Reliable gossiping in inter-vehicle communication

MIKLÓS MÁTÉ, ROLLAND VIDA

Budapest University of Technology and Economics,
Department of Telecommunications and Media Informatics, High-Speed Networks Laboratory
{mate, vida}@tmit.bme.hu

Keywords: inter-vehicle communication, gossiping, urban environment, reliability

Due to the increasing traffic density in urban areas, a computer-aided robust collision avoidance and traffic control system should be established, based on decentralized inter-vehicle communication. Vehicles group themselves into a special ad hoc network with high mobility and low link reliability. Traditional ad hoc routing solutions cannot cope with these conditions, while flooding based approaches consume too many resources. Our proposed scheme, Localized Urban Dissemination (LUD), is a location aided gossiping protocol, which concentrates the information spreading to areas where it is most likely to be useful. The reliability of simple gossiping, however, is not enough for emergency message dissemination, but our simulations prove that a simple modification in the packet forwarding scheme can overcome this limitation.

1. Introduction

The primary goal of Intelligent Transportation Systems (ITS) is to increase road safety by detecting emergency situations in advance and notifying the drivers about the traffic events. Such systems can be efficient only if the vehicles communicate with each other and share the measurements of their sensors to take coordinated actions. Inter-vehicle communication might be realized either in an infrastructure-based manner, in a pure ad hoc fashion, or as a mixture of these methods. In this paper we examine how emergency messages might be disseminated in a Vehicular Ad Hoc Network (VANET).

A VANET is a special kind of ad hoc network, as the vehicles have much higher average speed than the nodes of a sensor network, but their mobility pattern is restrained by the road network. In a VANET messages related to road safety and cooperative traffic jam avoidance should be distributed by a flooding-based protocol, as they are not addressed to a single destination, but to all cars that are interested in receiving them. These are typically the ones that need to change their speed or path in order to decrease the jam or to avoid the danger. It is an important design goal to determine where such vehicles can be found, because the broadcast in the ad hoc network is expensive, so the flood should be localized to the area of interest.

The vehicles that must be informed are in a certain vicinity of the source of the message; all the vehicles that receive the warning will take counteractions, and after a certain distance the message becomes irrelevant. We can safely assume that all vehicles are equipped with GPS receivers; therefore, a spatial flood limitation is a viable solution.

In the followings we present our Localized Urban Dissemination (LUD) protocol, which limits the message flood into areas where vehicles are interested in the message with high probability [1]. Unlike most of the similar protocols, the target area in LUD is not determined by the source, but certain forwarding nodes decide if the message is worth forwarding in a certain direction or not. This distributed decision scheme makes our solution radically different from the restricted flooding protocols based on a predefined hop count.

The rest of the article is organized as follows. Section 2 gives an introduction of the gossiping scheme, and Section 3 explains how it can be used to disseminate emergency messages in urban environments. Section 4 gives a detailed description of the operation of the LUD protocol and its properties, and the way its reliability can be increased. Finally, Section 5 summarizes the results, draws the conclusions and shows our future plans.

2. Gossiping

Historically, gossiping was first introduced in distributed databases to reduce the cost of synchronization [2]. It works as follows. All nodes know the list of the nodes in the system, and in each timeframe they randomly choose a subset to synchronize with. This reduces the number of messages exchanged in a timeframe, and as a consequence the convergence time may also decrease due to the significantly shorter periods and the marginally slower information propagation if the peers are selected carefully [3].

Gossiping in a wireless multi-hop ad hoc network (MANET) requires special peer selection strategies, as these networks exhibit properties that are different from the ones based on a fixed infrastructure. In the latter it is usually safe to assume that the cost to reach every peer is the same, and it is easy to set up a point-to-point connection between any pair of nodes (for example a TCP connection). In a MANET, however, the cost of reaching a peer dramatically increases with distance, and the wireless medium is inherently broadcast based. More-over, handling link unreliability and energy constraints are important only in a MANET. One of the possible ways of gossiping in MANETs is to flood the messages and drop them randomly with a predefined $p_{drop}>0$ probability. This random peer selection scheme favors the nodes that are close to the source and also utilizes the broadcast nature of the wireless medium.

In inter-vehicle communication the nodes are usually placed along a line (the road), and we observed that in this topology the gossiping scheme effectively limits the distance a message can reach. Namely, if the rebroadcast probability is p, then the expected value of the hop count is $1/(1-p)$, which is not infinite if $p<1$. This is exactly what is expected to be needed in emergency message propagation. The question is: how to set p to get the optimal target area?

