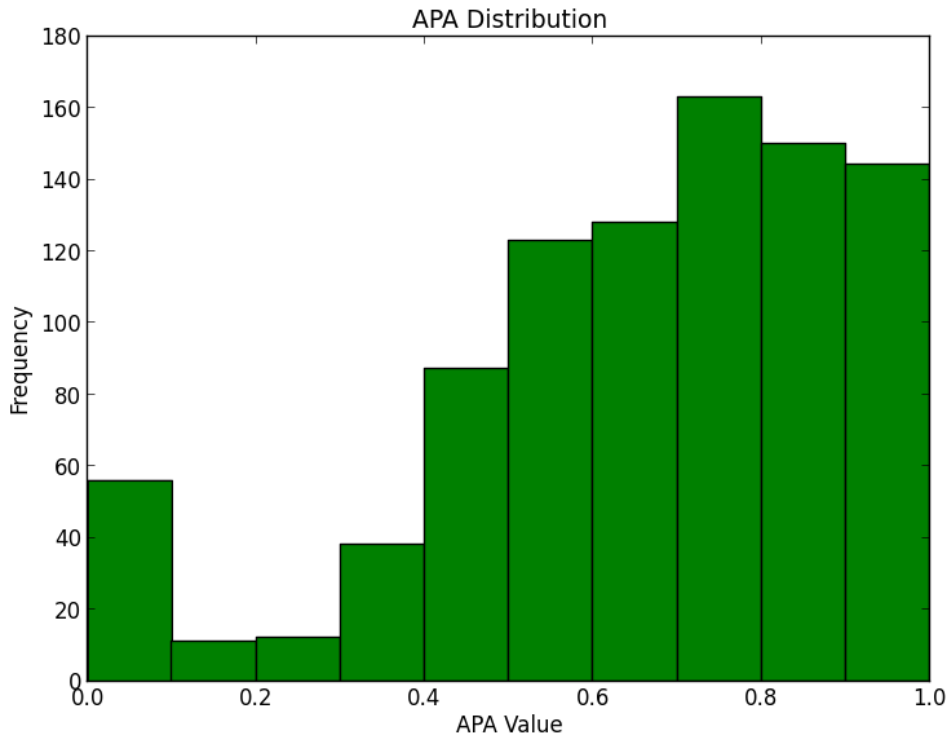


Figure 4 APA Distribution for the scenario under test

To highlight the impact of the APA value on the simulation results, we depict in Figure 5 the throughput for different APA values in the scenario under test. We can observe in Figure 5 that when the APA value increases the throughput also increases (See Table 2 for more details). The reason for this is that the available communication paths between the source and destination nodes are higher. We also depict in Figure 5 the confidence intervals of the measurements (blue vertical line in Figure 5) for each obtained throughput value. The confidence intervals are considerably lower for higher values of APA. Notice that if we select source and destination pairs randomly we could pick pairs with different APA values, and in consequence, the simulation results will be more disperse. Let us illustrate this situation with an example, if we select randomly two communication pairs with very different APA values such as 0.2 and 1.0, we will obtain

0.20 Kbps and 0.86 Kbps respectively for the throughput (See Table 2 for more details). The mean of both throughput values will be 0.53 Kbps, and the confidence interval 0.65, which is quite high related to the obtained mean. Consequently, the obtained mean does not reflect the performance of AODV in the scenario under test. The reason is that two different APA values represent two different network conditions from the source node viewpoint. Conversely, if we select the communication pairs with similar APA values, we will guarantee that the network conditions in terms of connectivity will not change from one communication pair to another. As a result, we will achieve less disperse results.

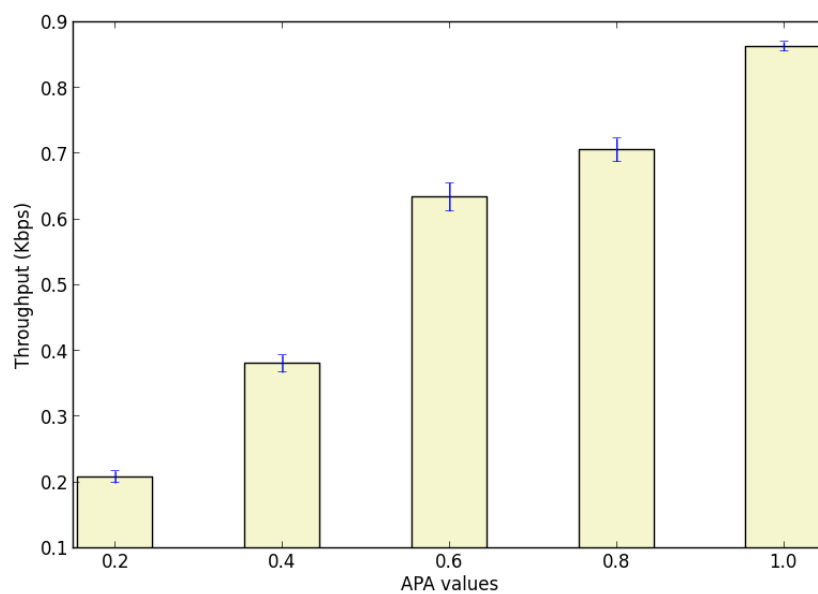


Figure 5 Throughput vs APA values

Table 2 Throughput vs APA values in the scenario under test

APA value	0.2	0.4	0.6	0.8	1
Throughput (Kbps)					
Mean	0.2076	0.3799	0.6337	0.7051	0.8625
Confidence interval	0.0091	0.0132	0.0216	0.0178	0.0070

Once introduced the importance of the APA value in the simulation results, the next step in the proposed communication pair selection is to present a mechanism to choose a number of pairs that have similar values of APA. It is not a trivial task because in general the APA value depends on the topological characteristics of the simulation scenario such as the density and mobility of nodes. In particular, the density of nodes plays an important role. As a rule, the higher the density of nodes, the higher the APA value on average in the network. Consequently, the selection will be different according to the density of nodes in the network. As a primary condition, we should guarantee that there are a significant number of source-destination pairs in the network which can be selected as a valid pairs. For this reason we have to fix an APA value for each density of nodes which assures this condition, which we name after *target APA* from now on. We represent the inverse of the cumulative distribution function (1-CDF) for the APA values in the aforementioned simulation scenario in Figure , but in this case, varying the number of nodes in the network in order to obtain different values of density. Notice that the 1-CDF describes the probability that the APA value (considered as a random variable) is higher or equal than a given value. It is obvious that the most restrictive case is for the lowest density level (50 nodes) in which the APA values are in general very low. We state that the 1-CDF value of the target APA for a given scenario should be at

least 20 % in order to have enough number of source-destination pairs to be selected.

This condition is marked in Figure 6 with black points.

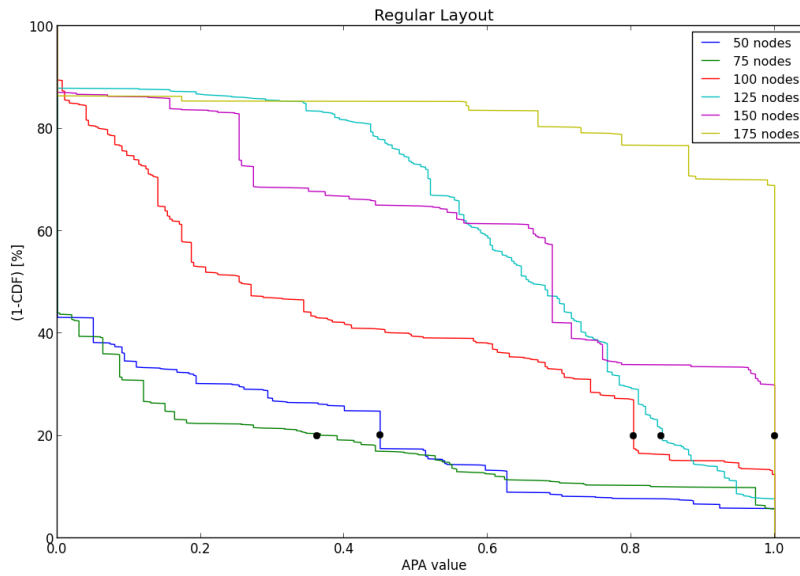


Figure 6 Inverse cumulative distribution function of the APA for different density levels

Table 3 Target APA values

Scenario	Target APA
50 nodes	0.370
75 nodes	0.452
100 nodes	0.811
125 nodes	0.846
150 nodes	1.0
175 nodes	1.0

Following the APA condition described above, the Table 3 contains the minimum target APA values that accomplish such condition for each density level.

The next step in the communication pair selection is to select those that are separated by the same number of hops on average. Notice that the APA value provides an idea about the availability of a communication path between the source node and the destination

node, but it does not take into account the separation in terms of the number of hops in the scenario under test (See Figure 7).

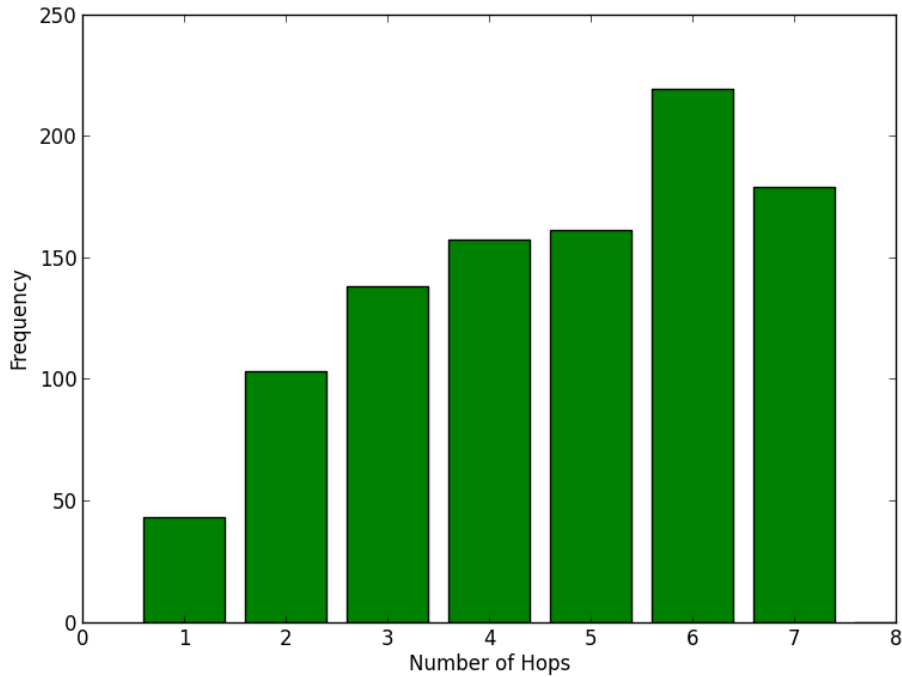


Figure 7 Number of hops distribution for the scenario under test

The minimum number of hops to reach a destination node is 1, this situation corresponds to the situation in which the destination node is within the source node coverage, and the maximum number of hops is 7. To ensure a reliable and fair evaluation of the routing protocols, we should fix a similar number of hops for each selected communication pair since the number of hops impacts significantly in the obtained simulation results as shown in Figure 8. Figure 8 shows that the throughput of the network is reduced (See Table 4 for more details) as the number of hops in the selected communication pairs increases. In order to illustrate the importance of the separation in number of hops, we use another well-known performance metric such as the NRL (Normalized Routing Load) (red line in Figure 8). This metric is defined as the

ratio between routing packets and the total number of delivered application packets. In Figure 8, the NRL increases (See Table 4 for more details) as the number of hops increases because the routing protocol needs to generate a larger number of control packets to discover/maintain the routes. This metric is directly related to the energy consumption [21]. The confidence intervals for both metrics also increase (blue and red vertical lines in Figure 8 respectively) as the number of hops between source and destination nodes increase. If we select source and destination pairs randomly, we could pick pairs separated by very different hop numbers. Consequently, the simulation results will be more disperse. Let us consider an example of this situation in which we focus on the NRL metric. If we select randomly two pairs separated by very different numbers of hops such as 2 and 7, the NRL values obtained for these values are 1.24 and 9.96 respectively (See Table 4 for more details). The mean of both values will be 5.6 and the confidence interval 8.54. Notice that the obtained confidence interval is too high compared with the mean. Consequently, the obtained mean does not reflect the performance of AODV in the scenario under test.

