Adaptive Retransmission Policy for Reliable Warning Diffusion in Vehicular Networks

Francesco Giudici, Elena Pagani, and Gian Paolo Rossi
Information and Communication Department, Università degli Studi di Milano, Italy
E-mail: {fgiudici,pagani,rossi}@dico.unimi.it

***Abstract*— Use of wireless technologies is becoming pervasive in everyday life. Recently, research began analyzing their use on board of vehicles for several kinds of applications, ranging from traffic safety to fleet management and cooperative work, to entertainment and Internet browsing. In this work, we focus on safety applications and in particular on the approach proposed in the framework of the PATH project [1] for reliable diffusion of warnings to advertise problems in vehicular traffic. The approach in [1] is based on *static* parameters describing the environment. Unfortunately, in real environments those parameters may dynamically change over time. In this work we present performance measurements obtained by varying the parameters, to evaluate how the performance of the approach depends on environmental conditions.**

I. Introduction

Vehicles equipped with wireless network interface cards (WNICs) start to be available. This equipment can be used for several applications, ranging from fleet management and cooperative workgroup to entertainment and Internet navigation. In this work we focus on the problem of *vehicular safety*. Wireless networking can be exploited to provide communication among vehicles, in order to notify the occurrence of problems – e.g., accidents, icy street, obstacles on the road – to other oncoming vehicles. *Warnings* are addressed to all vehicles approaching the place where the problem occurred, so as to allow drivers to perform the appropriate actions. Warning traffic has service requirements that must be guaranteed by the system. *Low latency* is needed to guarantee that the warning can be detected by a driver so that s/he has sufficient time to properly react to the event notified. *High reliability* is needed to guarantee that all interested vehicles actually receive a warning. On the other hand, wireless links have some unfavorable characteristics, such as long latency for channel set-up and low reliability. The former is due to the time needed to two devices to synchronize and agree about the policy for channel access (such as the used code or frequency hopping pattern). Some solutions have already been proposed to exploit wireless technologies to supply vehicular safety, some of which supported by car producers and government institutions. In this work, we focus on the *static* approach proposed by the PATH project [1] for reliable diffusion of warnings, with the aim of both understanding how environmental characteristics impact on the obtainable reliability, and devising mechanisms to dynamically adapt the approach in order to optimize performance according to changes in those characteristics. The California PATH Project involves among other things, researches on an infrastructure to boost vehicular safety. The proposed solution aims at achieving reliable dissemination of warnings through *repetitions*, i.e., multiple retransmissions of a warning so as to overcome channel failures and collisions with other messages. PATH seems the most promising approach, and it is under study for adoption on U.S. highways. However, it assumes that an optimal number of repetitions exists for a certain scenario. Unfortunately, the vehicular environment is highly dynamic. The density of vehicles in a certain area, the number of concurrent warning sources and the vehicle speed vary over time, and so should do the number of repetitions. In this work, we analyze by simulations how the number of repetitions needed to achieve reliability varies depending on the characteristics of both the vehicular and the data traffic; an alternative adaptive policy is discussed. A Vehicular Collision Warning Communication (VCWC) [3] has been proposed, focusing primarily on achieving a low transmission latency. VCWC does not take reliability aspects into considerations. Both the above proposals rely on the DSRC (Dedicated Short Range Communications) multi-channel architecture [4]. DSRC has been explicitly designed for use in vehicular systems. DSRC proposes communication services for both private applications and public safety, with the possibility of using high power transmission when latency is important. DSRC is now in the process of standardization by IEEE as the WAVE (Wireless Access in Vehicular Environments) project; it is also known as the ISO CALM (Communications Air Interface Long and Medium range) standard [5]. The European Project CarTALK/Fleetnet [6], [2] uses UTRA-TDD (UMTS Terrestrial Radio Access with Time Division Duplexing) as the channel architecture, thus adopting a frequency range requiring licensing.