3. Characteristics of an Urban Environment

In an urban environment the road topology is not just a lonely road, but a dense network of streets and junctions. Yet, this environment is similar to a lonely highway in the sense that the message flood follows the roads, as the buildings block the propagation of radio signals. The big difference is that from point A to point B there can be several different paths; thus, emergency messages do not need to reach a certain point, like the highway exit, because vehicles a few blocks away can already change their route in case of an accident.

It is true that there are multiple paths between the two points, but of course not all of them are taken with the same probability by the vehicles. This is due to the fact that vehicles arriving at a junction prefer certain outgoing directions over others. Some roads are one-way, some lead to important places, and turning left is usually forbidden in large intersections. When disseminating emergency messages these conditions must be taken into account, because the radio resources are scarce, and broadcasting has huge overhead [4]. To make a message dissemination protocol efficient, the traffic conditions must be considered when defining the coverage area.

4. Localized Urban Dissemination

Our proposed emergency message dissemination protocol, called Localized Urban Dissemination (LUD), is a gossiping-based emergency message dissemination protocol [1].

Gossiping makes it very easy to make a certain road segment included in, or excluded from the coverage area. The p rebroadcast probability should be changed in the junctions depending on the traffic conditions. Thus, the vehicles need to be equipped with a digital map, and set their role to *Decider* or *Forwarder*, according to their current position. The ones arriving at a junction become Deciders, and must recalculate the rebroadcast probability of the packets they forward. The nodes that are not in a junction are Forwarders; they simply rebroadcast the messages with the probability written in the packet header.

4.1. The Decision Scheme

The heart of any gossiping-based protocol would be the decision scheme that sets the rebroadcast probabilities of the packets. The decision scheme of LUD sets the p rebroadcast probability so that the probability of a message forwarded on a given path reaching a certain point P becomes equal to the probability of vehicles from that point going to the source of the message on the same path. If we call the first event A and the other one B, then the formula will look as follows:

$$\mathbf{P}(A) = C\mathbf{P}(B)\,, \tag{1}$$

where a C scaling factor is inserted to let the source scale the size of the coverage area. We will see that the resulting scheme is memoryless, and this scaling factor disappears after the first junction.

Forwarding the message along a path is a geometric process, because it is a series of independent Bernoulli trials. Its parameter is the probability of the success on the elementary trials, which is the p rebroadcast probability. The probability of the before mentioned event A is

$$\mathbf{P}(A) = \sum_i p_i^{\,h_i}\,, \tag{2}$$

because the message reaches a certain point only if all nodes on the path have chosen to forward it. The rebroadcast probability may be different for each road segment, and the different road segments are h_i hops long. A hop can be longer than the inter-car distance if there are multiple cars in the radio range of the transmitter node. The LUD protocol requires a Medium Access Control (MAC) protocol to be used that can select the farthest receiver in the direction of the flooding to rebroadcast the message. Such protocols are CBF [5] and CFB [6] for example.

The probability of vehicles going to the source of the message can be described with two parameter sets. The first is a Q_i matrix of steering probabilities for each junction; an element $q^i_{j,k}$ for junction i represents the probability that cars coming from the neighboring junction j go to neighboring junction k. The second dataset consists of s_i stop probabilities for each road segment to model finite journeys. A car reaches the source of the message along the path of the message only if it chooses the appropriate roads and it does not stop in between.

Turning this into an equation we get:

$$\mathbf{P}(B) = \sum_i q^i_{j,k}(1-s_i)\,, \tag{3}$$

if we assign the identifiers to the junctions on the path as shown in *Fig 1*.

The Decider being in junction D decides how likely it is that vehicles coming from junction D+1 are interested in the message, because the Decider itself came from junction D+1. The equation it must solve is

$$\sum_{i=0}^{D} p_i^{\,h_i} = C \sum_{i=D}^{1} q_{j,k}^{i} (1 - s_i)\,, \qquad (4)$$

where the quantity in question is q_D .The reversed indexing on the right side refers to the order the vehicles going to the source encounter the junctions.