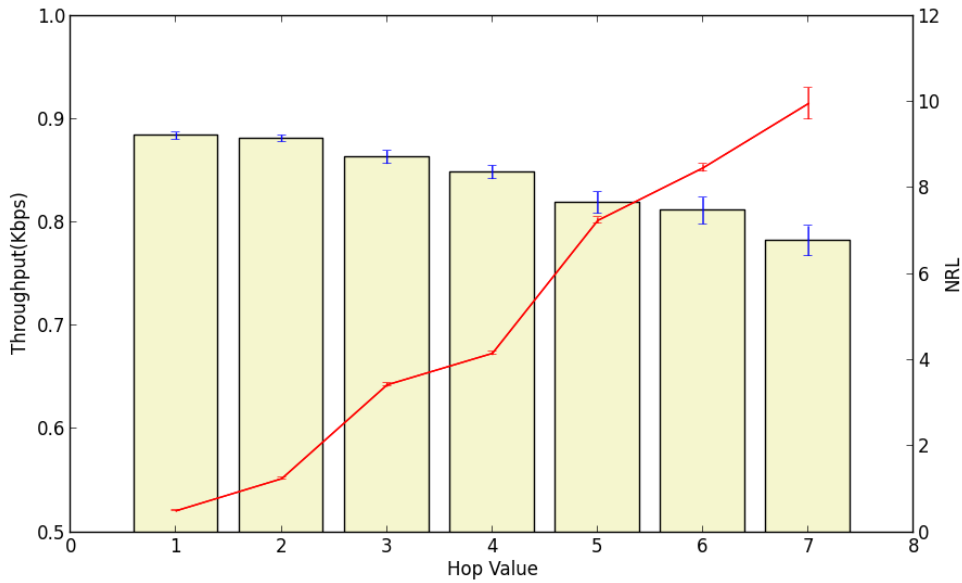


Figure 8 Throughput and NRL vs number of hops in the scenario under test

Table 4 Statistics measures for Throughput vs number of hops

Number of hops	1	2	3	4	5	6	7
Throughput (Kbps)							
Mean	0.8835	0.8812	0.8629	0.8484	0.8189	0.8111	0.7820
Confidence interval	0.0033	0.0031	0.0063	0.0066	0.0104	0.0128	0.0143
NRL							
Mean	0.5006	1.2448	3.4239	4.1562	7.2428	8.4652	9.9593
Confidence interval	0.0169	0.0251	0.0449	0.0390	0.0704	0.0858	0.3664

Following the same procedure for the APA selection, we need a selection mechanism to choose communication pairs with similar separation in terms of the number of hops. Again, we use the inverse cumulative distribution (1-CDF) for different density levels in order to get more insight into the distribution of the number of hops. The obtained results are shown in Figure 9.

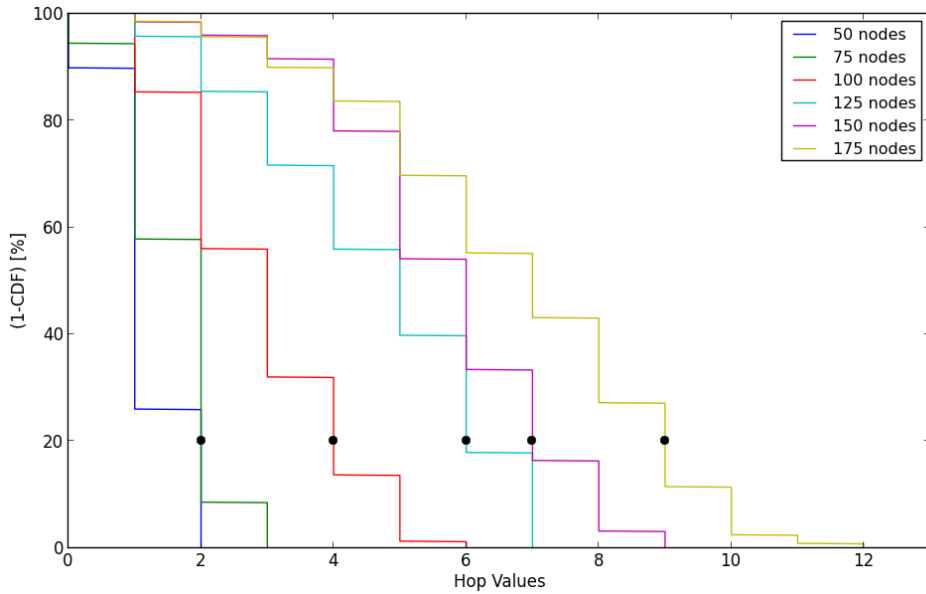


Figure 9 Inverse cumulative distribution function of the number of hops for different density levels

In this case, we should select the highest possible number of hops since we are interested in evaluating the performance of routing protocols, but we must also ensure that there are enough source destination pairs. For this reason, we should focus on the most restrictive case, which corresponds to the density level of 50 nodes, as it is very difficult to select communication pairs that are separated by more than 2 hops (See Figure 9). As with the APA value, we state that the 1-CDF value of the target number of hops for a given scenario should be at least 20 % in order to have enough number of source-destination pairs. This condition is marked in Figure 9 with black points. Table 5 shows the target number of hops values that accomplish a minimum number of pairs with that condition in the scenario under test. From here on, when we refer to the proposed methodology we must ensure that the two above conditions (APA and number of hops) are fulfilled.

Table 5 Target hops values

Scenario	Target Hops
50 nodes	2
75 nodes	2
100 nodes	4
125 nodes	6
150 nodes	7
175 nodes	9

Another important feature that should be controlled in the selection of communication pairs is the possibility of several communication pairs having the same source and destination nodes. Although the actual source and destination nodes will depend on the underlying application in the VANET, for a reliable evaluation, we should guarantee that the selected pairs have similar properties in terms of repetitions of source and/or destination nodes. A great variability in the repetition of selected nodes can impact significantly on the simulation results. For instance, if destination nodes are very frequently repeated then the queue of these nodes can saturate, and this fact can cause dropped packets.

To highlight the importance of the repetition of the source and destination nodes, we depict in Figure 10 and Figure 11 the throughput for different numbers of repeated source and destination nodes in the scenario under test.

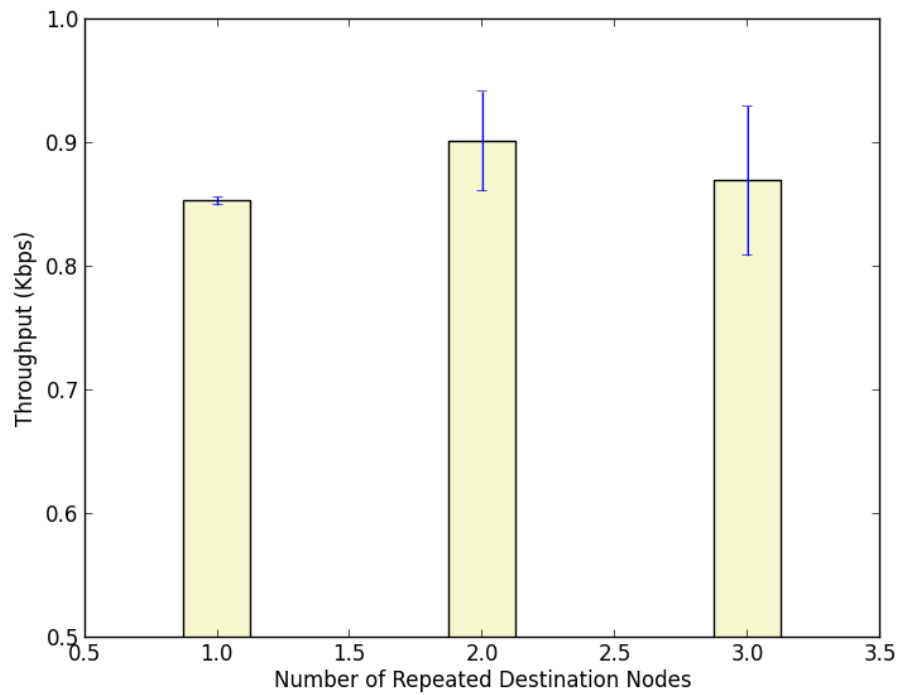


Figure 10 Throughput vs repeated destination nodes

Table 6 Throughput vs repeated destination nodes values in the scenario under test

Number of Repeated Destination Nodes	1	2	3
Throughput (Kbps)			
Mean	0.8529	0.9012	0.8690
Confidence interval	0.0034	0.0399	0.0504

As we can see in Figure 10, when the number of repeated destination nodes increases the confidence intervals are higher (See Table 6 for more details) because the number of application packets lost in the destination nodes buffers are higher.

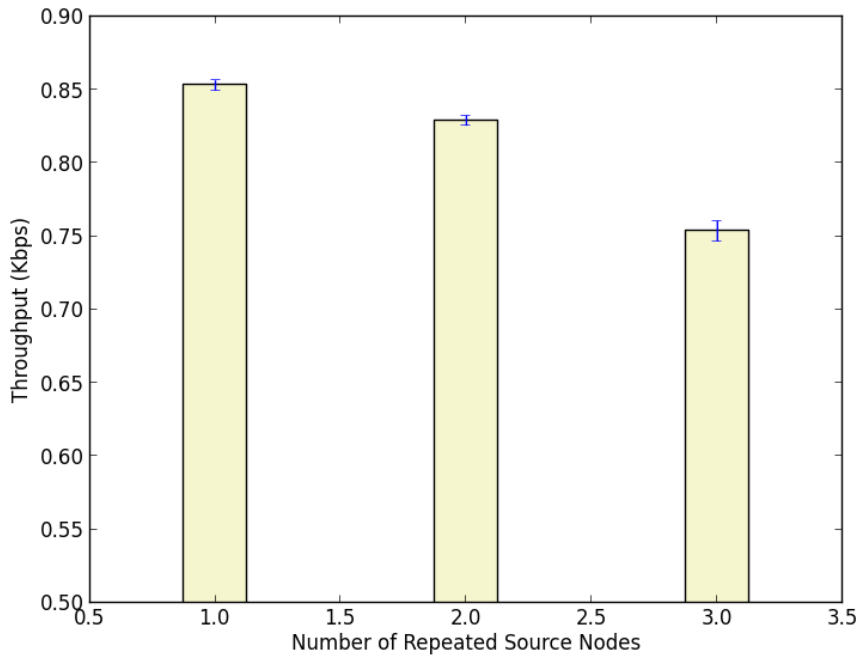


Figure 11 Throughput vs repeated source nodes

Table 7 Throughput vs repeated source nodes values in the scenario under test

Number of Repeated Source Pairs	1	2	3
Throughput (Kbps)			
Mean	0.8529	0.8289	0.7535
Confidence interval	0.0034	0.0035	0.0070

We can observe in Figure 11 that as the number of repeated source nodes increases the results are more disperse (See Table 7 for more details) because there are some application packets lost in the intermediate nodes buffers due to traffic congestion. However, the results are lesser scattered than in the case of repeated destination nodes (See figure Figure 10). This is due to the fact that the destination nodes buffers are more congested than the intermediate nodes, causing more collisions and contention in the shared wireless medium.

When random selection of communication pairs is used, we cannot control whether the destination and source nodes are repeated or not. Such repetition does not mean that the simulated scenario is unrealistic. However, this situation can favor the performance of a routing protocol with respect to others. For example in the case that one routing protocol is specifically designed to take into account this network condition. By applying the proposed selection we can guarantee that all routing protocols are evaluated under the same network conditions.

Finally, to illustrate the importance of the APA, the number of hops, the Warm Up period, and the repetition of source/destination nodes, we compare the simulation results in the scenario under test with and without considering the aforementioned simulation practices. Table 8 details the nomenclature that will be used to describe the following simulation results.

Table 8 Chosen nomenclature

Acronym	Meaning
R.P.	Randomly selected pairs
R.P with M.P.	Randomly selected pairs with measurement period
S.P.	Selected pairs based on APA and number of hops
S.P with M.P.	S.P with measurement period
S.R.	Randomly selected pairs with repeated source
D.R.	Randomly selected pairs with repeated destination
S.R. with M.P. + S.P.	Repeated source nodes selected based on target APA and number of hops with measurement period
D.R. with M.P. +S.P.	Repeated destination nodes selected based on APA and number of hops with measurement period

Figure 12 shows the throughput obtained in the scenario under test with 25 different source-destination pairs. We depict these results with boxplot graphs and we also highlight with a green point the obtained mean of the samples. The boxplot graphs are used to better understand how values are spaced out in different sets of data. The bottom line of the box represents the first quartile (Q1), the top line represents the third quartile

(Q3) and the distance between them is the interquartile range ($IQR=Q3-Q1$). Another important aspect in the boxplot graphs are the whiskers which depend on the IQR. The upper whisker is determined by the equation $Q3 + 1.5*(IQR)$ while the bottom one is determined by $Q1 - 1.5*(IQR)$. Therefore if we have a set of data bits scattered, the IQR and the whiskers will be low. In this paper we will focus on the IQR to measure the level of dispersion. Figure 12 includes results for *R.P.*, *R.P. with M.P.*, *S.P.* and *S.P. with M.P.* cases (See Table 8 for more details).

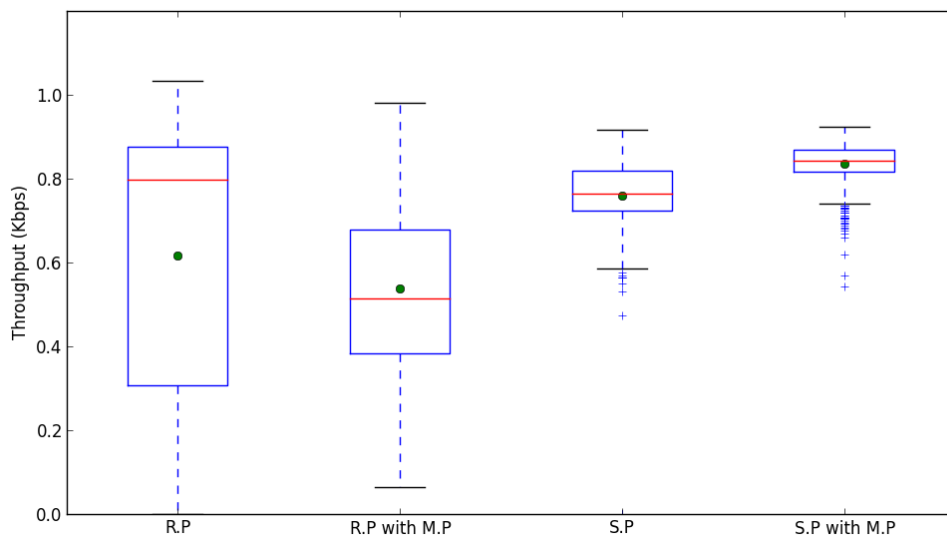


Figure 12 Throughput vs source destination pair selection

Using the Warm Up period (See *R.P with M.P* in Figure 12) we can observe in Figure 12 that the distance between first and third percentile is smaller than without these periods (*R.P* in Figure 12). This means that we have less scattered measures. To highlight the importance of Warm Up period we focus on the selected pairs based on the APA and number of hops case (*S.P* in Figure 12). We can observe that in this case the IQR is lower than the case when the selection technique is not used (See *R.P and R.P with M.P* in Figure 12) because the set of values is lesser scattered. If we also use the

measurement period (*S.P with M.P* in Figure 12) the IQR is lesser. Consequently, we are able to obtain reliable simulation results by using the proposed methodology.

