A Collaborative Protocol for Anonymous Reporting in Vehicular Ad Hoc Networks

Carolina Tripp Barba, Luis Urquiza Aguiar, Mónica Aguilar Igartua, Javier Parra-Arnau, David Rebollo-Monedero, Jordi Forné*, Esteve Pallarès

Department of Telematics Engineering, Universitat Politècnica de Catalunya (UPC)
C. Jordi Girona 1-3, 08034 Barcelona, Spain

Abstract

Vehicular ad hoc networks (VANETs) have emerged to leverage the power of modern communication technologies, applied to both vehicles and infrastructure. Allowing drivers to report traffic accidents and violations through the VANET may lead to substantial improvements in road safety. However, being able to do so anonymously in order to avoid personal and professional repercussions will undoubtedly translate into user acceptance. The main goal of this work is to propose a new collaborative protocol for enforcing anonymity in multi-hop VANETs, closely inspired by the well-known Crowds protocol. In a nutshell, our anonymous-reporting protocol depends on a forwarding probability that determines whether the next forwarding step in message routing is random, for better anonymity, or in accordance with the routing protocol on which our approach builds, for better quality of service (QoS). Differently from Crowds, our protocol is specifically conceived for multi-hop lossy wireless networks. Simulations for residential and downtown areas support and quantify the usefulness of our collaborative strategy for better anonymity, when users are willing to pay an eminently reasonable price in QoS.

Keywords: Anonymous reporting, anonymous-communication systems, Crowds protocol, multi-hop routing, vehicular ad hoc networks

1. Introduction

Road safety has become an important issue for governments and vehicle manufacturers in the last twenty years. Vehicular ad hoc networks (VANETs) [1] have recently emerged as a platform to support intelligent inter-vehicle communication to improve road safety. VANETs aim to provide vehicles and roads with capabilities to make roads more secure and to make driving time on the road more enjoyable, enabling communications among nearby vehicles (vehicle-to-vehicle communication) as well as between vehicles and nearby fixed equipment (vehicle-to-infrastructure communication). Concordantly, intelligent transportation systems (ITSs) have appeared to leverage the power of modern communication technologies, applied to both vehicles and infrastructure, in order to improve road safety.

Allowing drivers to report traffic accidents and violations through the VANET may lead to substantial improvements in road safety. Being able to do so anonymously in order to avoid personal and professional repercussions will undoubtedly increase user acceptance of such valuable service. Consider for example the potential risk incurred by a user who files a complaint against somebody who is also responsible for

*Corresponding author. Tel.: +34 93 401 1871.
Email addresses: ctripp@entel.upc.edu (Carolina Tripp Barba), luisfelipe@lfurquiza.com (Luis Urquiza Aguiar), monica.aguilar@entel.upc.edu (Mónica Aguilar Igartua), javier.parra@entel.upc.edu (Javier Parra-Arnau), david.rebollo@entel.upc.edu (David Rebollo-Monedero), jforne@entel.upc.edu (Jordi Forné), esteve@entel.upc.edu (Esteve Pallarès)

processing the corresponding violation. If such complaint were not anonymous, the reported individual may attempt to take action against the reporting user. Not to mention the fact that user behavior may be profiled on the basis of location and other sensitive information contained in the report.

Particularly in ad hoc networks, users may prefer not to place their trust on intermediaries such as anonymizing proxies [2] and mix networks [3, 4]. Privacy-enhancing technologies based on user collaboration avoid the need for these trusted third parties (TTP). On the other hand, it is crucial that the anonymity-enforcing mechanisms implemented be aware of their impact on network performance that translates into quality of user experience (QoE). Although there exists a number of collaborative anonymity systems in the literature [5, 6], to the best of our knowledge none of them is perfectly suited to the specific requirements of vehicular networks highlighted here.

