

Fourth International Conference on Selected Topics in Mobile & Wireless Networking
(MoWNet'2014)

Performance assessment of an epidemic protocol in VANET using real traces

Raul Amici^a, Marco Bonola^a, Lorenzo Bracciale^{a,*}, Antonello Rabuffi^a, Pierpaolo Loreti^a,
Giuseppe Bianchi^a

^aDIE, Università di Roma "Tor Vergata", Rome, Italy

Abstract

Many vehicular ad-hoc network protocols have been validated using complex urban mobility simulators or by means of the few publicly available real mobility traces. This work presents an extensive measurement campaign of the positions of a fleet of 370 taxi cabs moving around the city of Rome, Italy. Due to its street network and its traffic conditions, Rome presents a characteristic mobility pattern representative of an ancient city with heavy road congestion, and therefor provides a valuable test case to experiment VANET protocols. We exploit these traces to run a set of experiments to assess the performance of a simple epidemic protocol that we compare with the basic random waypoint model in order to quantify how far the performance metrics are from this baseline. The results show the possible outcomes of implementing data dissemination through an opportunistic network that uses taxi cabs as an information vector.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Peer-review under responsibility of organizing committee of Fourth International Conference on Selected Topics in Mobile & Wireless Networking (MoWNet'2014)

Keywords: Delay tolerant network; VANET; Real Traces; Smart City; Epidemic routing

1. Introduction

The recent advances in technologies¹ and in the standardization of the elements that compound the Intelligent Transportation System^{2,3} bring new attention on the theme of vehicles to vehicles communication for a smarter and safer road traffic. In this field, many researchers proposed several protocols to design and implement new Vehicular ad-hoc networks (VANET)^{4,5}, usually testing them through simulations.

However, this validation methodology presents several complexities and some possible pitfalls due to the time and space dependent factors that influences the road conditions, from the weather to the presence of traffic lights, roads

* Corresponding author

E-mail address: lorenzo.bracciale@uniroma2.it



Fig. 1. The different road topologies for the city of San Francisco, Shenzhen, and Rome

with multiple lanes, car accidents and so on.

Many complex simulators have been build⁶ to deal with all these elements, and are able to accurately map the road condition once they are properly feed with the right tuning parameters. Nevertheless this tuning is far from being trivial especially for network researchers, who often make use of simplified models to validate VANET protocols. For this reason, some protocols have been validated against real traces collected by monitoring the taxi cabs position in the San Francisco Bay Area and in Shenzhen⁷.

This methodology exhibits the shortcoming of not-providing any flexibility degree (e.g. for varying the number of vehicles or their speed) but relieves network researchers from the burden of deeply characterize the urban mobility. In this work we present a study on the performance of a simple epidemic routing protocol using real traces acquired by an extensive measurement campaign in the city of Rome: 370 taxi cabs that report their position every 7 seconds for a period of 6 months.

Due to the road topology of Rome, which is far from the grid topology available in literature (figure 1), and its traffic conditions, the evaluated scenario and the proposed traces expand the testing cases available for validating the performance of the VANET protocols. At the same time the presented traces has been submitted to the Crawdad archive⁷ to make them available for other researchers.

This study can be useful in the smart city scenario for assessing the feasibility and the performance of a metropolitan VANET where taxi cabs carry the messages emitted by some monitoring devices (such as pollution sensors or traffic sensor etc) towards a sink node (e.g. a server of the municipality). In this way, the devices could communicate opportunistically to a cab and do not require any 3G/4G connections using the “better than nothing” network infrastructure, hence reducing the maintenance costs if the delay performance are tolerated by the specific service.

The paper is organized as follows: in section 2 we present the state of the art regarding the VANET validation strategy; in section 3 we describe the trace acquisition procedure as well as the filtering methodology to partially correct some position errors, typical to all traces reported via smartphones; then we characterize the statistics of the traces and finally in section we present several popular performance metrics such as the coverage and the message propagation speed, and we compare them with the popular naive random waypoint to provide a baseline.

2. Related works

Synthetic traces and mobility models for VANET

Usually, assessment of VANET protocols is performed by combining two different methods for: i) generating movements of vehicles in the reference area using a **mobility simulator**; ii) simulating the communication among nearby vehicles with a **network simulator** (e.g. with NS2).