Figure 13 shows the simulation results for *S.R*, *D.R*, *S.R with M.P + S.P* and *D.R with M.P + S.P* cases (See Table 8 for more details). Again, the simulation scenario is the scenario under test with 25 communication pairs, and the maximum number of repetitions for both the source and destination nodes is 3. It means that nodes can be selected as a source or destination in 3 communication pairs. According to the results shown in Figure 13, the repetition of source nodes affects more significantly the simulation results than the repetition of the destination nodes. Additionally, including the proposed selection based on APA and the number of hops, the simulation results are even better in terms of dispersion.

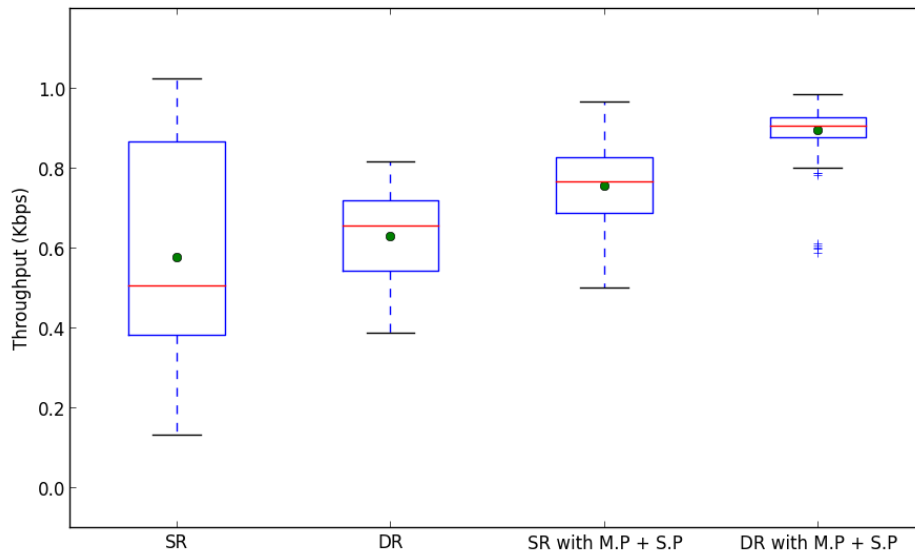


Figure 13 Throughput vs source destination pair selection

3) *Number of simulations*

Another important aspect to be considered when conducting simulation-based studies is the number of simulations that should be carried out for each data point in the results. Clearly, the more simulation trials, the more representative data sample we obtain. However, the simulation results also incur in computing time consumption. Consequently, a trade-off between the number of simulations and the computing time should be reached. We should not devote more time than the necessary to conduct simulations. Thus, the number of simulations should be selected in order to obtain a representative data sample without requiring excessive simulation time. Figure 14 shows the throughput results and the required computing time for different number of traffic seeds (number of simulations). The simulation scenario is the same one described in the previous section (the scenario under test) with 25 source-destination pairs of communications. As expected, we can observe that the computing time is higher as the number of seeds increases as well (See Table 9 for more details). When we use the proposed pair selection and the measurement period, the computing time is lower than when we do not use them. The main reason is that the network is less congested because the number of routing packets is lower, due to the APA based selection. Since we aim to obtain reliable simulation results, we want a good confidence interval. This means obtaining non-dispersed results with the lowest computing time. Figure 14 shows two different cases, the first one corresponds to the proposed communication pair selection, and the second one is for the case of using random pairs.

The obtained results are not similar in terms of the obtained average mean. By using the proposed methodology we obtain better results than not using it (See Table 9 for more details). In addition, the proposed communication pairs selection needs a lower number

of simulations to obtain better and lesser disperse results, as it is shown by the lower confidence intervals. To highlight the importance of our methodology in terms of computing time consumption, we focus on the 5 seed case in which we achieve a very good confidence interval without requiring excessive simulation time. We define a new metric to evaluate this situation, the ratio between the confidence interval and the mean. From now on, we will name it after *normalized confidence interval*. If we use the proposed measurement period and the communication pair selection mechanism, the *normalized confidence interval* takes the value of 0.01 for 5 seeds. The number of simulations to reach this value without the measurement period and the proposed communication pair selection mechanism is 80 simulations. It means that we need 16 times more simulations to obtain the same dispersion. In terms of computing time, we save 279.3 s ($287.9 \text{ s} - 8.6 \text{ s}$) when using the proposed methodology.

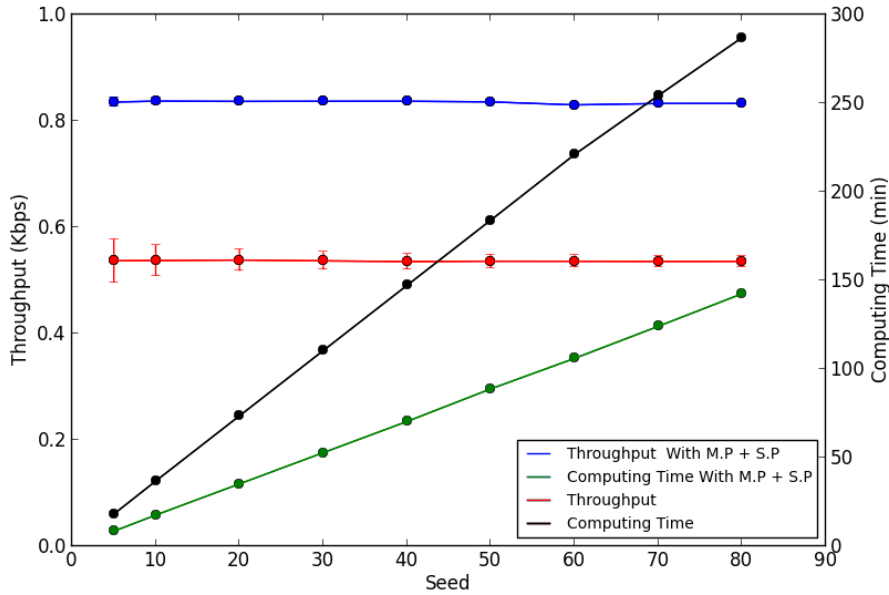


Figure 14 Throughput and computing time vs number of simulations

Table 9 Statistics measures for Throughput vs number of simulations

Number of simulations (seeds)	5	10	20	30	40	50	60	70	80
Throughput (Kbps)									
With M.P + S.P									
Mean	0.8353	0.8372	0.8367	0.8369	0.8370	0.8358	0.8301	0.8327	0.8328
Confidence interval	0.0089	0.0061	0.0043	0.0041	0.0034	0.0030	0.0027	0.0026	0.0025
Computing Time (min)	8.6	17.6	35.2	52.8	70.4	88.8	106	124.1	142.3
Without M.P + S.P									
Mean	0.5369	0.5374	0.5381	0.5368	0.5354	0.5362	0.5360	0.5365	0.5370
Confidence interval	0.0411	0.0291	0.0203	0.0165	0.0143	0.0127	0.0115	0.0109	0.0089
Computing Time (min)	16.5	36.8	73.6	110.4	147.2	184.0	221.0	254.3	287.9

4) Mobility in VANETs

A critical issue in VANETs simulation studies is the need of a mobility model that reflects the real behavior of vehicular traffic in urban scenarios. There are some mobility generators that are able to create mobility patterns that emulate such scenarios, some of them are VanetMobSim [22], SUMO [23], FreeSim [24] and CityMob [25]. In this study we use CityMob for Roadmaps (C4R) [26] as the selected mobility generator, which allows us to simulate vehicular traffic in different locations using real maps. C4R uses two tools to generate the mobility model. On the one hand, it uses OpenStreetMap [27] to get the real roadmaps and SUMO [23] to generate the vehicles and their movements within the scenario.

The functionality provided by C4R is twofold: it defines the vehicle movements on the streets, and it limits their mobility according to the vehicular congestion and traffic rules [28]. To simulate the vehicle movements in a VANET scenario, C4R provides the following mobility models: Krauss [29], Krauss modified [30], Wagner [31], Kerner [32], Downtown model [18] and Intelligent driver model (IDM) [18]. Moreover, C4R

allows users to modify some parameters to customize the mobility model [28], such as the attraction rate, downtown rate, departure, simulation time or number of traces.

The next objective of the proposed methodology is to determine how to select a representative VANET scenario to evaluate routing protocols. According to the classification made in [33], the cities can be categorized according to the density of their streets and junctions as simple, regular and complex layouts. Three cities that fall in such classification are Los Angeles, Washington and Tokyo respectively. We have studied the APA and the number of hops values found in these three layouts.

Figure 15 shows the distribution of the APA and the number of hops for each layout. We can observe that in general the number of hops is higher for more complex layouts. Regarding the APA distribution, we observe similar distributions for regular and complex layouts, where APA values higher than 0.5 are more probable.

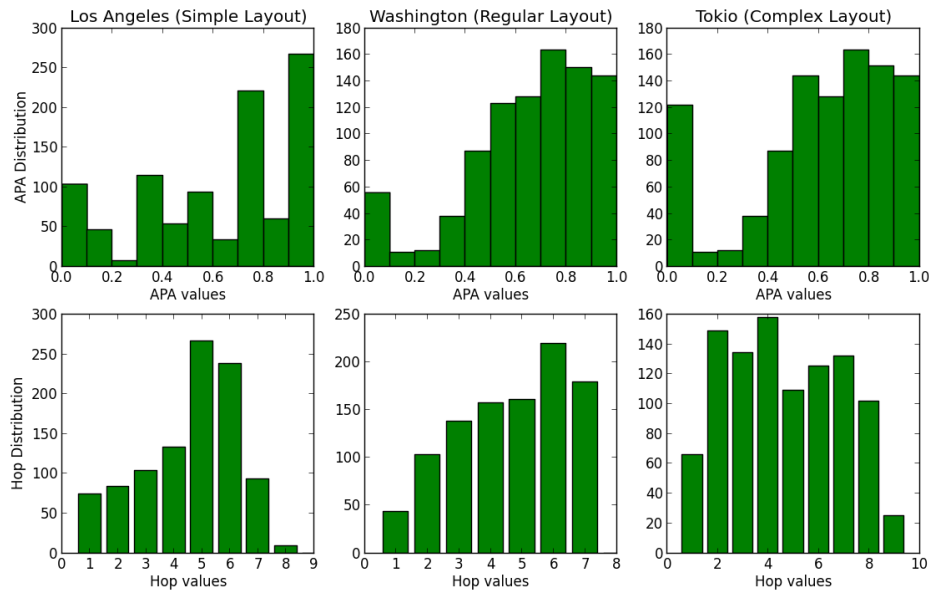


Figure 15 APA and number of hops distribution for different layout

Moreover, and regardless of the layout, we can select APA values within $[0, 1]$ interval according to Figure 15. Similarly, for the three layouts we can select pairs separated by

the same or similar number of hops. Consequently, controlling the APA and separation in number of hops between the source and destination nodes, we can expect similar simulation results for the three layouts as long as we apply the proposed methodology. This fact is shown in Figure 16, which represents the simulation results obtained by the proposed methodology (With P.M. in Figure 16) (See Table 10 for more details) and the simulation results without using the proposed methodology (Without P.M. in Figure 16). When the proposed methodology is applied, the results are less disperse and very similar to each other.

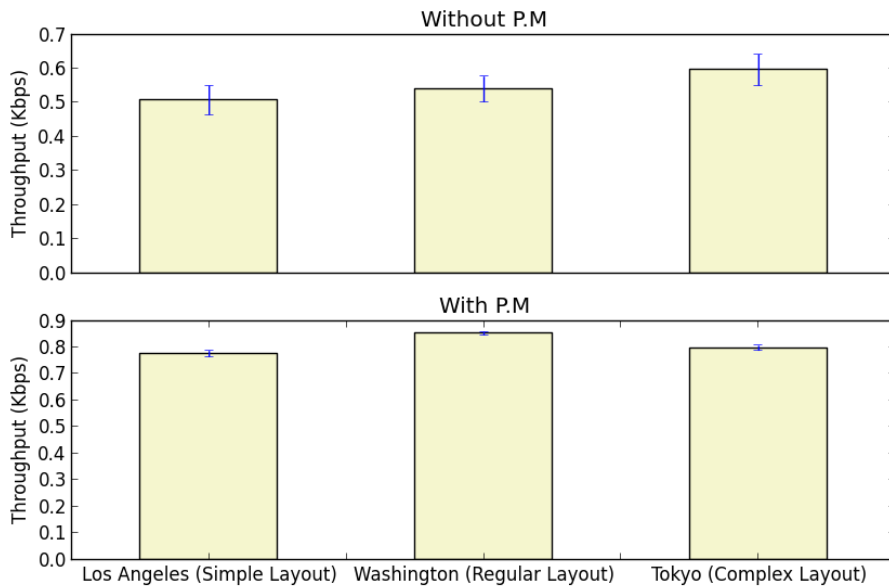


Figure 16 Throughput vs roadmap profile classification

Table 10 Statistics measures for Throughput vs layout classification

Layout Profile	Los Angeles(Simple)	Washington (Regular)	Tokyo (Complex)
<i>Without P.M.</i>			
Mean	0.5073	0.5394	0.5958
C.I.	0.0428	0.0374	0.0450
<i>With P.M.</i>			
Mean	0.7754	0.8527	0.7967
C.I.	0.0113	0.0063	0.0093

Next step is to study the inverse of the cumulative distribution function for the APA values. In Figure 17, we can observe that we obtain similar results for the three layouts. It means that for the same APA, the connectivity of the network is very similar and does not depend on the layout. This might be explained by the fact that the simulations do not consider obstacles like buildings or other vehicles.