With these challenges in mind, the main objective of this paper is to propose a new collaborative protocol for enforcing anonymity in multi-hop VANETs. The approach here presented is closely inspired by Crowds [5], a protocol according to which each user probabilistically decides to send a message directly to a common receiver, or else to forward it to a peer, who is asked to repeat the process. Our protocol differs from the original Crowds in that, first, it does take into account transmission losses, and secondly, it is specifically conceived for multi-hop VANETs, rather than for wired networks. Precisely, this second difference makes our approach dependent on the underlying multi-hop routing protocol, since the authority processing the reports may not be within communication range of every user. Motivated by this, this work assesses the suitability of our approach in combination with two standard routing protocols, AODV [7] and GPSR [8], and under two urban scenarios. Our extensive performance evaluation contemplates not only privacy, but also the impact on quality of service (QoS) of the privacy mechanism. On the one hand, QoS is measured in terms of packet loss, end-to-end delay and average number of hops; on the other, we measure anonymity as the attacker's probability of error when guessing the identity of the sender, in keeping with [9].

Sec. 2 examines the state of the art on anonymous-communication systems and reviews the routing protocols AODV and GPSR. Sec. 3 first describes the adversary model and anonymity metric assumed in this work. Afterwards, this section presents our anonymous-reporting protocol. Then, Sec. 4 is entirely devoted to the empirical evaluation of our approach under two distinct urban scenarios. Finally, conclusions are drawn in Sec. 5.

2. State of the Art

As stated previously, our main contribution is an anonymous-reporting protocol that, on the one hand, is inspired by the anonymous-communication protocol Crowds [5], and on the other, builds on a *generic* multi-hop routing protocol. In this section, we first provide a broad perspective of anonymous-communication systems, and secondly, describe in detail two widely-used routing protocols, one of them intended for mobile ad hoc networks, and the other specifically conceived for vehicular networks.

2.1. Anonymous-Communication Systems

In this subsection, we explore the underlying technologies of anonymous-communication systems. With this aim, we examine those systems based on the original mix devised by Chaum, and afterwards, analyze Crowds, a popular collaborative protocol for anonymous Web transactions.

In anonymous communications, one of the goals is to conceal who talks to whom against an adversary who observes the inputs and outputs of the anonymous communication channel. Mix systems [10, 3, 11] are nodes that forward messages so that it is unfeasible for an attacker to link an outgoing message to its corresponding input message. The idea behind Chaum's mix [3] is conceptually simple. Users wishing to submit messages to other peers encrypt the intended recipients' addresses by using public key cryptography and send these messages to the mix. The mix collects a number of these encrypted messages and stores them in its internal memory. Afterwards, these messages are decrypted and the information about senders is removed. In a last stage, when the number of messages kept reaches a certain threshold, the mix forwards *all* these messages to their recipients in a random order.

In the literature, this process of collecting, storing and forwarding messages when a condition is satisfied is normally referred to as a *round*. An important group of mixes called *pool* mixes operate on this basis.

Depending on the *flushing* condition, we may distinguish different types of pool mixes. Possibly, the most relevant form of pool mixes are *threshold* pool mixes [10], where the condition is imposed on the number of messages stored, as in the case of Chaum's mixes. The main difference is that threshold pool mixes do not flush all messages in each round, but keep some of them. Clearly, this strategy degrades the usability of the system—any incoming message can be stored in the mix for an arbitrarily long period of time. But on the other hand these systems achieve a better anonymity protection. The reason is that the set of possible incoming messages linkable to an outgoing target message increases substantially, as it includes all messages that entered the mix before this target message was flushed.

Another important group of pool mixes outputs messages based on time [12]. Essentially, these *timed* mixes forward all messages kept in the memory every fixed interval of time called timeout. The major advantage of these mixes is that the delay experienced by messages is upper bounded, in contrast to the case of threshold pool mixes. The flip side is that the unlinkability between incoming and outgoing messages may be seriously compromised when the number of messages arriving in that interval of time is small. Motivated by this, some of the current mix designs implement a combination of the strategies based on threshold and those based on time. Namely, these systems flush messages when a timeout expires, provided that the number of messages stored meets a threshold [13].

The use of networks of mixes has also been thoroughly studied in the literature. The reason is evident—on the one hand, routing messages through several mixes makes it more difficult for an attacker to track messages, and on the other hand, it improves the availability of the anonymous-communication system. Depending on the network topology, we may classify the existent approaches into *cascade mixes*, *free-route networks* and *restricted-route networks*. The application of cascade mixes was already suggested by Chaum in his original work [3]. Fundamentally, this approach contemplates the concatenation of mixes to endue the system with higher robustness. In contrast to this alternative where messages are routed through a fixed path, free-route networks recommend that users choose random paths to route their own messages [14]. In the end, restricted-route networks consider the case where every mix in the network is connected to a reduced number of neighboring mixes [15].