In the simplest case, VANET protocols are validated using only synthetic movement traces (produced by mobility simulators) so that cars can exchange data if their distance is below a given threshold (that depends from many factors, but it is usually set according to the adopted technology, such as 802.11b/g/a or the more recent 802.11p¹).

A realistic mobility model should take into consideration many aspects of real world mobility (such as one way streets, traffic lights, obstacles, weather conditions, drivers behaviour etc.) that are complex to model and to take into account. For this reason, several simulators have been developed.

IMPORTANT framework⁸ is one of the first simulators for MANET protocols. It combines several mobility models including the naive random waypoint (RWP) algorithm, the Reference Point Group Mobility (RPGM) model for group mobility, the Freeway mobility model (that focuses on the mobility of a single high speed street) and the Manhattan mobility model where vehicles move on a grid. According to any of these models, the simulator can produce a connectivity graph that can be used for assessing the performances of different routing protocols. As pointed out by the authors, RWP presents several limitations because it neglects temporal, spatial and geographical dependencies of the vehicles. However it is still one of the “most commonly used mobility model in MANET research community”.

VanetMobiSim^{9,10} is an extension of the CANU Mobility Simulation Environment (CanuMobiSim) for vehicular mobility. It can import maps from the US Census Bureau TIGER/Line database, or randomly generate them using Voronoi tessellation. It aims to simulate a near-to-reality scenario, hence it offers the support for many mobility features such as multi-lane roads, separate directional flows, differentiated speed constraints and traffic signs at intersections. For this reason, it must be tuned using several complex parameters that are often hard to quantify.

TRANSIM¹¹ is an integrated set of tools developed to conduct regional transportation system analyses. It integrates several aspects of mobility, including the synthetic behaviour of public transportation and many algorithms to mimic the regional population for matching the real population demography. It uses cellular automata to simulate interactions among vehicles.

Another alternative is SUMO¹², an open source mobility simulator that uses a Gipps-model to simulate the main features of traffic flow, while taking into account a wide range of editable features such as traffic lights etc.

Due to the complexity of these simulators, often the researchers resort to simpler models such as the naive Random Waypoint, or Random Waypoint City Model¹³ (that includes city maps), or STRAW¹⁴, which adds vehicular congestion and simplified traffic control mechanisms.

All these models could be enriched by adding more parameters (coordinated traffic lights, hazards, etc, see⁶ for a complete survey), paying the price of adding a greater complexity to the already complex environment.

In¹⁵ authors analyse the different impact of the features introduced to simulate urban mobility. They then draw the conclusion that some features, such as waiting at the intersections, affect the simulation results in a more significant manner than other features (multiple lines, coordinated traffic lights).

In this paper we do not deal with mobility models but we take the vehicular positions “as they are”, analysing several real mobility traces acquired by tracking a fleet of taxi cabs. We compare them to the basic random waypoint model not to demonstrate that the model does not fit with the reality, but to quantify how far the performance metrics are from that. Moreover, RWP, due to its simplicity, has been used in several works as a baseline¹⁶.

Real traces for VANET

Some works use real traces¹⁷ to study the storage capability of a VANET¹⁸ or the dynamics of network topologies¹⁹. Usually these real traces are provided by tracking public transportation vehicles. Among them, the traces obtained by tracking taxi cabs are particularly important as they better characterize the status of the city streets.

In this field, the most used public domain traces available in literature are the GPS traces of 533 taxis collected in 20 days in the San Francisco Bay area, USA⁷ and the traces of 13,799 taxi cabs collected in 9 days in the city of Shenzhen. In this paper we present an analysis performed in the city of Rome, that exhibits a different topology if compared to the Shenzhen and San Francisco Bay area (almost) grid topology 1.

Another alternative is to use some other traces available by some public services (eg. NOKIA Sports Tracker) acquired by logging mobile phone positions. However these traces lasts only for tiny segments of time, making them not suitable for an extensive analysis. Moreover, it is not guaranteed that the samples in the trace refer to vehicles and not to pedestrians, train etc.