II. System Model

In this work we consider a system composed by vehicles equipped with wireless network interface cards (WNICs). Vehicles have a GPS system. We focus on vehicle-to-vehicle communication; no roadside communication infrastructure is needed. A vehicle can notify road hazards to all oncoming vehicles; communication is broadcast and addressed to one-hop neighbors only. Warnings may contain information about the zone affected by the notified problem. The wireless technology is based on the 802.11 standard [7]. In particular, the channel structure is determined by the DSRC proposal [4]. Vehicles are equipped with *On-Board Units* (OBUs)

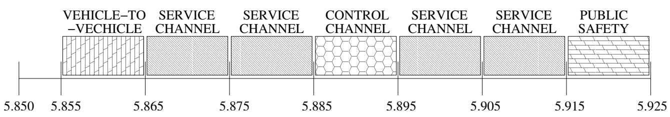


Fig. 1. Layout of the DSRC channel

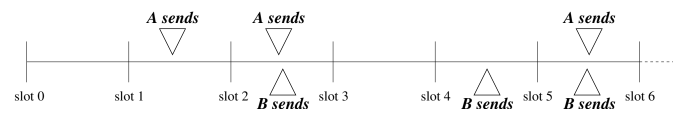


Fig. 2. Example of PATH protocol execution

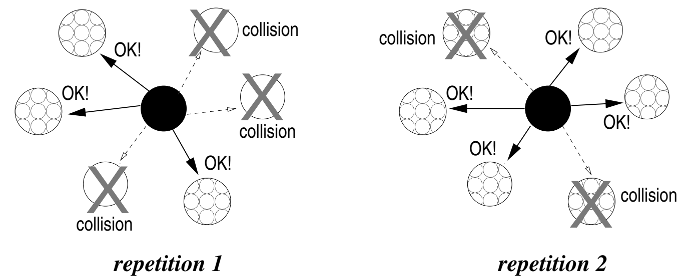


Fig. 3. Example of successful execution of PATH protocol

that support communications among vehicles. DSRC uses the transmission range 5.850 to 5.925 GHz. The transmission range is divided into 7 channels of 10 MHz each (fig.1); the data rate supported is up to 27 Mbps. MAC and physical layers are provided by the IEEE 802.11p proposal [8], [9]. DSRC has an average communication range of around 300 mt., and up to 1000 mt. Channel access is performed through CSMA. Channels have different aims: four of them are *Service* channels that can be used mainly for common data and private applications, but also for public safety. All Service channels are accessed in a shared way by all vehicles. The *Control* channel is mainly used to exchange control information needed to synchronize vehicles for access to the other channels and to announce the correspondence among applications and Service channels. It is also used to exchange high priority messages for vehicular safety. Time to access the Control channel must not be greater than 100 msec.; the channel must not fail in case of congestion. A device must listen to the Control channel for intervals of at least 200 msec., and it cannot be off the Control channel for more than 50 msec. A *vehicle-to-vehicle communication* channel exists, devoted for instance to publish information about the mobility pattern of a vehicle, in order to forecast the possibility of accidents and forewarn drivers. The last channel is dedicated to the exchange of warnings for *public safety*. All channels are used for several types of traffic and can be accessed by multiple vehicles concurrently; hence, collisions are possible. In this work we assume that all warnings notifying a problem in vehicular traffic are sent on the Control channel [10]. We analyze the problems involved with performing retransmissions on that channel in order to guarantee high reliability while at the same time avoiding channel congestion.

III. CALIFORNIA PATH PROJECT

In the framework of the California PATH project, six protocols have been proposed [11] to diffuse warnings for vehicular safety within a bounded time, while guaranteeing high reliability. Warning messages must be reliably received by all source's neighbors with a low latency. A warning has associated a *packet lifetime* τ within which it must be reliably received by all neighbors before becoming useless because too late to allow drivers to appropriately react: it is an upper bound on the transmission latency. Due to the broadcast diffusion of warnings, acknowledgments cannot be used to control reliability, to avoid ack implosion at the sender; for the same reason, the RTS/CTS mechanism cannot be used. The protocols proposed consists in considering the packet lifetime interval as divided into n *slots* such that $n = \lfloor \tau / T_x \rfloor$ where T_x is the transmission time of a warning, depending on the packet size and the channel bandwidth. Hence, a packet can be sent in a slot. The protocols presented in [11] differentiate in three respects:

- nodes can perform carrier sensing before accessing one of the chosen slots, or not. In the former case, if the channel is busy, the repetition scheduled for that slot is dropped, thus actually decreasing the number of repetitions performed;
- nodes are synchronized on slot beginning, or slots are locally determined according to the instant a warning is generated;
- when a node has a warning to send, either it a priori randomly chooses K slots among the n and tries to send the packet in those slots, or in each of the available slots transmits a warning with a uniformly distributed probability of K/n.