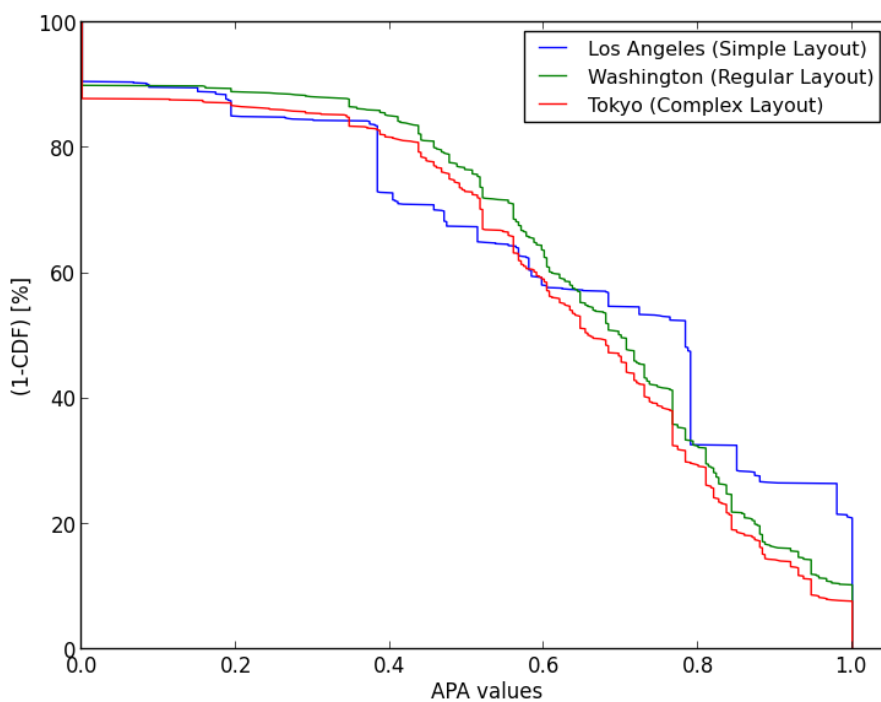


Figure 17 1-CDF for APA distribution for the three considered layouts

The last step in this analysis of the mobility of nodes is to study the cumulative distribution function (1-CDF) for the separation of nodes in terms of the number of hops. As we can observe in Figure 18, the results for the three layouts are very similar. However, for the Complex layout case the maximum number of hops is higher than for the other layouts. From now on in this study, we will use a city with a regular layout

such as Washington because it has an intermediate behavior in terms of the APA and number of hops distribution.

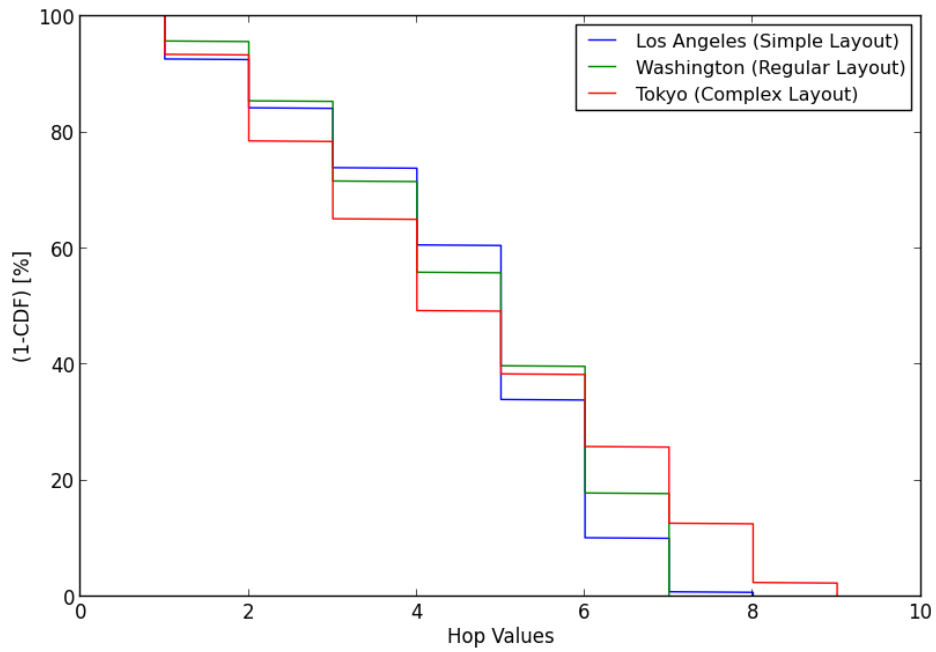


Figure 18 1-CDF for the number of hops separation for the three considered layouts

Figure 19 depicts the area of Washington that will be used in the simulation results section. The left-hand part of Figure 19 shows a real capture of the area used in our simulations, the right-hand part shows the model obtained from C4R and the bottom part depicts the movements of the vehicles over the selected area, with each color representing a different node.

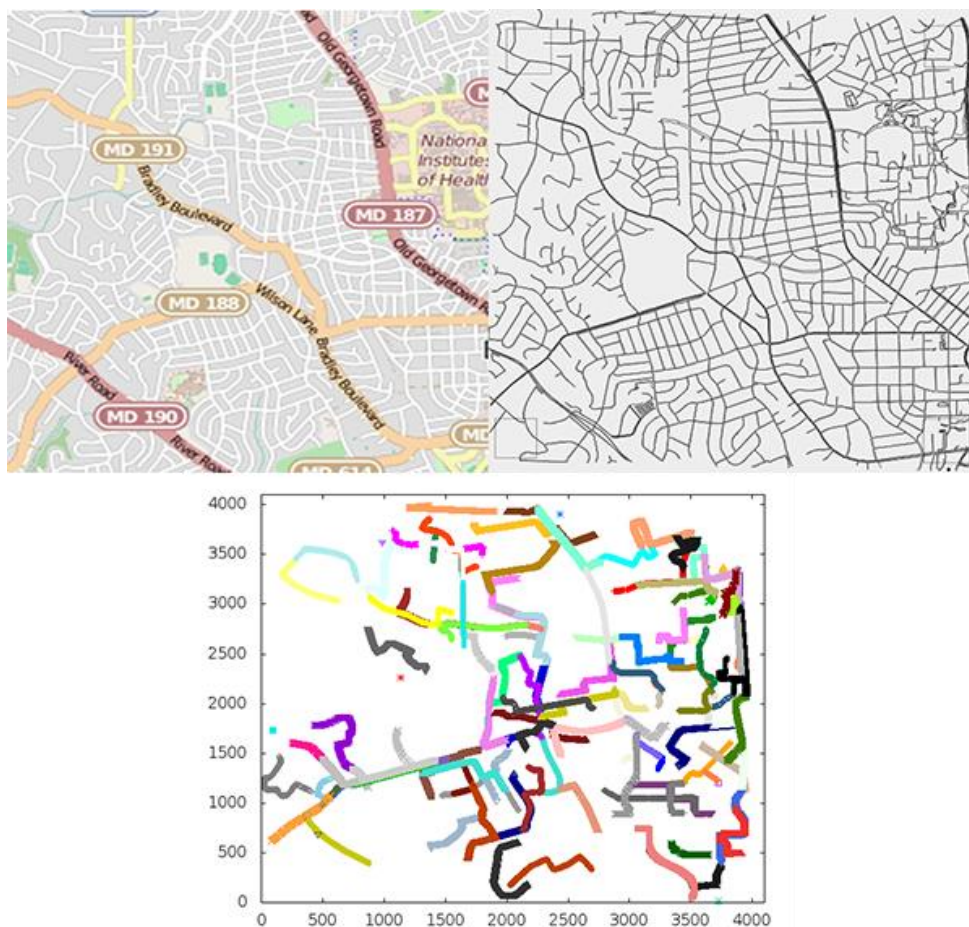


Figure 19 Washington layout and node movements in the scenario under test with 125 nodes

To model the behavior of the drivers we use the IDM model [18]. The chosen values for each parameter are those that ensure a maximum speed of 30Km/h, which is the speed limit in urban environments. We also ensure a minimum security distance between two vehicles trying to model the real driving conditions in urban scenarios. In this case we

are simulating a normal driving behaviour as described in [18] with a minimum security distance of 2 meters.

5) Performance metrics

Another important aspect to be considered when evaluating routing protocols is which performance metrics should be used in order to represent an unbiased performance of the routing protocols. It is important to use metrics that exhibit the performance of the routing protocols in different conditions. The following performance metrics are some of the most used in simulation-based studies [34].

Throughput (THR): It is the sum of the data packets in the simulation period.

$$THR (Kbps) = \frac{\sum \text{Delivered Application Packets}}{\text{Simulation time}} \quad (1)$$

Average End-to-End Delay (E2E): It is defined as the time taken for a data packet to be transmitted across an ad hoc network from the source to the destination node.

$$E2E (s) = \frac{\sum (\text{Delivered Time} - \text{Transmitted Time})}{\text{Number of packets successfully delivered}} \quad (2)$$

Normalized Routing Load (NRL): It is the ratio of the total routing packets to the total delivered data packets.

$$NRL = \frac{\sum \text{Routing Packets}}{\sum \text{Delivered Application Packets}} \quad (3)$$

Packet Delivery Fraction (PDF): It is the ratio of the number of packets delivered to the receiver, to the number of packets sent by the source.

$$PDF (Kbps) = \frac{\sum \text{Delivered Application Packets}}{\sum \text{Sent Packets}} \quad (4)$$

Jitter (JIT): It is the delay between two consecutive packet deliveries at a node.

$$JIT = \frac{\sum \text{delays}_{\text{packet delivered}}}{\text{Packet}_{\text{Application delivered}}} \quad (5)$$

Additionally, in this paper we propose a new performance metric, the Route Activity Time (RAT), which is aimed at evaluating the capability of a routing protocol to maintain an active route between the source and the destination nodes. The formal definition of RAT is as follows:

Route Activity Time (RAT): It is the period of time during which a communication path is available between the source and the destination nodes. In routing protocols based on request, reply and error messages, such as AODV, DSR, and DYMO, it is the period elapsed between the time at which the reply message arrives at the source node and the time at which an error message of such route is generated.

$$RAT(s) = \text{Error}_{\text{time}} - \text{Reply}_{\text{time}} \quad (6)$$

Notice that the RAT metric measures how the routing protocols manage the path availability. In theory, we control the APA values by selecting the communication pairs, however, the real time in which a communication path is established between a source and a destination node will depend on the underlying routing protocol and the network conditions.

Table 11 summarizes the desirable values for each metric used to evaluate the routing protocols performance.

Table 11 Desirable values for the performance metrics

Metric	THR	E2E	NRL	PDF	JIT	RAT
Desirable values	<i>High</i>	<i>Low</i>	<i>Low</i>	<i>High</i>	<i>Low</i>	<i>High</i>

Although we have described six different performance metrics (1)-(6) to evaluate the performance of the routing protocols, we will only use four of them, THR(1), E2E(2),

NRL(3) and RAT(4) in the next section, since the rest of them provide equivalent information. By using the THR metric, we measure the performance of the routing protocols in terms of the number of delivered packets. With E2E we evaluate the average delay of the application packets. Using the proposed RAT metric, we measure how the routing protocols maintain the communication routes between the source and destination nodes. The NRL metric measures the number of routing packets used by the routing protocols and provides an idea about the power consumption of the routing protocol. Regarding PDF and JIT metrics, we do not use these metrics for the following reasons. First, with the PDF metric we also measure the number of delivered packets so this metric will show us similar results of THR metric. Second, we do not use the JIT metric because it gives us an idea about the network delay and we are actually using E2E to measure this performance.

6) *Simulation analyses*

The objective at this point is to decide which analyses we should carry out for obtaining a good performance evaluation of routing protocols. In general, the number of nodes is a common parameter to vary in simulation-based studies in order to evaluate routing protocols under different density levels (connectivity). However, there are other parameters that also affect considerably the performance of routing protocols. For instance, the congestion is a common issue in multi-hop networks because nodes should share the wireless medium, and consequently, routing and application packets should compete for the wireless medium. The congestion in the network can be modified by varying some parameters of the communication flows between the source and destination nodes such as data rate, size of packets, and number of flows. Among the mentioned parameters, we focus on the number of flows since in this paper we are

giving relevance to the selection of source-destination pairs. Therefore, we propose two different analyses. First, a density analysis based on varying the number of nodes while maintaining the same number of communication flows. Second, a congestion analysis, focused on varying the number of communication flows while maintaining the same number of nodes. With the first analysis, we evaluate the routing protocols under different connectivity levels and with low congestion conditions. In the second analysis, we set a medium-high value of density (high connectivity) and vary the congestion of the network to observe how routing protocols perform under different levels of congestion.