Equation (4) has a very interesting property, because it describes a geometric process. It is known, that

$\mathbf{P}(X > x + y \mid X > x) = \mathbf{P}(X > y)$

if X follows a geometric distribution. In our case the Decider calculates the probability of the message reaching the next junction *if* it reached the current one. This conditional probability causes the dissemination process to forget the past, and all that remains of equation (4) is

$$p_D^{\,h_D} = q_{D+1,D-1}^{D} (1 - s_D)\,. \qquad (5)$$

To calculate p_D, the Decider must know the q_D turning probability for the junction and the s_D stop probability on that road segment. These might be derived from the ranks of the roads and maybe some other data the digital map can provide (e.g., stopping on a main road is highly improbable, and so is turning into a side street). Turning lanes and one-way roads are also indicated by the map. This method, however, is far from being perfect, because there are lots of things that influence the paths of the vehicles, and most of them are not present on the maps or their effect is not easy to determine.

The precision is very important when setting p, as the error caused by the insufficient knowledge of the traffic conditions can be severe. The expected hop count is $1/1-p$, which means small changes in p might trigger great leaps in the size of the coverage area.

If a traffic monitoring system collects the necessary data, a Traffic Conditions DataBase (TCDB) can be built that contains the q and s values that describe the usual traffic conditions. We expect that the navigation system that uses our dissemination protocol can eliminate most of the traffic jams, hence the usual conditions will change slowly, and the causes of the sudden deviations from the usual conditions are handled efficiently. The TCDB should be downloaded and regularly updated by the navigation device of all vehicles. The updates must not be sent on the same channel as the emergency messages, but some other means of wireless Internet access, like Wi-Fi hotspots in the parking lots or near the traffic lights. A vehicle with an outdated TCDB should not take on the role of a Decider.

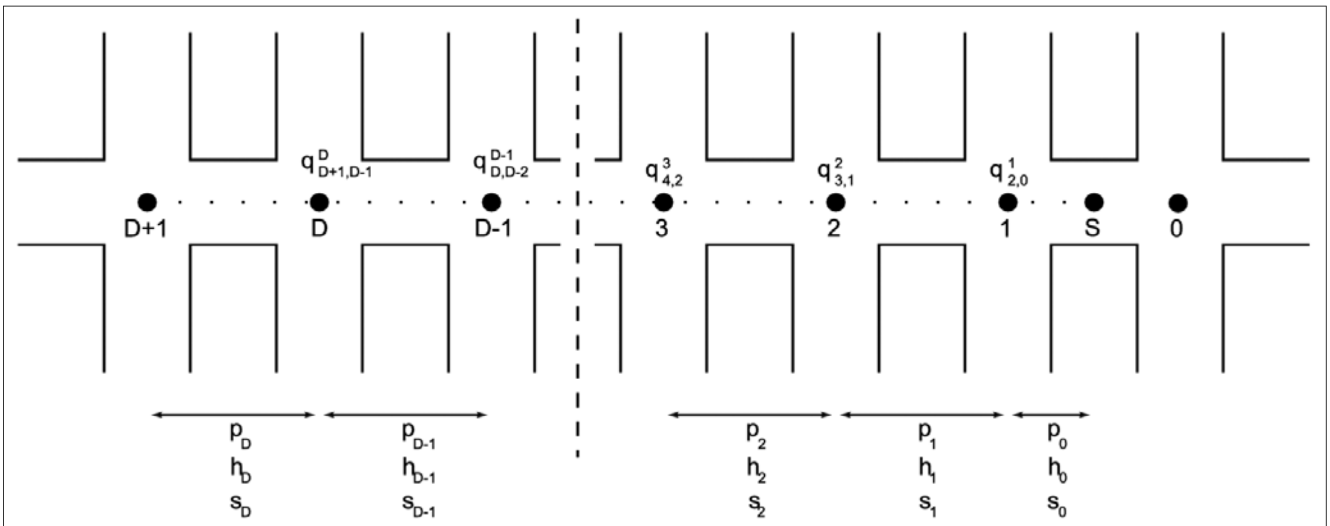
Equation (5) also contains h_D, the length of the next road segment the message might be forwarded on. The length in meters can be read from the digital map, and the Decider came to the junction from that road, so it should have some signal strength and next hop distance measurements. The LUD protocol, as mentioned earlier, needs a MAC protocol that selects the farthest node in the given direction, which implies that the MAC layer is capable of providing the necessary data.

4.2. Forwarding

The decision scheme and the gathering of its input data are the most important parts of the LUD protocol, but there are also some less obvious, but very interesting details of the protocol. The characteristics of the radio channel and the properties of the urban environment provide some challenges, but they also offer opportunities to improve the efficiency of the emergency message dissemination. The five fields the routing header of the packets should contain are: a unique identifier of the notification, the rebroadcast probability, the location of the source, the target junction (the one the message is heading to), and the previous junction (where its p was last set).