A warning has been reliably delivered when each node in the communication range of the source has received it at least once. It is worth to notice that, because of broadcast transmissions, a node can receive duplicates. The protocol fails if one or more source's neighbors exist that do not receive any warning. As an example, in fig.2, two sources are sending their warnings, and they performs 3 repetitions. They send their warnings in the slots chosen a priori and they collide in slots 2 and 5, while transmissions in slot 1 for source A and in slot 4 for source B are successful. Let us notice that, because of the hidden station problem, not necessarily a transmission is successful for *all* neighbors. In each repetition a source could reach only a subset of its neighbors. In fig.3, a situation is shown in which none of two repetitions is successful, but they together reach all destinations. What matters here is that throughout the K repetitions all neighbors have been reached at least once. According to simulation results discussed in [11], best performance has been achieved with slots for repetitions randomly chosen a priori, asynchronous nodes, and nodes performing carrier sensing before sending a packet in a slot (Asynchronous Fixed Repetition with Carrier Sensing, AFR-CS protocol). In fig.4, pseudo-code for AFR-CS is provided.

```
when (a warning must be sent) do
    for (i = 1 to K) slot[i] ← randomly chosen slot;
    for (i = 1 to K)
        when (current slot = slot[i]) do
            carrier sensing;
            if (slot free) then send i-th repetition;
        od
    end for
od
```

Fig. 4. Pseudo-code of AFR-CS

In [11], an analytical evaluation has been performed according to which the optimal number of repetitions is 7, under the hypothesis that warning generation follows a Poisson distribution and for a specific scenario with communication range of 80 mt., average distance among vehicles 30 mt., 4 lanes, and 75 interferers around each receiver. It is extremely important to carefully estimate an appropriate value for the number of repetitions. Reliability must be obtained, but without risk of congesting the Control channel, which *must* remain available for its other usages. However, analytical evaluation of K does not seem appropriate: in the considered *highly dynamic* environment it is impossible to characterize an average situation so as to optimize K. The analysis bases on assumptions that are not necessarily valid in a real environment, such as poissonian generation of warnings or estimation of the number of interferers. The number of interferers is not the same for all recipients. As a consequence, this approach can be used in a real environment only by configuring parameters basing on an average situation, which could be far from the actual situation, thus leading either to low reliability – for too low K – or to both congested network and low reliability – for too high K. As an alternative, a vehicle should consider the current state of the vehicular traffic, and dynamically adapt the number of retransmissions needed to disseminate its own warnings basing on local observations about the number of neighbors and the load of warning traffic in its surroundings in the recent past.