3. Evaluation scenario

In this section we describe some technical details regarding the mobility traces and how they were acquired. Given that they were obtained by tracking the real movement around the city of a set of taxi drivers, it is particularly important to properly define the context in terms of the location points acquiring procedure and in terms of driver behaviours.

Traces acquisition methodology

Each driver is equipped with an Android OS tablet device running an app that updates the current GPS position every 7 seconds towards a server. On the application side, the position is updated using the *getLastKnownLocation* method of the *LocationManager* Android object and it is filtered against its precision, using the *getAccuracy()* function. This function returns the estimated accuracy in meters with 68% of probability (according to the official Android documentation). A sample is accepted only if the accuracy is less than 20m and it is discarded otherwise.

The collected traces are further filtered to mitigate some localization errors. In particular, after a first post processing stage we noticed some “oddities” which are easily recognizable due to a unreasonable distance between two subsequent points. In some cases such distance is longer than 125m that corresponds to a speed greater than 64 km/h. A reasonable upper bound considering that the speed limit in the downtown is 50 km/h. After a further post processing analysis we noticed that these oddities usually happen when the drivers are in some part of the city where the GPS service quality is poor (e.g. tunnels, high buildings etc.). In these case we considered three strategies: (i) first, if the anomaly duration is less than 42 s (i.e. 6 points) we simply discard the “bad” samples; (ii) second, if the anomaly duration is greater than 42s (but less than 8 minutes) we correct the traces by introducing artificial samples according to the short path between the endpoints of the gap. For this task we use the Open Streetmap database; (iii) finally, if the gap is greater than 8 minutes, we consider it as a service interruption (lunch break, shift closure, etc) and we take no further action.

Description of the sampled scenario

We limit our analysis to the center of Rome, where the density of the taxi cabs is relevant. We considered an area of 8x8 [km x km] whose bounds are given by the coordinates pairs (41.856, 12.442) (41.928, 12.5387).

This scenario is characterized by very thin and congested roads, high traffic volume and slow speed. The area has been analysed using a 32x32 grid. A grid cell covers a square area of 250x250 [m x m].

Statistical parameters of the evaluated traces

We took a sample of five months (October 2013 to February 2014) to derive the statistical parameters coming out from the traces. During day time, the taxi drivers can either move to serve customers (we call these movements “rides”) or stay in parking areas if they are idle (we refer to this as the waiting time). As for the taxi cab speed, we processed data from 1359 rides and obtained an average car speed of 12.29 km/h with a variance of 19.15. The mean waiting time is 600 s. The CDF of the average speed for each ride is shown in figure 2. As we can see, there are few cabs moving with an average speed greater than 25 km/h. This is a reasonable average speed considering stop times (e.g. traffic lights) and traffic congestion. We point out that this value is about half of the speed limit in the area (50 km/h).

Figure 3 shows the car density in the different part of the reference area. In particular, for each cell, the overlay shows the probability of finding at least one car in each cell of the grid (in a reference time period). If we consider this period equals to 6 hours (figure 3 a)) there is an high probability that the taxi cabs cover almost the entire area with the exception of some zones not covered by the road network such as graveyards, public gardens, railway stations etc. As we consider a smaller reference period of 1 hour (figure 3 b)) this probability is greater than 0.8 only in the very center of the area and alongside the principal streets. If we shorten even more the period to 10 minutes as represented in figure 3 c) the great part of the area has a probability less than 0.2. As a baseline result to compare with, in figure 4 we plot the same study for the case of random waypoint movement and a reference period of 10 minutes. As we can