To summarize the procedure described in this section, Table 12 provides the most important values of the proposed methodology and also the benefits of each of them.

Table 12 Summary of the simulation parameters used in the proposed methodology

Simulation parameter	Selection	Benefit obtained
S.P. = 300 s	W. P. = 50 s (M.S.P. =250 s)	Using W.P. we improve mean of the used performance metrics
Selection Pairs based On APA	Depending on the scenario and based on APA target	Applying the proposed methodology based on APA and Number of hops we reduce the dispersion.
Selection Pairs based On Hop Number	Depending on the scenario and based on number of hops target	
Performance Metrics	THR(1), E2E(2), NRL(3), RAT(6)	We can evaluate different features of the evaluated routing protocol.
Scenario	Washington (Regular layout)	We can emulate real scenarios.
Analysis	Congestion and density	We evaluate the routing protocols under different network conditions.

4. Evaluation of the proposed methodology

This section includes the results of the proposed methodology when evaluating routing protocols in VANETs. The aim is to show that the proposed methodology leads to more reliable simulation results. For this purpose, we compare the obtained results of several well-known and widely used routing protocols for multi-hop ad hoc networks such as AODV, LAR and DYMO with and without the proposed methodology.

1) *Simulation environment settings*

We use NS-2.34 [35] under a Debian Linux operating system. NS-2 is a simulation tool for replicating real life networking environments. To simulate urban mobility of vehicles, we use CityMob for Roadmaps (C4R) [25]. Table 13 summarizes the general simulation settings used. It is important to highlight some specific aspects of VANET simulations such as the MAC protocol used, which is the IEEE 802.11p [36]. This is based on the 802.11a standard and has the same structure. The main difference, compared to 802.11a, is its bandwidth, which is narrower in order to make the signal more resistant to fading and multipath propagation in the automotive environment. Another important difference is the operating frequency; in the case of 802.11p standard is 5.9 GHz, as opposed to the 802.11a, which is a standard operating at 5 GHz in Europe [37]. Regarding the propagation model, we use the two-ray ground reflection model [28] because it gives more accurate prediction for long distances than the free space model. We select a transmission range of 500 m, which is a typical transmission range for VANET scenarios (the standard IEEE 802.11p can reach up to 1000 m). The simulation time is 300 s, it is high enough to guarantee a good evaluation of routing protocols. The warm up period has been selected according the study conducted in the

previous section 3.1. The type of traffic is Constant Bit Rate (CBR), which is typically used in multi-hop scenarios with UDP transport layer. CBR traffic is suitable for real time applications. The transmission rate of application packets is 1 packet per second and the size of each packet is 512 bytes. The simulation scenario is the one described in section 3.4. The maximum speed of nodes is 30 km/h. This maximum value is suitable for urban scenarios, where the limited speed is about 50 km/h.

Table 13 Simulation parameters

Parameter	Value
Simulation Time	300 s
Warm Up period	50 s
Routing Protocols	AODV,DYMO, LAR
Transmission range	500 m
MAC Protocol Type	IEEE802.11p
Number of Nodes	50, 75, 100, 125, 150, 175
Numbers of Sources	5, 10, 15, 20, 25
Traffic Types	CBR
Transport Layer	UDP
Maximum Packet in Queue	50
Packet Size	512 bytes
Packet Rate	1 packet/s
Area Size	4000*4000 m ²
Mobility model	Washington Layout
Propagation model	Two-ray ground
Maximum speed of nodes	30 km/h

2) Simulation results

In this subsection we show the obtained results of the two different proposed analyses, the density analysis and the congestion analysis.

Density analysis

In this analysis we vary the number of nodes for each of the three used routing protocols, in steps of 25 each, within the interval [50, 175]. The number of communication flows is fixed to 5 (low congested scenario), the APA values are selected according to Table 3 and the maximum number of hops according to Table 5.

Figure 20, Figure 21 and Figure 22 show the simulation results for THR, RAT, E2E, NRL performance metrics, with their confidence intervals for each number of nodes. Figure 20, Figure 21 and Figure 22 include the results using the proposed methodology (*Results with P.M* in Figure 20, Figure 21 and Figure 22) and without using the proposed methodology (*Results without P.M* in Figure 20, Figure 21 and Figure 22). We depict the results for 5 and 30 simulations (seeds in Figure 20, Figure 21 and Figure 22) with and without applying the proposed methodology to check the differences between them in terms of the mean and the confidence interval. By using the proposed methodology, the results are in general, very similar in terms of mean. Nevertheless, the confidence intervals are low for both numbers of simulations. But if we randomly select the source destination pairs, there are important differences in the mean values of some metrics and also in the confidence intervals which are large in spite of running many simulations.

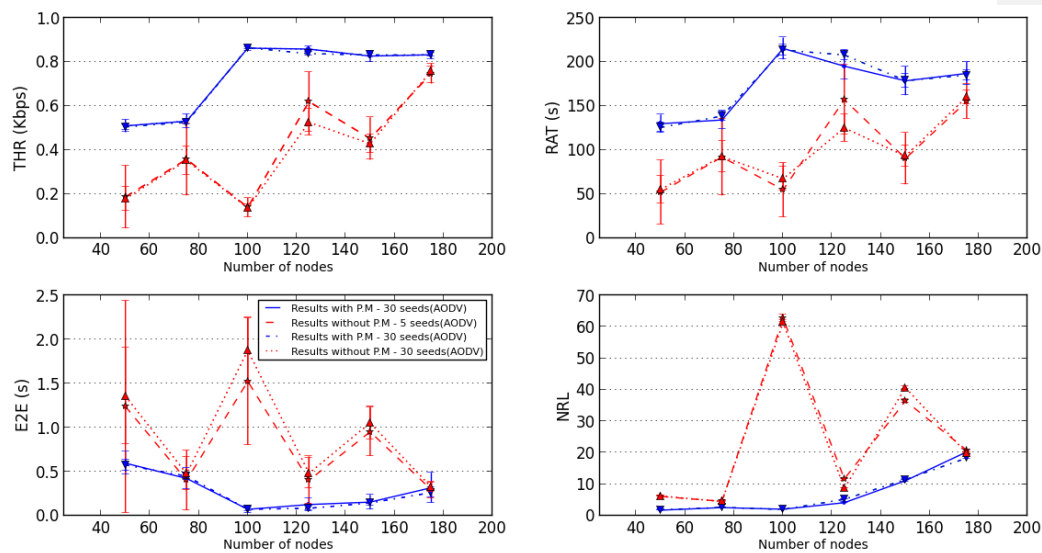


Figure 20 Results of the density analysis – AODV

Figure 20 clearly shows that the results obtained by AODV when using the proposed methodology are better and more reliable in terms of dispersion and tendency. Table 14 details the exact values of the mean and the confidence intervals of each performance metric for each number of nodes. We can observe that as the number of simulations increases the results are lesser scattered in both cases (With and Without P.M.). However, the differences are small when applying the proposed methodology; it means that by using the proposed methodology we obtain good results, in terms of dispersion, with a low number of simulations. On the one hand, in the P.M. case there are not important differences among the metrics, the results are similar. On the other hand, without P.M. case there are differences in some of them, for instance in the RAT, E2E and NRL metrics.

Table 14 Statistics measures for performance metrics vs number of nodes - AODV

Number of Nodes		50	75	100	125	150	175
AODV							
THR (Kbps)							
<i>With P.M.</i>							
5 seeds	Mean	0.5098	0.5310	0.8625	0.8582	0.8266	0.8325
	C.I.	0.0288	0.0328	0.0121	0.0119	0.0253	0.0184
30 seeds	Mean	0.5050	0.5262	0.8643	0.8377	0.8315	0.8317
	C.I.	0.0126	0.0112	0.0041	0.0077	0.0078	0.0080
<i>Without P.M.</i>							
5 seeds	Mean	0.1875	0.3583	0.1399	0.5235	0.4524	0.7475
	C.I.	0.1428	0.1639	0.04315	0.03791	0.09642	0.04556
30 seeds	Mean	0.1790	0.3527	0.1370	0.5202	0.4287	0.7595
	C.I.	0.0551	0.0653	0.0167	0.0156	0.0415	0.0194
RAT (s)							
<i>With P.M.</i>							
5 seeds	Mean	129.922	134.097	215.288	195.189	178.542	186.949
	C.I.	10.236	10.610	12.326	15.519	16.296	12.953
30 seeds	Mean	125.676	138.484	213.515	207.896	178.794	184.845
	C.I.	4.970	4.947	5.981	6.0117	7.616	6.138
<i>Without P.M.</i>							
5 seeds	Mean	52.221	91.979	54.925	134.405	90.367	155.215
	C.I.	36.460	43.448	30.859	12.670	29.243	19.459
30 seeds	Mean	55.469	92.662	67.431	126.133	93.120	160.211
	C.I.	15.591	17.707	13.299	5.745	12.170	7.799
E2E (s)							
<i>With P.M.</i>							
5 seeds	Mean	0.5963	0.4231	0.0743	0.1269	0.1545	0.3174
	C.I.	0.1314	0.11706	0.0408	0.0758	0.0832	0.1701
30 seeds	Mean	0.5762	0.4482	0.0737	0.08347	0.1453	0.2627
	C.I.	0.0646	0.0411	0.0193	0.0254	0.0276	0.0566
<i>Without P.M.</i>							
5 seeds	Mean	1.2394	0.4026	1.5247	0.6782	0.9525	0.2945
	C.I.	36.460	43.448	30.859	12.670	29.243	0.0944
30 seeds	Mean	1.3568	0.4762	1.8766	0.820	1.0538	0.3237
	C.I.	0.5468	0.1870	0.3704	0.0837	0.1841	0.0539
NRL							
<i>With P.M.</i>							
5 seeds	Mean	1.7628	2.5814	2.0633	4.1438	11.1339	20.2829
	C.I.	0.0532	0.0391	0.0669	0.1047	0.2496	0.7114
30 seeds	Mean	1.8860	2.6462	1.9964	5.2214	11.5078	18.3319
	C.I.	0.0242	0.0131	0.0189	0.0133	0.1205	0.2064
<i>Without P.M.</i>							
5 seeds	Mean	6.1176	4.5756	62.7165	26.2550	36.4604	20.8193
	C.I.	0.0615	0.0280	1.2346	0.3189	0.5694	0.3104
30 seeds	Mean	6.2507	4.5251	61.4516	25.6496	40.6906	19.8507
	C.I.	0.0223	0.0255	0.6222	0.1706	0.3172	0.1527

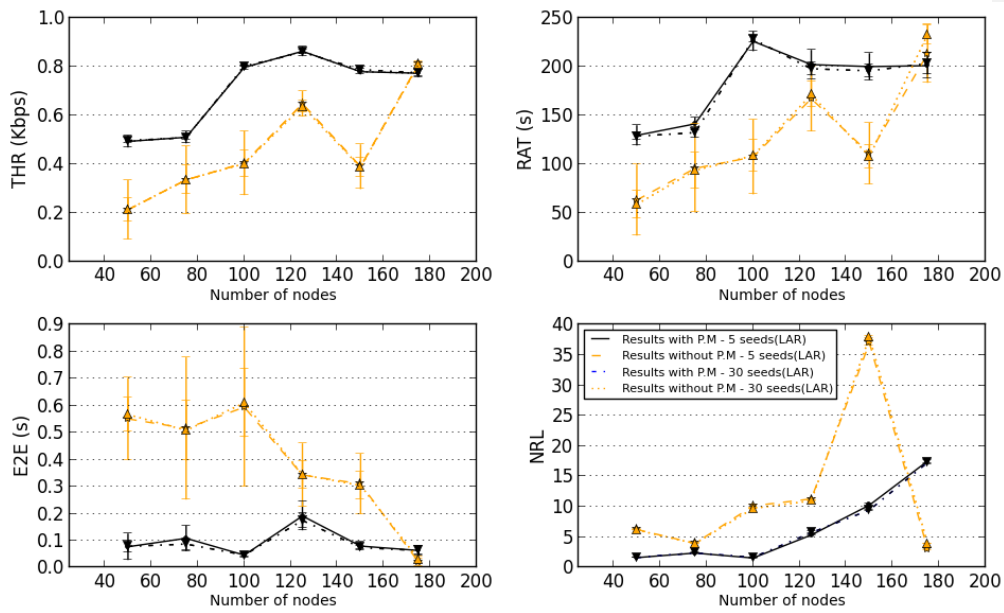


Figure 21 Results of the density analysis – LAR

Figure 21 shows the simulation results obtained by the LAR routing protocol. Again, the results are better in terms of mean and confidence interval when the proposed methodology is used. Table 15 details the obtained results for the considered performance metrics. Clearly, the dispersion of the results is much better when applying the proposed methodology.