Every multi-hop routing algorithm needs to avoid routing loops. Even if the lifetime of the messages is limited, the bandwidth waste of delivering a message to nodes that already received it is not acceptable. In the case of LUD the problem is severe, as it is expected that the emergency messages are generated in large quantities. By exploiting the information sources of the protocol, the

Fig. 1. Parameters for the decisions along a path



routing loops can be avoided with a simple, stateless test: the p rebroadcast probability must be set to zero if the Decider sees that the next junction of the message (where the Decider came from) is closer to the source than the current one. The position of the source must be included in the packet header for this to work, but the message has to contain it anyway, because it carries a notification about a local change in the traffic conditions with the location of the event.

The messages are broadcasted along all possible paths. Multiple instances of them are spawned in a junction, because the vehicles arriving from all directions become Deciders and they all set the rebroadcast probabilities for the road segment they came from. Locking the dissemination of a message to the road segment the two junctions define (the position of the last decision and the supposed next junction) is beneficial for the multi-path dissemination. The coverage area becomes entirely defined by the decisions, and there is less overhead if the messages cannot "wander" around.

The source can choose in which direction the message should start to spread if the Deciders let messages into new road segments only if their target junction is the current one. The most common scenario in IVC is the backward propagation: the vehicle that detects something notifies the other vehicles that follow it. In LUD the source sets the target junction of the message to the one behind it, and the previous junction to the one ahead of it. Using the notations of Fig. 1, the message reaches junction 0, but the Deciders drop it due to the target mismatch.

The packets are distinguished by the unique identifier of the notification they carry, and the identifier of the two junctions that define their actual road segment. When two instances of the same notification go along a road segment they are indistinguishable, and they fuse into a single instance. The coverage area becomes smaller due to this, and Equation (4) is also not entirely true anymore, because some of the possible propagation paths are cancelled. However, the gain in the number of duplicated messages being eliminated makes it worth it.

4.3. Reliability of the dissemination

Emergency message propagation, like any other system, has its own reliability criteria. These criteria can be organized into two categories: the ones the MAC protocol is responsible for, and the ones the dissemination protocol must meet.

The calculation of the size and shape of the coverage area, as we presented in the previous section, considered the only reason of packet drops being the coin flip trial of the Forwarders. The wireless medium is, however, highly unreliable and spontaneous packet drops caused by interference, noise and fading are inevitable. There are numerous MAC protocol proposals for VANETs that ensure a particular level of reliability (e.g., [7]), and it should also be possible to combine them with one of the directed broadcast MAC protocols mentioned earlier. The tradeoff between reliability and propagation speed is the most important design aspect of the MAC protocol, as emergency messages need to be propagated fast, and no vehicles should remain uninformed, but this is out of the scope of this paper.

A major advantage of the memoryless nature of the decision scheme lies in its simplicity, but unfortunately it also has a serious drawback. The standard deviation (the square root of the variance) of the geometric distribution is $\sqrt{p}/(1-p)$, which almost equals its mean $(1/(1-p))$. This means that the dissemination is highly unreliable, because the message can be dropped at any time, and reaching zero or very few hops has a too high probabili-