Apart from the systems based on mixes, other approaches attempt to anonymize the communication channel by relying on user collaboration [5, 16, 6]. An archetypical example is the Crowds protocol [5], a protocol originally designed to preserve the anonymity of users browsing the Web. The idea behind this protocol comes down to the notion of "blending into a crowd". That is, a set of users is organized to form a group or *crowd* which, acting as a single system, forwards requests to a Web site on behalf of its members. Specifically, suppose a user wishes to browse a Web site. With this aim, the user selects uniformly at random a member of the crowd, including themselves, and submits the request to that member. That member, in turn, decides with probability p to send the request directly to the Web site and with probability $1 - p$ to send it to another member, again chosen uniformly at random. In any of these two cases, the crowd member stores only the identifier of its predecessor, so as to enable a communication path for the Web site's response. This probabilistic forwarding is repeated until the request achieves its destination. The upshot of this process is that the Web site and any member of the group cannot guess the actual sender of a request, since they cannot distinguish its originator from a user who is simply submitting it on behalf of another user.

2.2. Routing Protocols in VANETs

As we mentioned in the introductory section, our anonymous-reporting protocol necessarily builds on a multi-hop routing protocol. This is because in vehicular networks users may not have a direct link with the infrastructure point in charge of processing their reports; and consequently, they must collaborate in the routing and forwarding processes. In this subsection, we summarize the main features of two multi-hop routing protocols, namely AODV [7] and GPSR [8]. The former was originally designed for mobile ad hoc networks (MANETs) but it is often used in performance evaluations of VANETs for comparative purposes with a well-known protocol. The latter was specifically conceived for vehicular networks. These two protocols will be used in Sec. 4 to assess the performance of our proposal.

Both protocols share one common characteristic—nodes periodically send *hello* messages to detect and monitor neighbors. The ad hoc on demand distance vector (AODV) [7] is a reactive protocol that uses the

Bellman-Ford distance vector to operate in a mobile environment. It determines a route to a destination only when a node wants to send a packet. Routes are maintained as long as they are needed by the source and while there is connectivity between nodes in the path. Route discovery is based on query and reply cycles, and route information is stored in all intermediate nodes along the route in the form of route table entries. The following control packets are used: route request messages (RREQ) are broadcasted by nodes requiring routes to other nodes, route reply messages (RREP) are unicasted back to the source of RREQ, and route error messages (RERR) are sent to notify nodes of the loss of a link, which will trigger a route recovery process. AODV implements a buffer to temporary store those packets for which the current node did not find a next forwarding node, to protect those packets instead of dropping them seeking to reduce losses. This protocol has been evaluated in different scenarios for VANETs [17, 18, 19] which have demonstrated that its performance is not suitable in high-mobility or low-density scenarios due to its inherent end-to-end operation.

Greedy perimeter stateless routing (GPSR) [8] is a well-known geographical routing protocol and one of the first routing protocols especially designed for VANETs. GPSR uses two algorithms to forward packets: greedy forwarding, which sends packets to the neighbor node closest to destination and is used by default; and perimeter forwarding, which is used in case greedy forwarding cannot be used. In perimeter mode, GPSR seeks to exploit the cycle-traversing properties of the right-hand rule to forward packets around voids when no closest neighbor is found. The main drawback of GPSR is the use of outdated information to select the next forwarding node due to possible inconsistencies in the neighbor tables or in the destination node's location [19].

3. A Protocol for the Anonymous Reporting of Traffic Violations

This section presents the major contribution of this work, a protocol that enables users to report traffic violations anonymously in vehicular ad hoc networks. Before we get into the details of our protocol, Sec. 3.1 examines the particular scenario of vehicular networks assumed. Later, Sec. 3.2 specifies the adversarial model considered in this scenario, both in terms of the attacker's objective and its strategy to compromise user privacy. Afterwards, Sec. 3.3 proposes a measure of anonymity so as to evaluate our approach. Finally, Sec. 3.4 is devoted to the description of the protocol. We anticipate that the design parameters of our approach are an underlying multi-hop routing protocol and a forwarding probability. The evaluation of our proposal under different routing protocols and interesting values of this probability is the object of the next section, Sec. 4.