Table 15 Statistics measures for performance metrics vs number of nodes – LAR

Number of Nodes		50	75	100	125	150	175
LAR							
THR (Kbps)							
<i>With P.M.</i>							
5 seeds	Mean	0.4937	0.5106	0.7982	0.8645	0.7805	0.7742
	C.I.	0.0248	0.0248	0.0081	0.0199	0.0111	0.0147
30 seeds	Mean	0.4971	0.5092	0.8027	0.8607	0.7892	0.7758
	C.I.	0.0110	0.0085	0.0036	0.0086	0.0049	0.0129
<i>Without P.M.</i>							
5 seeds	Mean	0.2126	0.3354	0.4046	0.6482	0.3897	0.8138
	C.I.	0.1216	0.1399	0.1319	0.0522	0.0921	0.0088
30 seeds	Mean	0.2127	0.3359	0.4018	0.6372	0.3871	0.8103
	C.I.	0.0488	0.0577	0.0537	0.0231	0.0377	0.0049
RAT (s)							
<i>With P.M.</i>							
5 seeds	Mean	129.663	141.456	226.402	202.442	200.400	195.022
	C.I.	10.422	6.730	10.149	15.102	13.779	15.099
30 seeds	Mean	128.843	132.460	228.495	198.159	196.056	198.251
	C.I.	4.178	5.208	3.607	6.948	6.499	12.926
<i>Without P.M.</i>							
5 seeds	Mean	63.128	95.711	107.442	167.329	111.023	213.585
	C.I.	36.520	45.289	38.419	33.477	31.526	29.885
30 seeds	Mean	58.743	93.597	108.845	171.429	108.019	233.413
	C.I.	14.230	18.272	16.127	13.114	12.103	9.905
E2E (s)							
<i>With P.M.</i>							
5 seeds	Mean	0.0775	0.1088	0.0465	0.1914	0.0800	0.0637
	C.I.	0.0491	0.0476	0.0083	0.0529	0.0155	0.0184
30 seeds	Mean	0.0789	0.0874	0.0469	0.1750	0.0745	0.0635
	C.I.	0.0208	0.0209	0.0042	0.0268	0.0070	0.0155
<i>Without P.M.</i>							
5 seeds	Mean	0.5509	0.5156	0.5943	0.3431	0.3102	0.0200
	C.I.	0.1520	0.2632	0.2938	0.1173	0.1114	0.0133
30 seeds	Mean	0.5671	0.5084	0.6103	0.3442	0.3042	0.0304
	C.I.	0.0614	0.1107	0.1272	0.0503	0.0495	0.0088
NRL							
<i>With P.M.</i>							
5 seeds	Mean	1.6561	2.4092	1.5935	5.3808	10.2996	17.5420
	C.I.	0.0331	0.0158	0.0450	0.0611	0.2137	0.4014
30 seeds	Mean	1.6730	2.4378	1.7024	5.7667	9.4928	17.3109
	C.I.	0.0172	0.0085	0.0176	0.1775	0.0836	0.3438
<i>Without P.M.</i>							
5 seeds	Mean	6.2837	4.0053	10.1743	11.3230	37.0032	3.0392
	C.I.	0.0204	0.1096	0.2588	0.1663	0.4880	0.1148
30 seeds	Mean	6.2066	3.9803	9.7491	11.0045	37.9442	3.9316
	C.I.	0.0154	0.0432	0.1069	0.0637	0.1834	0.0625

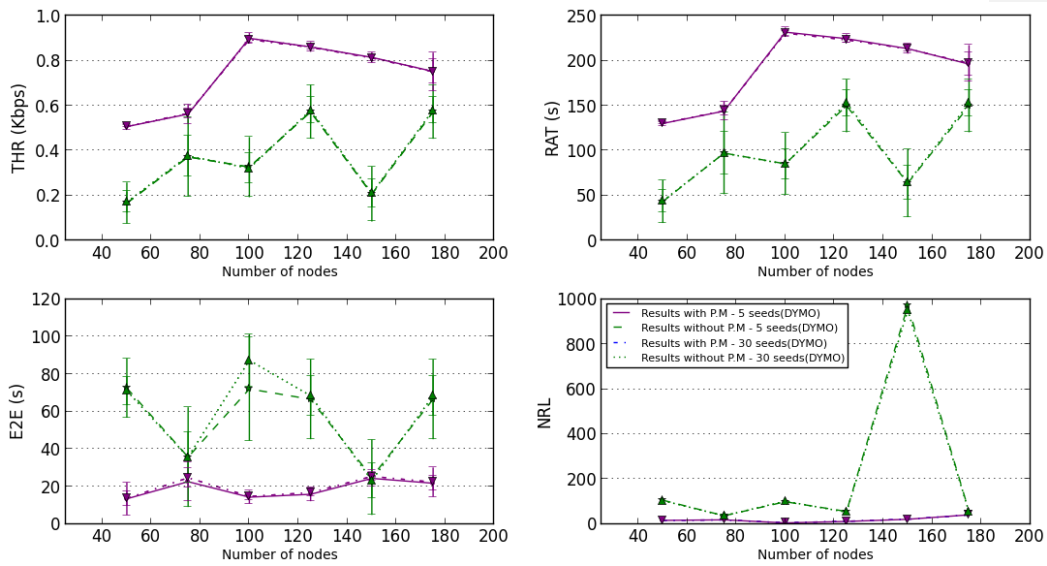


Figure 22 Results of the density analysis – DYMO

In Figure 22 we depict the performance metrics obtained by the DYMO routing protocol. We can see that the results are less dispersed when using the proposed methodology. (For more details see Table 16).

Table 16 Statistics measures for performance metrics vs number of nodes - DYMO

Number of Nodes		50	75	100	125	150	175
DYMO							
THR (Kbps)							
<i>With P.M.</i>							
5 seeds	Mean	0.5072	0.5630	0.8998	0.8614	0.8150	0.7526
	C.I.	0.0128	0.0432	0.0237	0.0214	0.0230	0.0531
30 seeds	Mean	0.5082	0.5658	0.8944	0.8594	0.8124	0.7509
	C.I.	0.0046	0.0220	0.0123	0.0114	0.0129	0.0854
<i>Without P.M.</i>							
5 seeds	Mean	0.1673	0.3726	0.3272	0.5729	0.2074	0.57296
	C.I.	0.0915	0.1760	0.1337	0.1195	0.1220	0.1195
30 seeds	Mean	0.17208	0.3769	0.3218	0.5801	0.2110	0.5801
	C.I.	0.0471	0.0906	0.065	0.065	0.0610	0.0585
RAT (s)							
<i>With P.M.</i>							
5 seeds	Mean	130.313	144.058	231.801	224.562	213.445	196.274
	C.I.	2.622	10.636	5.109	4.745	5.477	13.171
30 seeds	Mean	130.313	144.058	231.801	224.562	213.445	196.274
	C.I.	0.942	5.364	2.736	2.623	3.068	20.344
<i>Without P.M.</i>							
5 seeds	Mean	43.5044	97.3636	85.4762	149.915	63.9673	149.915
	C.I.	23.812	45.655	34.359	29.379	37.765	29.379
30 seeds	Mean	44.140	97.305	85.132	152.88	64.647	152.884
	C.I.	12.0723	23.180	16.961	14.523	18.705	14.523
E2E (s)							
<i>With P.M.</i>							
5 seeds	Mean	13.5061	22.6450	14.3689	15.9048	24.3982	21.8087
	C.I.	8.7146	10.5515	3.4104	3.6844	4.2207	3.8601
30 seeds	Mean	13.7178	24.7080	14.8118	16.7940	25.4729	22.4098
	C.I.	3.9479	5.0503	1.8438	2.0516	2.4619	7.8033
<i>Without P.M.</i>							
5 seeds	Mean	72.580	35.7447	72.0954	66.6104	25.0120	66.6104
	C.I.	15.8187	26.6106	27.5796	21.3251	20.1057	21.3251
30 seeds	Mean	71.082	35.7670	87.3907	68.4805	23.2911	68.4805
	C.I.	3.9479	5.0503	1.8438	2.0516	2.4619	7.8033
NRL							
<i>With P.M.</i>							
5 seeds	Mean	16.5322	18.9819	5.2775	11.8661	21.4601	41.2694
	C.I.	0.1695	0.0988	0.2035	0.4036	0.2306	0.6688
30 seeds	Mean	16.3235	18.8010	5.8642	12.6981	21.8431	40.9412
	C.I.	0.0621	0.0429	0.1149	0.2070	0.3080	1.4560
<i>Without P.M.</i>							
5 seeds	Mean	106.389	37.094	99.383	56.251	972.050	56.2519
	C.I.	1.553	0.194	1.770	1.195	10.362	1.1956
30 seeds	Mean	104.413	36.853	102.026	53.286	950.433	53.2866
	C.I.	1.120	0.083	1.033	0.668	7.617	1.4560

Congestion analysis

In this study we vary the number of connections, in steps of 5 connections each, within the interval [5, 25]. The number of nodes of the network is fixed to 125 nodes (high density scenario) and the APA values and the number of hops are the same than those selected in the previous analysis (See Table 3 and Table 5 for more details).

Figure 23 shows the simulation results for THR, RAT, E2E, NRL performance metrics with their confidence intervals for each number of nodes. Figure 23 includes the results using the proposed methodology (Results with P.M. in Figure 23) and without using the proposed methodology (Results without P.M. in Figure 23). Once again, we depict the results for 5 and 30 simulations and we verify that by using the proposed methodology more reliable results are obtained despite of using a lower number of simulations (See Table 17 for more details)

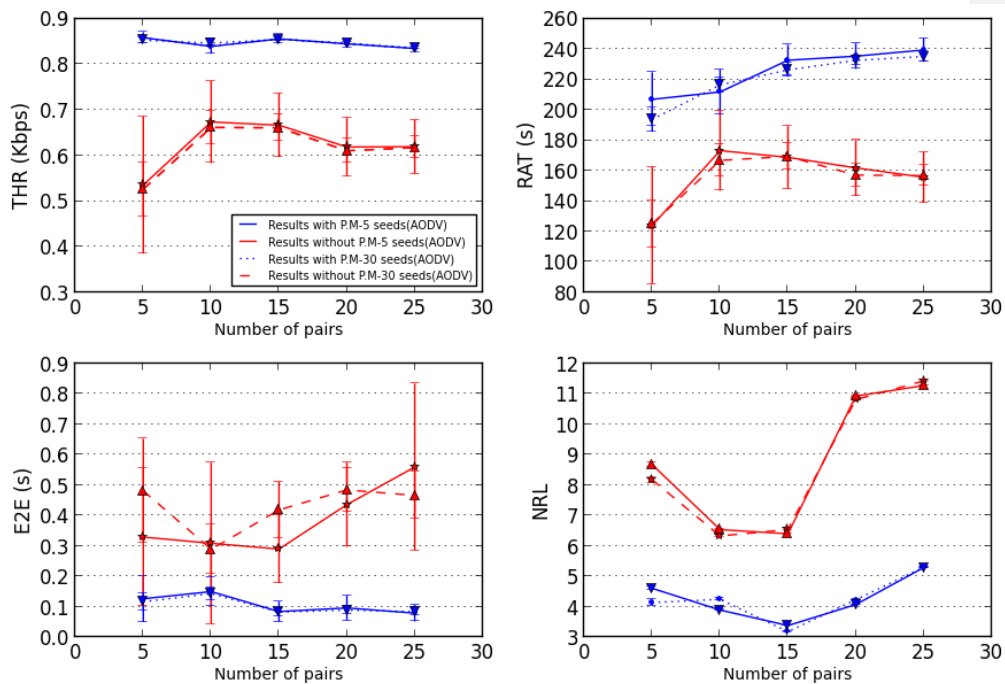


Figure 23 Results of the congestion analysis - AODV

Table 17 Statistics measures for performance metrics vs number of pairs - AODV

Number of Pairs		5	10	15	20	25
AODV						
<i>THR (Kbps)</i>						
<i>With P.M.</i>						
5 seeds	Mean	0.8582	0.8392	0.8557	0.8448	0.8346
	C.I.	0.0119	0.0160	0.0098	0.0082	0.0086
30 seeds	Mean	0.8526	0.8464	0.8546	0.8461	0.8359
	C.I.	0.0063	0.0061	0.0036	0.0036	0.0035
<i>Without P.M.</i>						
5 seeds	Mean	0.5355	0.6735	0.6664	0.6187	0.6193
	C.I.	0.1504	0.0898	0.0700	0.0637	0.0590
30 seeds	Mean	0.5262	0.6615	0.6608	0.6104	0.6173
	C.I.	0.0591	0.0360	0.0292	0.0261	0.0238
<i>RAT (s)</i>						
<i>With P.M.</i>						
5 seeds	Mean	206.923	211.773	232.659	235.301	239.376
	C.I.	17.771	14.754	10.157	8.420	7.762
30 seeds	Mean	193.780	216.385	226.776	232.539	235.134
	C.I.	8.019	5.125	4.387	3.316	3.182
<i>Without P.M.</i>						
5 seeds	Mean	124.043	173.201	168.820	161.880	155.708
	C.I.	38.520	26.256	20.532	18.419	16.549
30 seeds	Mean	125.124	166.788	169.387	157.127	156.794
	C.I.	15.413	10.331	8.578	7.420	6.730
<i>E2E (s)</i>						
<i>With P.M.</i>						
5 seeds	Mean	0.1269	0.1505	0.0850	0.0966	0.07975
	C.I.	0.0758	0.0459	0.0327	0.0409	0.0272
30 seeds	Mean	0.1173	0.1436	0.0831	0.0894	0.0844
	C.I.	0.0274	0.0202	0.0135	0.0130	0.0131
<i>Without P.M.</i>						
5 seeds	Mean	0.3297	0.3092	0.2902	0.4364	0.5598
	C.I.	0.2257	0.2650	0.1108	0.1386	0.2744
30 seeds	Mean	0.4817	0.2895	0.4189	0.4846	0.4668
	C.I.	0.1710	0.0802	0.0922	0.0723	0.0766
<i>NRL</i>						
<i>With P.M.</i>						
5 seeds	Mean	4.1438	4.2499	3.1861	4.2388	5.3047
	C.I.	0.1047	0.0385	0.0181	0.0222	0.01427
30 seeds	Mean	4.6115	3.8982	3.3883	4.0784	5.2930
	C.I.	0.0623	0.0186	0.0093	0.0069	0.0055
<i>Without P.M.</i>						
5 seeds	Mean	8.1798	6.3340	6.5440	10.8139	11.4277
	C.I.	0.0938	0.0269	0.0368	0.01402	0.0282
30 seeds	Mean	8.6758	6.5426	6.4049	10.9307	11.2774
	C.I.	0.0649	0.0224	0.0113	0.0107	0.0101

Figure 24 shows again that a smoother tendency is obtained for the congestion analysis when the proposed methodology is used. Similarly, the confidence intervals obtained with the proposed methodology are lower than those obtained when it is not applied. In this analysis (See Figure 24) the proposed methodology also provides better results as the network congestion increases (See Table 18 for more details).