IV. Enhancing PATH

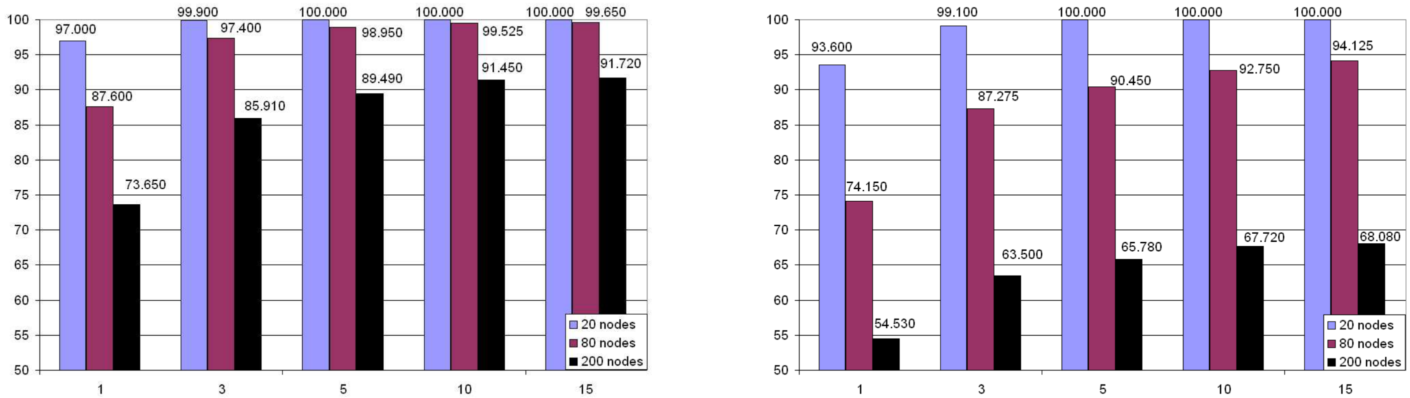
We performed simulations of PATH using the NS-2 package to highlight correlations among the vehicular and data traffic conditions and the achieved reliability. The parameters used to evaluate PATH performance are shown in Table I, and they are the same adopted in [12]. The aim of the measures is to evaluate an upper bound on the retransmissions needed to achieve reliability by stressing the system. In simulations, the message generation interval has been set equal to the packet lifetime so as to guarantee that each node is always sending a warning. The channel bandwidth has been set to 18 Mbps, in accordance with the considerations reported in [12]. In [5] a data rate of 6 Mbps is assigned to the Control channel, while the other channels have a data rate of 27 Mbps; we also performed measures with 6 Mbps rate. Devices are distributed over a small area so that are all in communication range. Packet size allows to communicate coordinates – according to a GPS system – indicating where a problem occurred. From packet size and 18 Mbps bandwidth, a slot time of 147 μsec. is obtained, slightly greater than the packet transmission time. As a consequence, the number of slots is $n = (\tau/\text{slot time}) = 681$. The communication range is in line with the DSRC characteristics. The message range, that is, the distance from the source at which the message should be propagated, equals the communication range, thus enforcing one-hop diffusion. Mobility impacts on the definition of reliability, because vehicles near the warning source can move out of communication range before receiving the packet, and vehicles can enter the source communication range within the packet lifetime. To accurately measure the reliability degree without having to deal with mobility issues, vehicles do not move in our simulations. Measures have been performed with 20, 80 and 200 nodes in the network, for variable number of repetitions.

TABLE I
PARAMETERS FOR PATH PERFORMANCE EVALUATION

Packet lifetime (τ)	100 msec.
Message generation interval	100 msec.
Packet size	250 Bytes
Control channel bandwidth	18 Mbps
Communication range	250 mt.
Message range	250 mt.
Mean distance among neighbors	$\leq$ 250 mt.
Slot time	147 μsec.

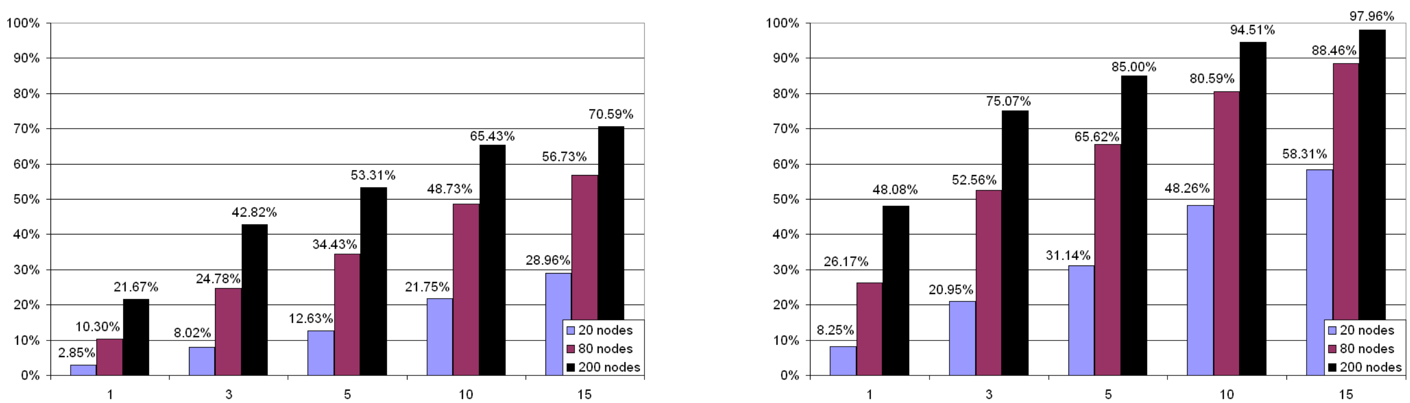
A. Performance Analysis and Optimization