3.1. Application Scenario

In our introductory section, we presented, in rather general terms, the motivating scenario of our work, i.e., vehicular ad hoc networks in *urban environments*. In this subsection, our purpose is to elaborate on the specific type of VANETs upon which our anonymous-reporting protocol builds.

As we mentioned in Sec. 1, this kind of networks allows users to both exchange messages among them and submit messages directly to the network infrastructure. These two forms of communications are referred to as vehicle-to-vehicle communication and vehicle-to-infrastructure communication, respectively. In addition, we commented that such communications enable users to report common traffic offenses such as speeding, red light violation or tailgating, as well as traffic incidents such as accidents and traffic jams. Further, it is expected that these networks will support entertainment applications in the near future. All these services are foreseen to be provided in the promising smart cities in the next years [20].

In this work, we contemplate a particular instance of these networks. First, in terms of applications, we restrict our analysis to the reporting of traffic violations, which does not exhibit real-time requirements as stringent as those of infotainment. Secondly, we consider packet losses due to collisions and a number of other causes explored in Sec. 4. And finally, we assume that there exists a single, fixed infrastructure point to which all users within a bounded area, e.g., neighborhood, send their reporting messages. We consider this last assumption describes a fairly realistic scenario, owing to the costly deployment of these networks.

In this specific scenario, we assume that the infrastructure point is not within the communication range of every user, what leads these users to use a multi-hop routing protocol so as to deliver their messages.

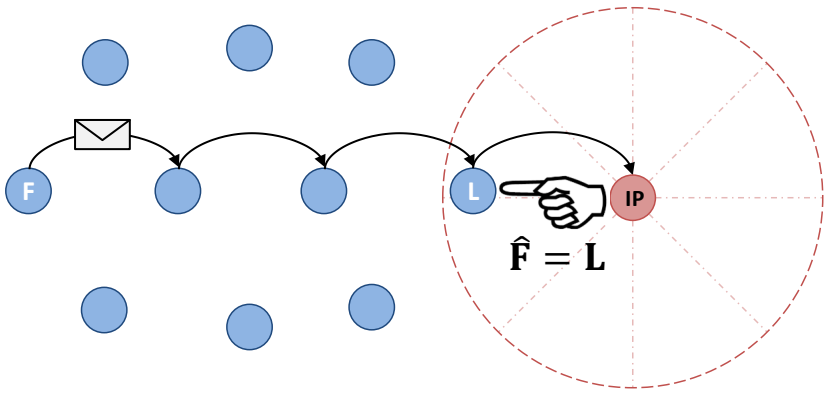

Figure 1: In our application scenario, users submit messages to the infrastructure point (IP) to report incidents of any kind. Concretely, users use a multi-hop routing protocol which enables them to communicate with the infrastructure even though they are not within the communication range of each other. Upon receipt of a message, the infrastructure point, playing the role of the privacy attacker, attempts to identify the actual sender of that message, F. Because the attacker has a limited view of the network, they know only the last user who forwarded it, L. From this information, the attacker estimates that the first sender is in fact the last one, $\hat{F} = L$.

Specifically, when a user witnesses a traffic violation, they immediately submit a message to the next user in the route determined by the routing protocol. The message includes identifying information about the offender, the GPS coordinates where the offense was committed and the time when it was observed. When this next user receives the message, they forward it to another user and so on until the message reaches its destination. When this finally happens, the infrastructure point adds the message to a traffic infractions list, which users can download afterwards to verify whether their reports have been received successfully.

3.2. Adversarial Model

Having described our application scenario, now we proceed to specify the concrete assumptions about the privacy attacker considered in this work. These assumptions refer to the capabilities, properties or powers of the attacker, and are known as the *adversarial model*. The importance of this model lies in that the level of privacy protection offered by a privacy-enhancing technology is measured with respect to it. In a nutshell, if the assumptions change, so does the metric.

In our scenario, it is the infrastructure point who plays the role of the privacy attacker. The users involved in the multi-hop routing protocol, however, are assumed to be partially trusted. This is in the sense that, while they are willing to route messages, they may attempt to compromise the privacy of users reporting traffic violations.