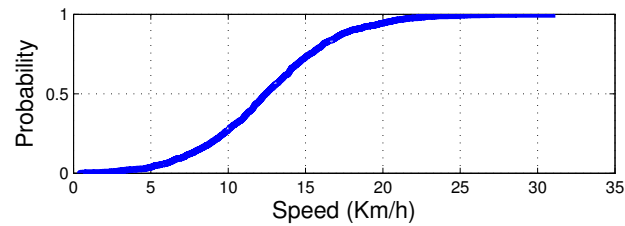


Fig. 2. CDF of the average speed for each considered ride

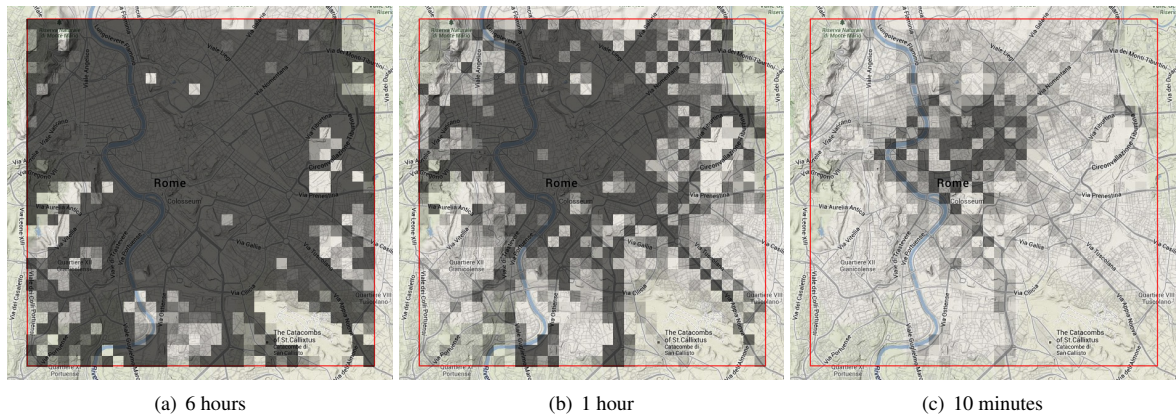


Fig. 3. Probability that at least one car enters in a cell varying the reference period using the real traces. The grey scale overlay represents the probability from 0 to 1 with step of 0.1: darken area means a greater probability

see, the area is already covered only after 10 minutes while the whole area appears as fully covered after just one hour (we omit the figure for brevity).

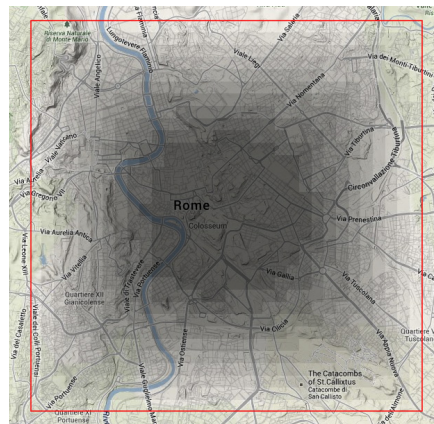


Fig. 4. Probability that at least one car enters in a cell varying the reference period using random waypoint. The grey scale overlay represents the probability from 0 to 1 with step of 0.1: darken area means a greater probability

Discussion

Rome is currently the city with the highest number of transportation vehicles on wheels in Italy. Its road topology is very irregular and very far from a grid topology.

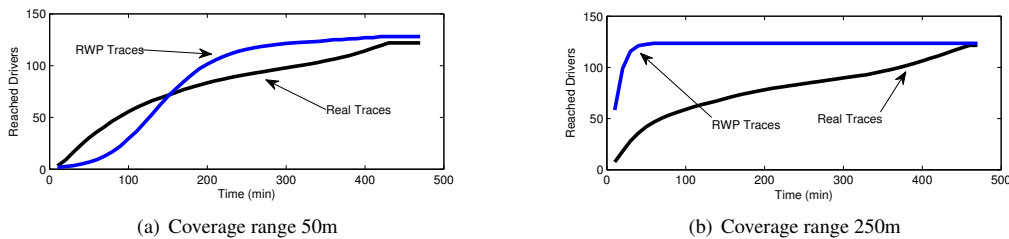


Fig. 5. Number of infected drivers as a function of the time.

We point out that there are some streets dedicated only to public transportation (bus and taxi), therefore the presented results can not be applied to general user mobility cases. Moreover taxi drivers behave differently from other citizens because their movements depend on the customer demand.

In general this demand varies during the week and across different months (for example it is higher during the summer). Moreover there could be sensible variations depending on holidays, weather conditions, special events etc. For this reason we limit the analysis in section 4 to just one month (that we recognize to be the most stable), from February 1st 2014 to March 2th 2014.