For both values of control channel bandwidth, simulation conditions exist in which 100% reliability cannot be achieved (fig.5). For high number of nodes, increasing the number of repetitions is not effective when the channel tends to congest, and the achieved reliability tends to stabilize. Channel congestion (fig.6) increases almost to saturation. For larger (18 Mbps) bandwidth and 15 repetitions still 1/3 of the channel is unused although reliability has already stabilized. This is due to a greater probability that nodes choose the same slots to perform repetitions. Indeed, the probability of a slot to be chosen by a certain node to send a repetition is the ratio (number of repetitions / n), which for 15 repetitions amounts to 0.02 for 18 Mbps and 0.06 for 6 Mbps. Hence, The probability for a slot of being not used by any of 200 nodes is $0.98^{200} \simeq 0.018$ in the former case, while $0.94^{200} \simeq$ 4E-6 in the latter. The carrier sensing mechanism helps in avoiding collisions (fig.7), but it confirms channel congestion. For high number of both nodes and repetitions, often a certain slot chosen a-priori cannot be used for sending a repetition because already in use. Behavior for 6 Mbps bandwidth is similar; for 200 nodes and 15 repetitions the probability of finding a slot already in use increases up to 92.56%. On the other hand, also under critical conditions the probability of collisions is negligible (fig.7(b)). It is worth to notice that, in case nodes



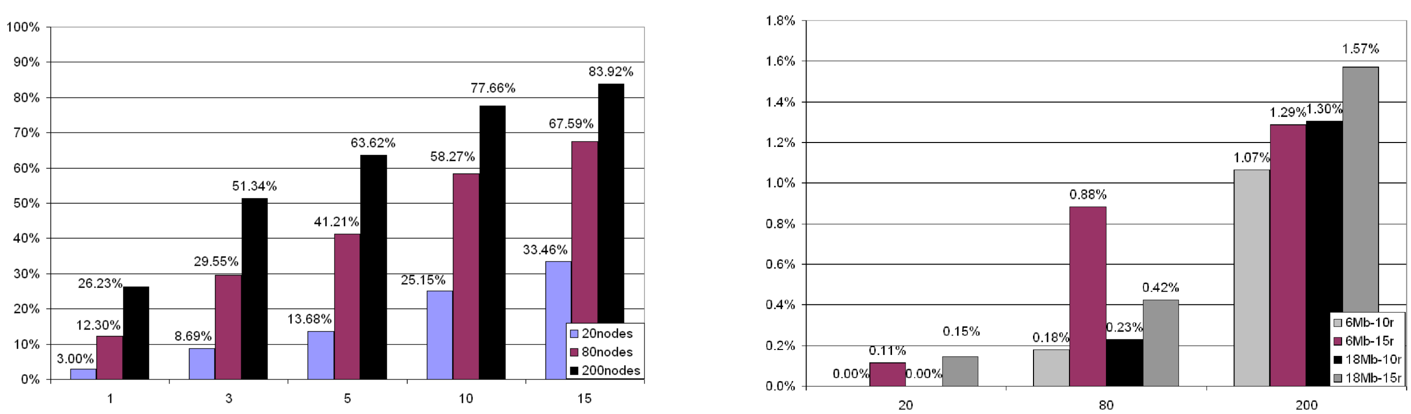
(a) (b)