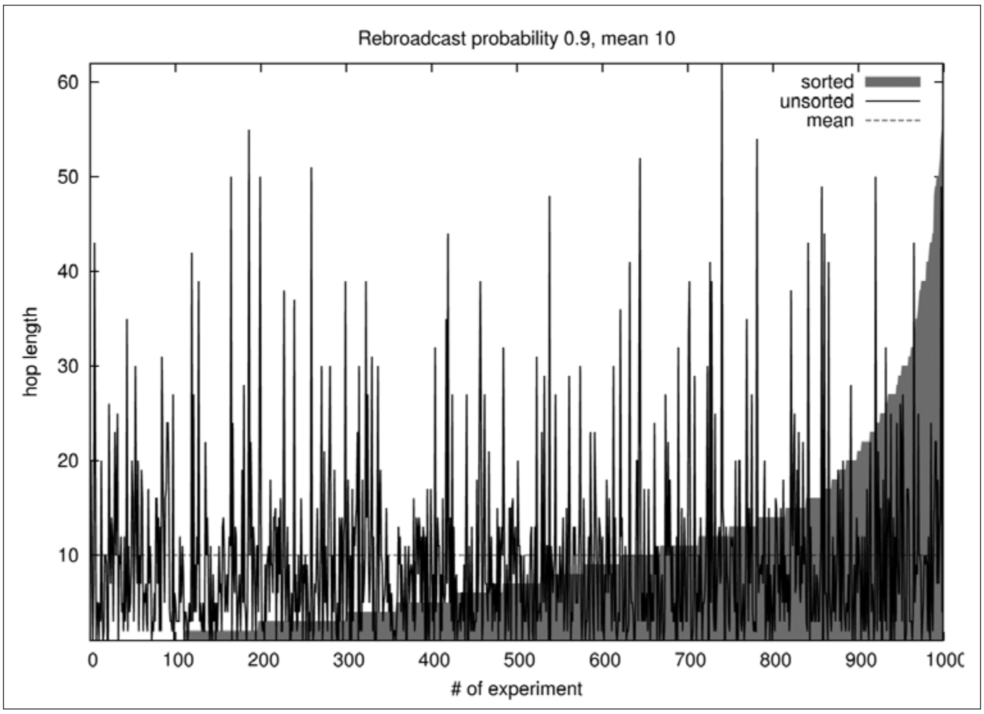


Fig. 2. Hop lengths in realizations of the geometric process

ty. The high variance of the geometric process is caused by the terminal verdict of the elementary trials.

According to simulations for a single road, shown in *Fig. 2*, there are high peaks in the hop lengths, and after sorting the results numerically the dominance of the ones that are smaller than the mean becomes clearly visible: the stairs are broad for small values, and the peak at the end is much higher than the mean. The difference between the resulting coverage area and the theoretical one should be minimized in order to improve the reliability of the dissemination.

Reducing the variance of a distribution is best done with averaging. There are two possible ways to produce an averaged dissemination area. One is to send more notifications of the same event – this is easy, as events are usually detected by more than one vehicle, and the only thing to do is not to suppress additional notifications. This solution, however, increases the number of messages flowing in the network, wasting the precious resources.

A better way to decrease the variance is to modify the elementary coin flipping trial of the Forwarders to make the resulting distribution the average of *k* independent runs. The modified trial is a voting game: the message only gets dropped if *k* nodes voted for dropping. This scheme results in a smoother distribution, as the verdict of an elementary trial is not a terminal one. However, it is not memoryless anymore, because the counter of the dropping votes must be included in the packet header. To restore the expected value of the hop count, the rebroadcast probability must be decreased to:

$$p' = 1 + k(p-1)\,. \qquad (6)$$

Fig. 3 shows the simulation results for the band of standard deviation around the mean for the geometric process and two averaged processes. The effect of the averaging, i.e., the decreased standard deviation means an increase in the reliability of the dissemination. It is also visible that using this modified scheme messages go at least *k* hops, but it has no harmful consequences.

5. Summary and conclusions

The Localized Urban Dissemination protocol provides a limited flooding by using a gossiping scheme that randomly drops packets with a given probability. In urban environments the buildings block the propagation of radio signals; hence the area that must be covered by the dissemination can be determined by using a digital map. Using information about the usual traffic conditions, the vehicles being in the junctions can decide if it is worth forwarding the message in the next road segment or not. The shape of the coverage area evolves dynamically as a result of this chain of decisions.

In theory, the LUD protocol is well suited to emergency message dissemination, as the distributed computing eliminates the long delay the source would need to spend on calculating the shape of the coverage area. It also uses locally available knowledge, like the actual traffic density, which improves the quality of the coverage area. Here, quality means that nodes that actually need the information should receive it, but the precious bandwidth of the wireless channel should not be wasted with superfluous packets.

Simulations have shown that gossiping limits indeed the message flood into the designated area. The disadvantage of the random packet drop is the high variance in the hop number the messages reach; however, a simple change in the behavior of the Forwarders, and a stored state in the packet headers can increase the reliability of the protocol to an acceptable level.