Finally, it is worth nothing that there are some municipality regulations that impose to give priority in rides assignment to the drivers waiting in specific taxi parking spots. This leads to a periodic behaviour in which the drivers start from a parking lot, serve the request and then go back to the parking lot.

4. Vanet applications

In this section we evaluate the application of a simple epidemic routing protocol on the collected real traces. The protocol works as follows: an initial message spawns on an uniformly distributed random car at a given time. That car is called “infected”. Then the message is passed to every not infected car whose position is within a given range (*coverage range*) from the infected one. Once a car receives the message, it becomes infected and participates in the data dissemination.

It is evident that after a certain period every car will receive the message hence the steady state of the process is not interesting. For this reason we present the transitory part of the diffusion process. We remark that this data dissemination is the fastest one available in the Delay Tolerant Networking field. Even if this routing presents several disadvantages such as the memory occupancy and the huge number of data circulating in the network, it can serve as an upper bound for any other routing scheme.

For all the presented results we analyse the movements of 370 taxi for a 30 days period. For each day we conduct four different simulations in which we vary the origin of the messages. Every simulation starts at a random time during the interval 0 am to 6 pm. We plot the average results together with the ones obtained through cars that move according to a simple random waypoint algorithm in which we set the speed to a constant value equals to 12.29 km/h and the wait time to 600s. These values are respectively the values of the average speed and the average wait time calculated on the real traces.

Figure 5 shows the number of reached (infected) vehicles as a function of the time comparing the real traces with the one produced by the random waypoint and varying the coverage range from 50m (figure 5 a)) to 250m (figure 5 b)). In figure 5 a) the infection process on the real traces presents an higher speed in the initial part. The reason of this behaviour can be accounted to the distribution of the drivers in the area, that is more concentrated in the center for the real case with respect to the random waypoint case, as we can see by comparing figures 3 c) with 4. If we increase the range from 50m to 250m, the RWP presents a more rapid diffusion due to the unrealistic assumption that the paths are independent. Despite several literature works that present a 250m or more coverage range, we point out that, for 2.4 and 5 GHz communications, in the proposed urban scenario, 50m seems more realistic because of the presence of several obstacles and a very high level of interferences. In the real case the coverage radius does not seem to be relevant on the speed diffusion process. We ascribed this behaviour to the habits of the drivers, that usually wait for rides in dedicated parking areas. This finds them very close for large portions of time during the day, making the dissemination more probable.

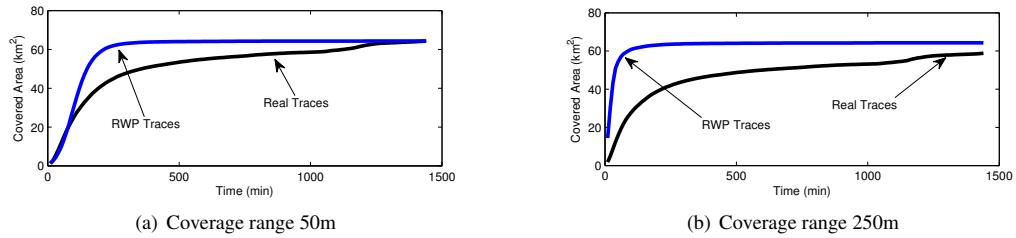


Fig. 6. The covered area as a function of the time.

Figure 6 shows the covered area defined as the sum of the areas of the cells where there is at least one infected driver before a given time. The covered area saturate to the maximum area, 64km^2 . For the considered scenario, while the RWP exhibits a relevant boost if we increase the coverage range, in the real case the difference is negligible (as already pointed out in Figure 5).