Fig. 5. Percentage of warnings reliably delivered with respect to number of repetitions issued by each node for (a) 18 Mbps or (b) 6 Mbps of channel bandwidth



(a) (b)

Fig. 6. Percentage of slots used for (a) 18 Mbps or (b) 6 Mbps of channel bandwidth, with respect to number of repetitions issued by each node



(a) (b)

Fig. 7. (a) Percentage of slots found busy with 18 Mbps with respect to number of repetitions issued by each node. (b) Percentage of collided packets with respect to number of nodes

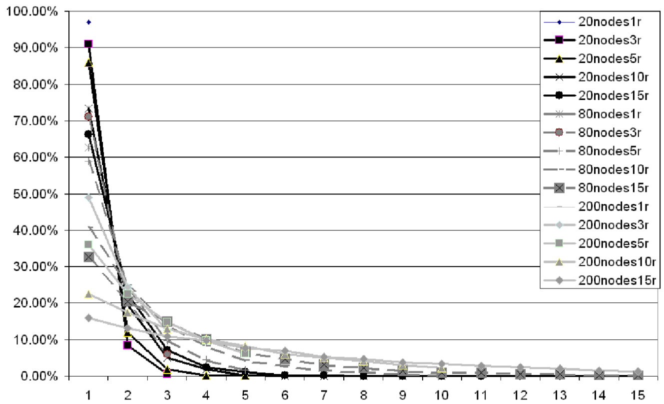
are not all in range, the *"hidden station"* phenomenon would increase collisions, which could be much more disruptive for reliability than repetitions suppressed because of busy channel. We analyzed the contribution of each repetition to the globally obtained reliability. In fig.8, the percentage of times in which reliable delivery has been achieved in the i-th repetition is reported. The percentage has been evaluated over a number of warnings equal to (50 $*$ number of nodes). For a few nodes, there is greater probability that they choose different slots to perform repetitions. Hence, all destinations are reached with a low number of retransmissions. By contrast, the greater the number of nodes, the greater the number of retries before achieving reliability. Indeed, nodes contend for using the same slots: with 18 Mbps bandwidth, 200 nodes and 15 repetitions for each warning, the number of slots needed to accommodate repetitions of all nodes is $200 \times 15 = 3000$ while only 681 slots are available in a packet lifetime. A slot could be thus chosen by 4-5 sources on average. One of them succeeds in accessing the channel, while the others omit to perform a repetition and wait for the next slot chosen. As a consequence, the average number of repetitions needed to achieve reliable delivery increases for increasing number of nodes. Moreover, increasing the number of repetitions, the probability of succeeding with the first repetition decreases; but on the other hand increases the probability of success in successive repetitions, thus yielding a better global reliability than with only a few repetitions.

Several considerations can be inferred from the results presented above. First of all, a statically determined number of repetitions (7 according to the analysis performed in [1]) is not always adequate. For instance, with low number of nodes less repetitions (3 -5) are enough to reliably diffuse warnings, without at the same time congesting the Control channel. The most appropriate decision for a node seems to be performing enough repetitions to reach a stable reliability without high congestion. Further dissemination of information about the traffic event signaled by the node could be obtained by warnings generated by other nodes detecting the same event, or by warning generated as a consequence of the original warning. On the other hand, congestion on the control channel *must* not occur to avoid making non-accessible the other channels. Indeed, the simplest solution to guarantee high reliability in any condition would be to exploit carrier sensing: a node continuously senses the channel and sends a repetition in each slot it finds unused, possibly till the desired number of repetitions has been reached. But this approach is absolutely not suitable in order to guarantee proper work of other channels. Each node should monitor the number of other nodes in its neighborhood that are generating warning, and the traffic load, and compute its number of repetitions according to those parameters. The number of repetitions should be dynamically adapted according to variations in the number of neighbors. The most promising solution, and the one we are going to evaluate with further simulations, allowing to reduce congestion and contention on slot usage, consists in having a node that refrains from performing all the repetitions initially scheduled in the event it senses the channel free for the first few (3-5) repetitions, thus letting the channel usable by other nodes. On the other hand, a node that senses the channel busy could re-schedule the suppressed repetition in one of the successive slots in order not to decrease its probability of success for inability in using the channel. Although re-scheduling could be computationally heavy. A better comprehension of the mechanisms coming into play in the described system could be obtained by devising a statistical model of the behavior of nodes and performing an analytical evaluation. Yet, many phenomena may impact on both achieved reliability and channel usage, which cannot be easily modeled analytically, nor reproduced in simulations. They are discussed in the next session.