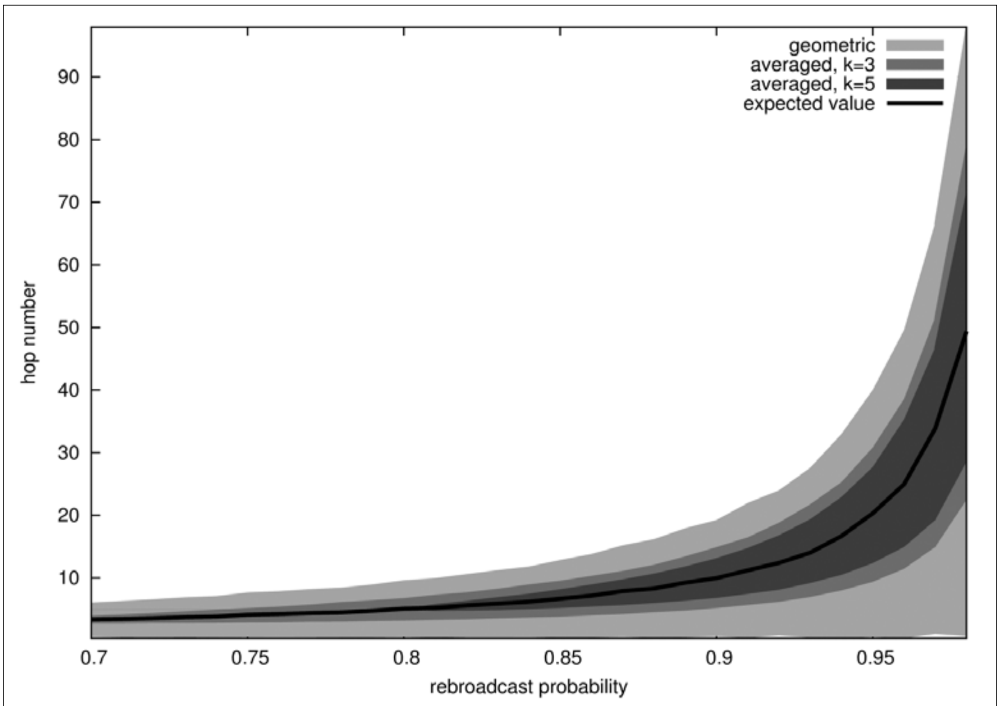


Fig. 3. Comparison of the standard deviation of the geometric and the averaged disseminations

In the future the continued theoretical inspection might reveal other interesting properties of the LUD protocol, and we expect that its efficiency can be further increased with additional small modifications to one of its algorithms. Packet level simulations will also be needed to analyze the effects of the packet collisions and node mobility on the multi-hop dissemination.

Authors

MIKLÓS MÁTÉ is a PhD student at the Department of Telecommunication and Media Informatics of the Budapest University of Technology and Economics, where he obtained his MSc degree in 2007. His research interests include scalable routing protocols and efficient information propagation in ad hoc networks.

ROLLAND VIDA is Associate Professor at the Budapest University of Technology and Economics. He obtained his BSc and MSc degrees in Computer Science from the Babes Bolyai University, Cluj-Napoca, Romania, in 1996 and 1997 respectively, and his PhD degree from the Université Pierre et Marie Curie, Paris, France, in 2002. Between 2003 and 2005 he obtained the György Békésy Postdoctoral Fellowship, and in 2007 the János Bolyai Research Fellowship. In the last 5 years Dr. Vida has acted as organizer, TPC member or reviewer for more than 30 international conferences, participated in several national and European research project, and taught different networking courses at universities in Romania, Slovakia and Hungary. In 2008 he was elected as Chair of International Affairs of the Scientific Association for Infocommunications, Hungary.

References

[1] M. Máté, R. Vida, "Probability-based information dissemination in urban environments", In Proc. of Eunice 2008, Brest, France, September 2008.

[2] A-M. Kermarrec, M. van Steen, "Gossiping in distributed systems", ACM SIGOPS Operating Systems Review, Vol. 41, Issue 5, 2007.

[3] S. Boyd, A. Ghosh, B. Prabhakar, D. Shah, "Randomized gossip algorithms", IEEE Transactions on Information Theory, Vol. 52, No. 6, pp.2508–2530. June 2006.

[4] S-Y. Ni, Y-C. Tseng, Y-S. Chen, J-P. Sheu, "The broadcast storm problem in a mobile ad hoc network", In Proc. of the 5th annual ACM/IEEE International Conference on Mobile Computing and Networking, Seattle, Washington, United States, August 1999.

[5] D. Chen, J. Deng, P.K. Varshney, "On the forwarding area of contention-based geographic forwarding for ad hoc and sensor networks" In Proc. of SECON'05, September 2005.

[6] H. Füßler, H. Hartenstein, J. Widmer, M. Mauve, W. Effelsberg, "Contention-Based Forwarding for Street Scenarios", In Proc. of WIT 2004, Hamburg, Germany, March 2004.

[7] G. Korkmaz, E. Ekici, F. Özgüner, "An Efficient Fully Ad-Hoc Multi-Hop Broadcast Protocol for Inter-Vehicular Communication Systems", In Proc. of IEEE ICC'06, Istanbul, Turkey, June 2006.