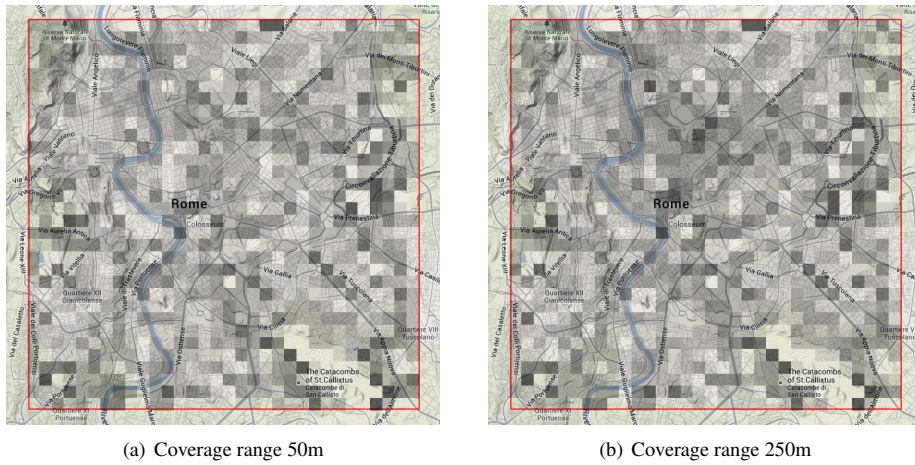


Fig. 7. Average message propagation speed for real traces

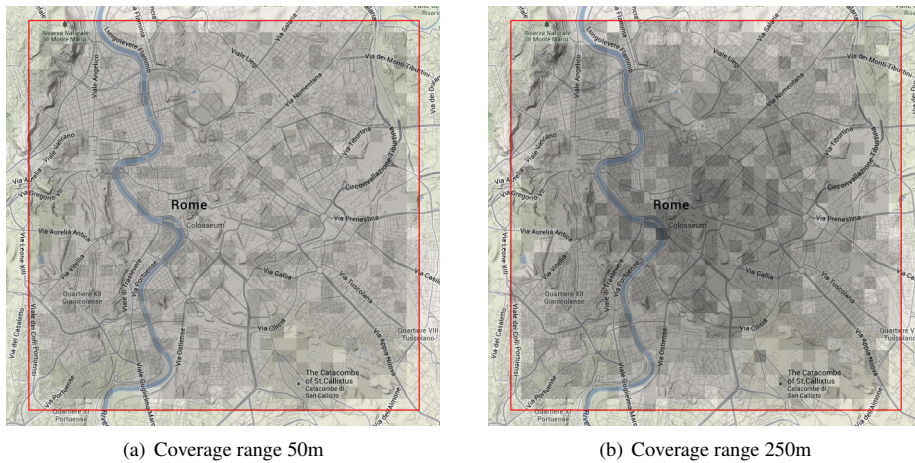


Fig. 8. Average message propagation speed for the random waypoint case

In figures 7 and 8 we show the average message propagation speed defined as the time elapsed from the instant when a message enters for the first time in a cell and when it will be passed to the first neighbouring uninfected cell.

This measure is important to understand the gradient of the message propagation speed. As shown in figure 7, in the real case the speed values are scattered in the whole area and they are not significantly affected by the increase of the communication radius. In the random waypoint case represented in figure 8, the speeds are more uniform and, as expected, they present greater values in the center, and this is exacerbated if we vary the communication range from 50m to 250m.

5. Conclusions

In this paper we presented the results of an extensive measurement campaign in which we collected the position of 370 taxi cabs for 6 months working in the center of Rome, a city characterized by high congestion streets and a road topology that is far from the grid one. We exploited the real traces to show some performance metrics related to the epidemic diffusion of a message that can be considered as an upper bound for message propagation time in delay tolerant and vehicular area networks. All the results are compared with the random waypoint case to provide a baseline. Given the characteristic of the acquired traces, all the results are deeply influenced by the major concentration of drivers in the city center, even though they also cover all the peripheral areas, but with a minor density. The presented results and traces can be used for supporting and validating emerging services in the smart city field or to assess VANET protocol performances in the presented scenario that, to the best of our knowledge, has not been reported yet in literature.