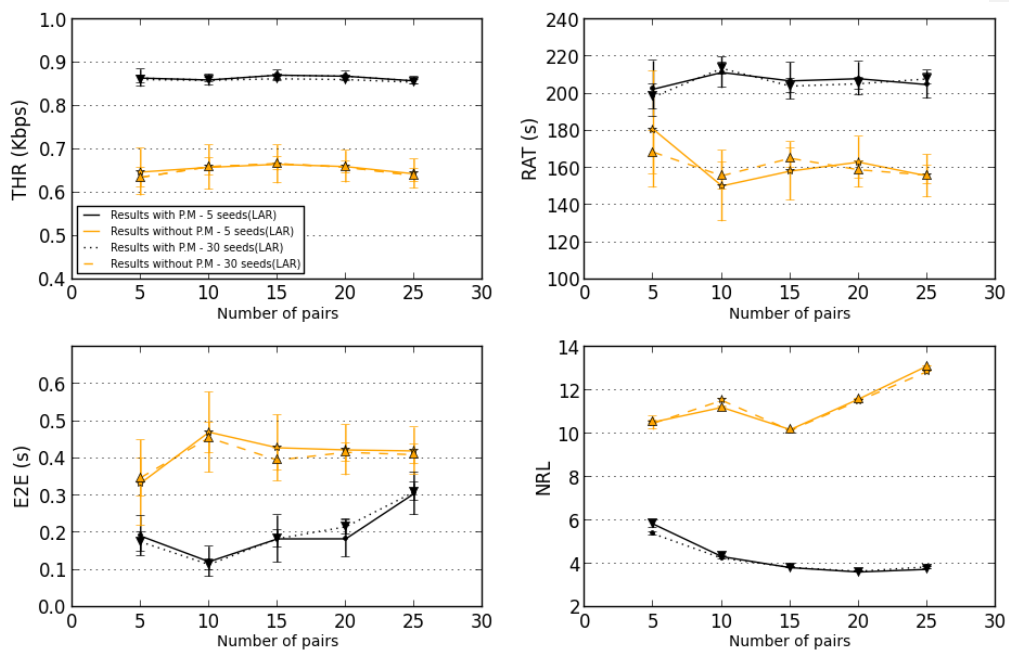


Figure 24 Results of the congestion analysis - LAR

Table 18 Statistics measures for performance metrics vs number of pairs – LAR

Number of Pairs		5	10	15	20	25
LAR						
<i>THR (Kbps)</i>						
<i>With P.M.</i>						
5 seeds	Mean	0.8645	0.8601	0.8712	0.8688	0.8581
	C.I.	0.0199	0.0133	0.0112	0.0098	0.0093
30 seeds	Mean	0.8610	0.8595	0.8628	0.8611	0.8549
	C.I.	0.0080	0.0054	0.0043	0.0036	0.0041
<i>Without P.M.</i>						
5 seeds	Mean	0.6477	0.6585	0.6656	0.6603	0.6436
	C.I.	0.0539	0.0513	0.0440	0.0365	0.0339
30 seeds	Mean	0.6356	0.6601	0.6669	0.6591	0.6396
	C.I.	0.0224	0.0199	0.0157	0.0132	0.0141
<i>RAT (s)</i>						
<i>With P.M.</i>						
5 seeds	Mean	202.442	211.412	206.878	208.050	204.882
	C.I.	15.102	8.350	9.819	8.958	7.529
30 seeds	Mean	198.321	213.621	204.071	205.322	207.927
	C.I.	6.601	3.203	3.925	3.100	3.084
<i>Without P.M.</i>						
5 seeds	Mean	180.722	150.339	158.320	163.162	155.743
	C.I.	31.306	19.054	15.669	13.654	11.530
30 seeds	Mean	168.474	155.804	165.221	159.049	156.168
	C.I.	11.681	7.366	5.427	4.736	4.811
<i>E2E (s)</i>						
<i>With P.M.</i>						
5 seeds	Mean	0.1914	0.1224	0.1837	0.1839	0.3057
	C.I.	0.0529	0.0405	0.0629	0.0483	0.0574
30 seeds	Mean	0.1757	0.1148	0.1840	0.2163	0.3102
	C.I.	0.0256	0.0146	0.0228	0.0207	0.0241
<i>Without P.M.</i>						
5 seeds	Mean	0.3338	0.4709	0.4289	0.4231	0.4204
	C.I.	0.1161	0.1082	0.0889	0.0680	0.0630
30 seeds	Mean	0.3486	0.4558	0.3957	0.4167	0.4104
	C.I.	0.0510	0.0414	0.0290	0.0251	0.0259
<i>NRL</i>						
<i>With P.M.</i>						
5 seeds	Mean	5.3808	4.2594	3.8384	3.6584	3.8578
	C.I.	0.061	0.047	0.015	0.006	0.006
30 seeds	Mean	5.8295	4.3325	3.8238	3.6221	3.7453
	C.I.	0.1699	0.0357	0.0063	0.0047	0.0035
<i>Without P.M.</i>						
5 seeds	Mean	10.4852	11.5379	10.1581	11.5247	12.8672
	C.I.	0.2965	0.0612	0.0271	0.0124	0.0283
30 seeds	Mean	10.539	11.216	10.192	11.598	13.123
	C.I.	0.0804	0.0252	0.0144	0.0080	0.0072

Regarding the DYMO routing protocol (See Figure 25), the results are also not dispersed. These are better (in terms of mean) and we can see a smoother tendency of the performance metrics when applying the proposed methodology (For more details see Table 19).

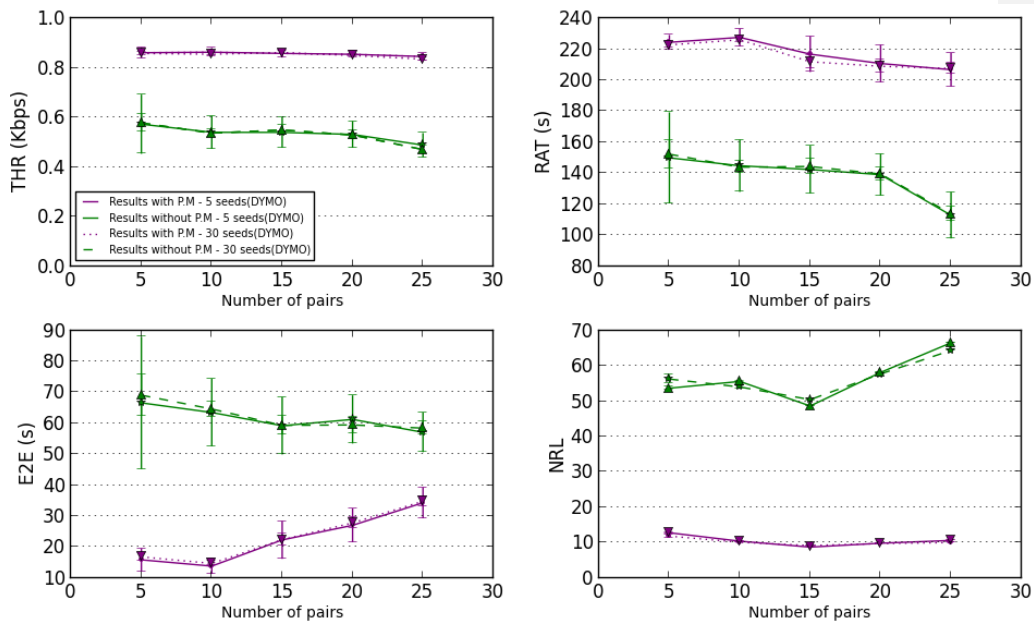


Figure 25 Results of the congestion analysis – DYMO

Table 19 Statistics measures for performance metrics vs number of pairs – DYMO

Number of Pairs		5	10	15	20	25
DYMO						
<i>THR (Kbps)</i>						
<i>With P.M.</i>						
5 seeds	Mean	0.8614	0.8636	0.8587	0.8551	0.8461
	C.I.	0.0214	0.01693	0.0164	0.0134	0.0146
30 seeds	Mean	0.8585	0.8554	0.8620	0.8502	0.8354
	C.I.	0.0071	0.0047	0.0053	0.0045	0.0056
<i>Without P.M.</i>						
5 seeds	Mean	0.5729	0.5400	0.5398	0.5315	0.4884
	C.I.	0.1195	0.0664	0.0620	0.0535	0.0484
30 seeds	Mean	0.5780	0.5368	0.5495	0.5290	0.4703
	C.I.	0.0360	0.0154	0.0200	0.0176	0.0159
<i>RAT (s)</i>						
<i>With P.M.</i>						
5 seeds	Mean	224.562	227.734	216.916	210.874	206.888
	C.I.	4.745	5.709	11.521	11.972	10.794
30 seeds	Mean	222.868	226.217	211.986	209.079	207.642
	C.I.	1.628	1.573	4.525	4.117	3.691
<i>Without P.M.</i>						
5 seeds	Mean	149.915	144.814	142.371	139.053	112.827
	C.I.	29.379	16.855	15.374	13.429	14.962
30 seeds	Mean	152.285	143.882	144.351	139.452	113.793
	C.I.	8.972	3.864	4.987	4.419	4.910
<i>E2E (s)</i>						
<i>With P.M.</i>						
5 seeds	Mean	15.9048	13.8787	22.3367	27.0592	34.3470
	C.I.	3.6844	2.5246	6.1019	5.4334	5.0076
30 seeds	Mean	16.8680	14.7454	22.4142	27.8414	34.8339
	C.I.	1.2619	0.7011	1.9091	1.7882	1.6750
<i>Without P.M.</i>						
5 seeds	Mean	66.6104	63.5052	59.2219	61.2661	57.1043
	C.I.	21.3251	10.8777	9.2749	7.7797	6.4067
30 seeds	Mean	69.0914	64.6261	59.3388	59.4877	58.4515
	C.I.	6.6084	2.4876	2.9497	2.5537	2.1852
<i>NRL</i>						
<i>With P.M.</i>						
5 seeds	Mean	11.8661	10.1422	9.2064	9.7781	10.0558
	C.I.	0.4036	0.0658	0.0688	0.0222	0.0162
30 seeds	Mean	12.8203	10.4430	8.7722	9.8621	10.6552
	C.I.	0.1274	0.0199	0.0185	0.0102	0.0092
<i>Without P.M.</i>						
5 seeds	Mean	56.2519	54.0320	50.4579	57.7146	64.3516
	C.I.	1.1956	0.3271	0.1449	0.0851	0.1476
30 seeds	Mean	53.6853	55.5944	48.6554	58.1439	66.5199
	C.I.	0.4110	0.0921	0.0678	0.0417	0.0497

3) Comparison of the routing protocols using the proposed methodology

This subsection is aimed at providing a fair and unbiased comparison between the three used routing protocols such as AODV, DYMO, and LAR. We evaluate the mentioned routing protocols under different network conditions using the proposed APA metric and number of hops. In Figure 26, we depict the throughput and NRL metrics for AODV, LAR and DYMO for different APA values. We vary the APA value in steps of 0.2. As it is shown in Figure 26, the throughput of the three routing protocols increases as the APA value also increases. The reason is that the routes between source and destination nodes are available during more time. As a consequence, the number of delivery packets is higher. According to the results, DYMO has the best performance metrics for high APA values while LAR is the best one for low APA values. However the NRL decreases when the APA value increases because it is not necessary to initiate new discovery phases. The reason is that the routes are more time available so the number of routing packets decreases. In terms of NRL, DYMO presents the worst behavior because the necessary routing information (routing packets) is higher.