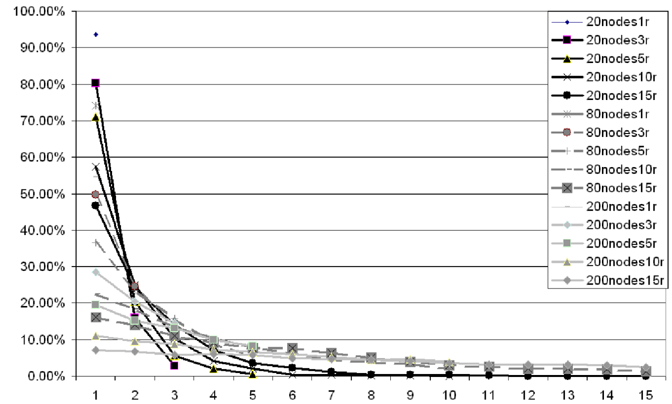
B. Behavior in Real Environments

Simulations show that the number of retransmissions needed to achieve reliability depends on the load offered to the network and on the density of devices. Because of mobility, each node may observe dynamic changes of these indexes as a consequence of its own movements and the movements of the devices in its communication range. As an explicit requirement of the DSRC architecture is that the Control channel is resilient to congestion, the number of retransmissions must be the lower bound needed to achieve reliable warning delivery. A dynamic policy – able to adapt to the current network state – is preferable to a static one both to supply reliability guarantees and to avoid congestion. As a matter of fact, in real environments many other situations occur, which are very difficult to reproduce with simulations. A warning reporting a traffic problem may be – almost simultaneously – generated by all vehicles near the position of occurrence and detecting the problem. One one hand, these *duplicate* warnings compete to use the channel, thus making more difficult guaranteeing a reliable delivery of all of them. On the other hand, as they signal the same problem, it is enough that a vehicle receives at least one of those warnings from one of the advertisers. Hence, *multi-path* propagation helps achieving reliability. It is difficult to evaluate the extent to which these competing effects impact on vehicles (drivers) behavior. A warning can trigger the generation of *cascading* warnings. In case a driver suddenly brakes, his/her vehicle V sends a warning to oncoming vehicles, let us say W and Z. Those vehicles in turn are forced to brake or slow down, generating on their behalf other warnings. This chain of events propagates the notification of a traffic problem over multiple hops. But vehicles in the communication range of V, W and Z receive different warnings concerning the same problem. This phenomenon contributes in increasing reliability. It is worth to notice that in all our simulations only warning traffic has been generated. In fact, in a real DSRC environment the Control channel is also used by other data traffic,[1] and those messages are not subject to retransmissions as they do not have reliability requirements. This has a twofold consequence: (i)

[1]E.g., announcements of services available on the other channels.



(a)



(b)

Fig. 8. Percentage of warnings that have been reliably delivered at the i-th repetition, for (a) 18 Mbps or (b) 6 Mbps of channel bandwidth

concurrency in medium access should be lower than that we reproduced, also in conditions of high vehicular density; (ii) data traffic with no reliability requirements risks to be pushed out of the network because of the aggressiveness of warning traffic. As far as the latter issue is concerned, as data traffic sent over the Control channel is needed to synchronize accesses to Service channels, if nodes cannot access the media then the whole system is disrupted. A solution could be to equip vehicles with two WNICs – according to WAVE specification. One antenna is devoted to safety applications while the other one is used for all other applications. In this case, concurrency among warnings could be accurately modeled by the presented simulations.