References

1. Campolo, C., Molinaro, A.. On vehicle-to-roadside communications in 802.11p/wave vanets. In: *Wireless Communications and Networking Conference (WCNC), 2011 IEEE*. 2011, p. 1010–1015. doi:10.1109/WCNC.2011.5779273.
2. Institute, E.T.S.. Etsi es 202 663 v1.1.0. *Transportation Research Part C: Emerging Technologies* 2009-11;.
3. CAR2CAR. . Car-2-car communication consortium 2013;URL: <http://www.car-to-car.org/>.
4. Li, F., Wang, Y.. Routing in vehicular ad hoc networks: A survey. *Vehicular Technology Magazine, IEEE* 2007;2(2):12–22.
5. Altayeb, M., Mahgoub, I.. A survey of vehicular ad hoc networks routing protocols. *International Journal of Innovation and Applied Studies* 2013;3(3):829–846.
6. Spaho, E., Barolli, L., Mino, G., Xhafa, F., Kolici, V.. Vanet simulators: A survey on mobility and routing protocols. In: *Broadband and Wireless Computing, Communication and Applications (BWCCA), 2011 International Conference on*. 2011, p. 1–10. doi:10.1109/BWCCA.2011.11.
7. Kotz, D., Henderson, T., Abyzov, I.. CRAWDAD data set dartmouth/campus (v. 2004-12-18) 2004;.
8. Bai, F., Sadagopan, N., Helmy, A.. The important framework for analyzing the impact of mobility on performance of routing protocols for adhoc networks. *AdHoc Networks Journal* 2003;1:383–403.
9. Härrri, J., Fiore, M., Filali, F., Bonnet, C.. Vanetmobisim 1999;URL: <http://vanet.eurecom.fr/>.
10. Härrri, J., Filali, F., Bonnet, C., Fiore, M.. Vanetmobisim: Generating realistic mobility patterns for vanets. In: *Proceedings of the 3rd International Workshop on Vehicular Ad Hoc Networks; VANET '06*. New York, NY, USA: ACM. ISBN 1-59593-540-1; 2006, p. 96–97. URL: <http://doi.acm.org/10.1145/1161064.1161084>. doi:10.1145/1161064.1161084.
11. Smith, L., Beckman, R., Anson, D., Nagel, K., Williams, M.. TRANSIMS: Transportation analysis and simulation system. In: *Conference: 5. National transportation planning methods applications conference, Seattle, WA (United States), 17-21 Apr 1995*. 1995, .
12. Krajzewicz, D., Erdmann, J., Behrisch, M., Bieker, L.. Recent development and applications of SUMO - Simulation of Urban Mobility. *International Journal On Advances in Systems and Measurements* 2012;5(3&4):128–138. URL: <http://elib.dlr.de/80483/>.
13. Kraaiier, J., Killat, U.. -the random waypoint city model-: user distribution in a street-based mobility model for wireless network simulations. In: *Proceedings of the 3rd ACM international workshop on Wireless mobile applications and services on WLAN hotspots*. ACM; 2005, p. 100–103.
14. Choffnes, D.R., Bustamante, F.E.. An integrated mobility and traffic model for vehicular wireless networks. In: *Proceedings of the 2nd ACM international workshop on Vehicular ad hoc networks*. ACM; 2005, p. 69–78.
15. Mahajan, A., Potnis, N., Gopalan, K., Wang, A.I.A.. Urban mobility models for vanets. In: *IN PROC. OF 2ND WORKSHOP ON NEXT GENERATION WIRELESS NETWORKS*. 2006, .
16. Mahajan, A., Potnis, N., Gopalan, K., Wang, A.. Evaluation of mobility models for vehicular ad-hoc network simulations. In: *IEEE International Workshop on Next Generation Wireless Networks (WoNGen)*. 2006, .
17. Piorkowski, M., Sarafijanovic-Djukic, N., Grossglauser, M.. A parsimonious model of mobile partitioned networks with clustering. In: *Communication Systems and Networks and Workshops, 2009. COMSNETS 2009. First International*. 2009, p. 1–10. doi:10.1109/COMSNETS.2009.4808865.
18. Liu, B., Khorashadi, B., Ghosal, D., Chuah, C.N., Zhang, H.M.. Analysis of the information storage capability of vanet for highway and city traffic. *Transportation Research Part C: Emerging Technologies* 2012;23:68–84.
19. Chen, Y., Xu, M., Gu, Y., Li, P., Cheng, X.. Understanding topology evolving of {VANETs} from taxi traces. *Advanced Science and Technology Letters* 2013;42(Mobile and Wireless):13–17.