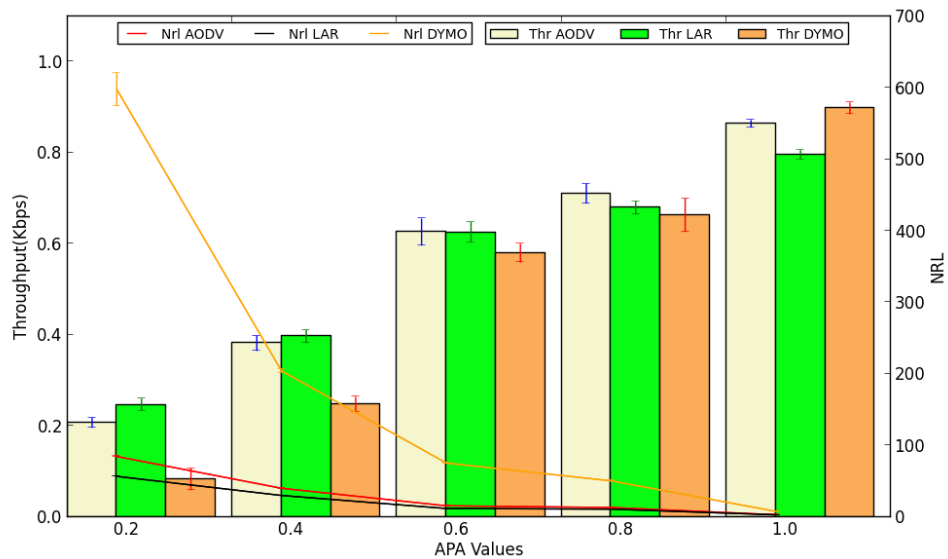


Figure 26 Throughput and NRL vs APA

Now we evaluate the routing protocols versus the number of hops between the source and destination nodes (See Figure 27). We also evaluate the three routing protocols in terms of throughput and NRL metrics. For the three routing protocols, the throughput metric is better for low numbers of hop values (See Figure 27). This situation corresponds to the one in which the destination node is near the source node (in terms of hops). Consequently, the probability of losing data packets is lower. Yet, when the number of hops increases the throughput decreases for the three routing protocols. DYMO has the best performance for low number of hops because this one is able to generate routes entries for each intermediate hop. However, when the distance between source and destination node is higher the routing data packets are also higher. In consequence, the delivery of packets decreases. Regarding the NRL, it increases for the three used routing protocols because the number of discovery phases is higher when the number of hops increases. As we mentioned before DYMO has the worst behavior because the routing information is higher.

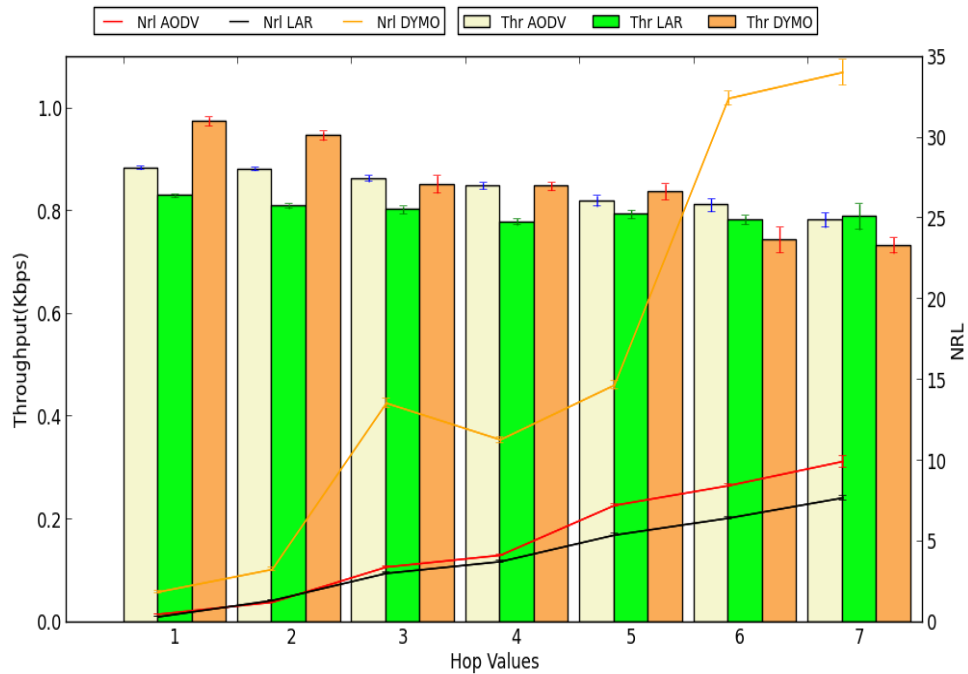


Figure 27 Throughput and NRL vs number of hops

5. Conclusions

Simulation-based studies of routing protocols for VANETs involve selecting, adjusting, and setting a large number of simulation parameters, which can affect significantly the simulation results. In this paper, we propose a methodology to conduct reliable simulations of routing protocols based on a set of good simulation practices for VANET scenarios. The proposed methodology includes aspects such as: i) *Measurement period*, to ensure that all the simulation measurements begin and end at the same time, consequently we improve the performance metric mean. ii) *Source destination pair selection*, to avoid discrepancies in terms of path availability and number of hops between the source and destination nodes and to obtain results lesser scattered, iii) *The number of trials*, to obtain reliable measurements. iv) *Mobility models*, based on maps

to emulate mobility of vehicles in urban scenarios. v) *The importance of the repetition of source and destination.* vi) *Performance metrics and simulation analyses*, to evaluate routing protocols under different conditions. We have shown the importance of selecting these simulation parameters carefully in order to obtain reliable simulation statistics and make a fair and unbiased evaluation of routing protocols in VANET scenarios. Finally, we have validated the proposed methodology by conducting a comparison of the AODV, LAR and DYMO routing protocols with and without using the proposed methodology. The obtained simulation results demonstrate that the proposed methodology provides better results in terms of reliability (confidence intervals), and a smoother tendency of the performance metrics.

Acknowledgements

This work was supported in part by the University of Seville under the V plan propio de investigación and by the Ministerio Español de Economía y Competitividad under the contract of Jesús Sánchez García associated to the project DPI2013-44278-R.

References

- [1] [R. Jurdak, C.V. Lopes, P. Baldi. A survey, classification and comparative analysis of medium access control protocols for ad hoc networks. Communications Surveys & Tutorials, IEEE. Vol.6, pp. 2-16. 2004](#)
- [2] [L. Junhai, Y. Danxia, X. Liu, F. Mingyu. A survey of multicast routing protocols for mobile Ad-Hoc networks IEEE. Communications Surveys & Tutorials. Vol 11, pp. 78-91. 2009.](#)

- [3] [B.T. Sharef, A.R. Alsaqour, M.Isamil .Vehicular communication ad hoc routing protocols: A survey. Journal of Network and Computer Applications. Vol. 40, pp. 363-396. 2013.](#)
- [4] [D. G. Reina, S. L. Toral, P. Jonshon, F. Barrero. A survey on probabilistic broadcast schemes for wireless ad hoc networks. Ad Hoc Networks. Vol. 25, pp. 263-292. 2015.](#)
- [5] O. Younes, N. Thomas. Modelling and performance analysis of multi-hop ad hoc networks. Simulation Modelling Practice and Theory. Vol. 38, pp. 69-97. 2013.
- [6] A. B. Cruz, T. Boutaleb, H. Tianfield. Simulation of service-oriented systems for Mobile Ad hoc Networks. Simulation Modelling Practice and Theory. Vol. 32, pp. 42-63. 2013.
- [7] B. Blywis, M. Günes, F. Juraschek, O. Hahm, N. Schmittberger. Properties and Topology of DES-Testbed (2nd Extended Revision). Telematic and Computer System. Freie Universität Berlin. 2011.
- [8] [H. Pucha, S. M. Das, Y. C. Hu. The performance impact of traffic patterns on routing protocols in mobile ad hoc networks. Computer Networks. Vol 51, pp. 3595-3616. 2007.](#)
- [9] [A.M. Mezher, J.J. Oltra, L.Urquiza-Aguiar, C.I. Paredes, C. Trip-Barba, M.A. Igartua. Realistic Environment for VANET Simulations to Detect the Presence of Obstacles in Vehicular Ad Hoc Networks. PEWASUN'14, pp. 77-84. 2014](#)
- [10] [C. Perkins, E. Royer, and S. Das, Ad hoc On-Demand Distance Vector \(AODV\) Routing. IETF, RFC 3561, 2003.](#)
- [11] [Y-B. Ko and N. Vaidya. "Location-Aided Routing \(LAR\) in mobile ad hoc networks." Wireless Networks, pp.307-321. 2000](#)

- [12] [D. B. Johnson, D. A. Maltz, Y-C. Hu, "The Dynamic Source Routing Protocol for Mobile Ad Hoc Networks" IETF, RFC 4728, 2004](#)
- [13] [S. Kurkowski, T. Camp, and M. Colagrosso. MANET Simulation Studies: The Incredibles. ACM's Mobile Computing and Communications Review. Vol. 9, pp. 50-61. 2005.](#)
- [14] [J. Yoon, M. Liu, and B. Noble. Random Waypoint Considered Harmful. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications \(INFOCOM 2003\), pp. 1312-1321. 2003.](#)
- [15] [S. Kurkowski, W. Navidi, and T. Camp. Constructing MANET Simulation Scenarios That Meet Standards. IEEE International Conference on Mobile Adhoc and Sensor Systems \(MASS 2007\), pp. 1-9. 2007.](#)
- [16] [T. R Andel and A. Yasinsac. On the credibility of Manet Simulations. Computer. Vol. 37, pp. 48-54. 2006.](#)
- [17] [D. Hiranandani, K. Obraczka, and J. J. García-Luna-Aceves. MANET protocol simulations considered harmful: the case for benchmarking. IEEE Wireless Communications, Vol. 20, pp. 82-90. 2013.](#)
- [18] [M. Treiber, A. Hennecke, and D. Helbing. Congested traffic states in empirical observations and microscopic simulations. Physical Review E. Vol. 62, pp.1805. 2000](#)
- [19] [J. Friginal, D. Andrés, J-C. Ruiz, M. Martínez. A Survey of Evaluation Plataforms for Ad hoc Routing Protocols: A Resilience Perspective. Computer Network. Vol.75, pp.395-415. 2014](#)
- [20] [F. Bai, N. Sadagopan, A. Helmy. IMPORTANT: framework to systematically analyze the Impact of Mobility on Performance of Routing protocols for Adhoc](#)

- NeTworks. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications (INFOCOM 2003), pp.825-835. 2003.
- [21] M. Hawa, S. Taifour, M. Qasem, W. Tuffaha. A dynamic ic cross-layer routing protocol fot mobile Ad hoc Networks International Journal of Electronics and Communications. Vol.66, pp.996-1005. 2012
- [22] R. Shukla and N. Tyagi. Performance of VANET Routing Protocols Using Realistic Mobility Model. Proceedings of the International Conference on Soft Computing for Problem Solving (SocProS 2011), pp. 549-555. 2011.
- [23] D. Krajzewicz, G. Hertkorn, C. Rossel, and P. Wagner. SUMO (Simulation of Urban MObility) - An open-source traffic simulation. In Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM2002), pp. 183–187. 2002.
- [24] J. Miller and E. Horowitz. FreeSim – A Free Real-Time Freeway Traffic Simulator. Intelligent Transportation Systems Conference, pp. 18-23. 2007.
- [25] F. J. Martinez, J.-C. Cano, C. T. Calafate, and P. Manzoni. Citymob: a mobility model pattern generator for VANETs. In IEEE Vehicular Networks and Applications Workshop (Vehi-Mobi, held with ICC), pp. 370-374. 2008.
- [26] C4R: CityMob for Roadmaps. Available at <http://www.grc.upv.es/Software/c4r.html> (last access: 07/03/2015)
- [27] OpenStreetMap, collaborative project to create a free editable map of the world, 2009. Available at <http://www.openstreetmap.org>.
- [28] M. Fogue, P. Garrido, F.J. Martinez, J-C. Cano, C.T. Calafate, P. Manzoni. A Realistic Simulation Framework for Vehicular Networks. SIMUTOOLS '12

Código de campo cambiado

Proceedings of the 5th International ICST Conference on Simulation Tools and Techniques, pp. 37-46. 2012.

- [29] [S. Krauss, P. Wagner, and C. Gawron. Metastable states in a microscopic model of traffic flow. Physical Review. Vol.55, pp.5597–5602. 1997.](#)
- [30] D. Krajzewicz, G. Hertkorn, C. Rossel, and P. Wagner. SUMO (Simulation of Urban MObility) - An open-source traffic simulation. In Proceedings of the 4th Middle East Symposium on Simulation and Modelling (MESM2002), pp. 183–187. 2002.
- [31] P. Wagner. How human drivers control their vehicle. The European Physical Journal B - Condensed Matter and Complex Systems . Vol. 52, pp.427-431. 2006.
- [32] B. S. Kerner. Experimental features of self-organization in traffic flow. Phys. Rev. Lett. Vol. 81, pp.3797-3800. 1998.
- [33] M. Fogue, P. Garrido, F.J. Martinez, J. C. Cano, C.T. Calafate and P.Manzoni, Identifying the Key Factors Affecting Warning Message Dissemination in VANET Real Urban Scenarios. Sensors. Vol.13, pp. 5220-5250. 2013
- [34] S. I. Chowdhury, W. I. Lee, Y-S Choi, G-Y Kee, and J-Y Pyun. Performance Evaluation of Reactive Routing Protocols in VANET. Communications (APCC), 2011 17th Asia-Pacific Conference on, pp. 559-564. 2011.
- [35] Available at <http://www.isi.edu/nsnam/ns/> (last access: 07/03/2015)
- [36] IEEE Draft Standard for Information Technology - Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC)

and Physical Layer (PHY) Specifications; Amendment: Prioritization of management frames environments, pp. 1-54. 2011.

- [37] W-Y. Lin, M-W. Li, K-C. Lan and C-H. Hsu. A Comparison of 802.11a and 802.11p for V-to-I Communication: A Measurement Study. 7th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness, QShine 2010, and Dedicated Short Range Communications Workshop, pp. 559-570. 2010.