V. Conclusions and Future Works

In this work, an approach is analyzed for warning dissemination in vehicular networks, with the purpose of deriving indications to make it adaptive. Measurements provide several ideas about how to dynamically change node behavior according to current vehicular traffic and network conditions, and what parameters to consider for this purpose. These ideas must be validated by further simulations. Moreover, other future developments can be imagined. A warning is addressed to one-hop neighbors of the source. Depending on the vehicle speed, this could be not enough, for instance if the speed is so high that the route covered by a vehicle before arriving to the place a problem occurred – or needed to a driver to brake before arriving there – is larger than the communication range. In these cases multi-hop propagation is needed. In the discussed simulations, warnings are unrelated one to another. A more careful analysis could be performed to highlight whether correlated, cascading warnings are effective to propagate a warning over multiple hops in acceptable time. An alternative approach we are exploring is to set-up ad hoc *safety networks* dedicated to the exchange of warnings, so that a vehicle always belong to a safety network and is able to receive warnings of interest. Such an approach must cope with the delays involved in creating, joining and merging ad hoc networks, and it seems to require amendments to the 802.11 standard. Further simulations must be performed to evaluate the mutual impact of warning traffic and all other traffic. On one hand, concurrency among several traffic flows would make more difficult to provide reliability guarantees. On the other hand, it is interesting to measure how repetitions for warning messages affect normal traffic, in order to ensure a fair bandwidth usage among flows, compatibly with respective service requirements. Moreover, the effects of mobility could be analyzed for different vehicle speeds, once an appropriate reliability definition is characterized for the case of changes of the destination group.

References

[1] California PATH Project, *Partners for Advanced Transit and Highways (PATH)*. http://www.path.berkeley.edu/.

[2] W.J. Franz, H. Hartenstein, B. Bochow, *Internet on the Road via Inter-Vehicle Communications*. Proc. Workshop der Informatik 2001: "Mobile Communications over Wireless LAN – Research and Applications", Sep. 2001, http://www.et2.tu-harburg.de/fleetnet/english/documents.html.

[3] X. Yang, J. Liu, F. Zhao, N.H. Vaidya, *A Vehicle-to-Vehicle Communication Protocol for Cooperative Collision Warning*. Proc. 1st IEEE Annual Intl. Conf. on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous'04), 2004, pp. 114-123.

[4] DSRC writing group, *Dedicated Short Range Communications (DSRC) – Tutorial*. http://www.leearmstrong.com/DSRC/DSRCHomeset.htm.

[5] IEEE, *Consolidated Report on the Requirements for Public Safety Security in WAVE Systems*. Draft 0.8, IEEE, June 15 2004.

[6] CarTALK Consortium, *CarTalk 2000 Project*. http://www.cartalk2000.net.

[7] IEEE Standards Association, *IEEE 802.11 Wireless Local Area Networks*. http://grouper.ieee.org/groups/802/11/index.html.

[8] IEEE Vehicular Technology Society, *5.9 GHz Dedicated Short Range Communication (DSRC) – Overview*. http://grouper.ieee.org/groups/scc32/dsrc/index.html.

[9] B. Cash, *WAVE Background Information – Wireless Access in Vehicular Environments for the 5.9 GHz band*. doc.: IEEE 802.11- 04/ 0121r0, Jan. 2004.

[10] R. Sengupta, Q. Xu, *DSRC for Safety Systems*. Intellimotion – Research Updates in Intelligent Transportation Systems, Vol. 10, n. 4, 2004, pp.2.

[11] Q. Xu, T. Mak, J. Ko, R. Sengupta, *MAC Protocol Design for Vehicle Safety Communications in Dedicated Short Range Communications Spectrum*. Proc. IEEE ITSC 2004, http://path.Berkeley.edu/dsrc.

[12] Q. Xu, T. Mak, J. Ko, R. Sengupta, *Vehicle-to-Vehicle Safety Messaging in DSRC*. Proc. 1st ACM Workshop on Vehicular Ad Hoc Networks (VANET'04), 2004, pp. 19-28.