As we discussed in the previous subsection, Sec. 3.1, in our application scenario users reveal their position, though *not* their identity, when submitting reports to the infrastructure point. In this work we assume that, upon receipt of a message, the attacker's objective is to ascertain the *identity* of its originator. Note that, under this adversarial model, the re-identification of users is not the only privacy threat—based on the locations of the infractions reported by such users, the attacker could track them and ultimately extract sensitive information such as health-related issues, salary or religion. Just imagine a user who regularly drives to a hospital specializing in the treatment of AIDS and often reports traffic offenses on their way.

In the end, our last assumption has to do with the attacker's strategy to determine the originator of a message. More precisely, we assume that the attacker strives to guess the identity of the *first* sender (originator) of a given message, knowing *only* the user who *last* forwarded it. This could be interpreted as an adversary with a local view of the network, possibly due to their limited coverage range. Mathematically, we model the first sender and the last sender by the random variables (r.v.'s) F and L respectively, both taking on values in the alphabet $\{1, \ldots, n\}$. Since the only information available to the attacker is L, we

consider that the attacker's estimator of F is directly

$$\hat{F} = L, \tag{1}$$

that is, the attacker chooses the last sender as the originator of the message. Fig. 1 illustrates the assumptions of our adversarial model.

3.3. Anonymity Metric

In this subsection, we specify an anonymity metric consistent with the adversarial model described in Sec. 3.2. This metric will allow us to assess the anonymous-reporting protocol proposed in this work.

As we mentioned in the previous subsection, our attacker has a limited perspective of the network and, consequently, estimates the last user in the forwarding chain as the actual sender of a given message. According to this model, we define *sender anonymity* $\mathscr{A}$ as the attacker's probability of error when guessing the identity of the author of a given message, conditioned on the event $R =$ "the message is received successfully",

$$\mathscr{A} = \mathrm{P}\{\hat{F} \neq F|R\} = \mathrm{P}\{F \neq L|R\}. \tag{2}$$

Intuitively, the higher this probability of error, the higher the degree of anonymity attained by a user reporting incidents.

In the literature there exist several proposals that, like ours, measure anonymity as a probability of error and, more generally, as an attacker's estimation error. In the case of mix networks, for example, [21] establishes a connection between the probability of error of an eavesdropper who wishes to ascertain who is communicating with whom, and the degree of anonymity provided by such networks. The issue of quantifying anonymity in the context of location-based services (LBSs) is explored in [22] and revisited shortly afterwards in [23]. Specifically, the authors propose to measure privacy as the adversary's expected estimation error for that particular context. The attacker's estimation error is also proposed as a privacy criterion in [9], but under a much wider perspective that encompasses not only the application scenario of LBS, but also the fields of anonymous-communication systems and statistical disclosure control. The authors interpret and compare numerous metrics from these areas as particular instances of their more general measure of privacy.

3.4. Description of the Protocol

In this subsection we present the main contribution of our work, an anonymous protocol for the reporting of traffic offenses in vehicular ad hoc networks. The primary purpose of this protocol, as its name indicates, is to enable users of these networks to report traffic infractions in a manner that neither the infrastructure point nor the users participating in the protocol cannot compromise the anonymity of reporting users.

Our approach builds on top of a generic multi-hop routing protocol. We refer to this protocol as the *underlying* routing protocol, or simply, as the routing protocol. As we shall see later on in Sec. 4, the choice of this protocol will determine the performance of our approach, both in terms of QoS and user anonymity. On top of this routing protocol, and operating at the *application* layer, we find our own protocol. Since our approach is greatly inspired by the popular anonymous-communication system Crowds [5], next we underline the main differences with respect to our proposal. We refer the reader to Sec. 2.1 for a complete description of the original Crowds protocol.

- First, Crowds is a uni-hop protocol inasmuch as every member can forward a request directly to a Web site, or in general, to an untrusted receiver. Our approach, which is specifically designed for a *mobile* scenario, does not contemplate this possibility, though. Basically, this is due to the limited communication range of vehicles in our scenario. As a consequence of this limitation, a user participating in the protocol cannot forward a message to any member of the system. Instead, the user has to content themselves with submitting it to one of their neighboring members, that is, those within their coverage range.

- Secondly, we do *not* introduce a mandatory initial forwarding step, as Crowds does. The reason is that such initial step would double the minimum possible message forward count from 1 to 2, imposing a price on average delay which, in the context of the intended applicability of our work, we deem more than significant.

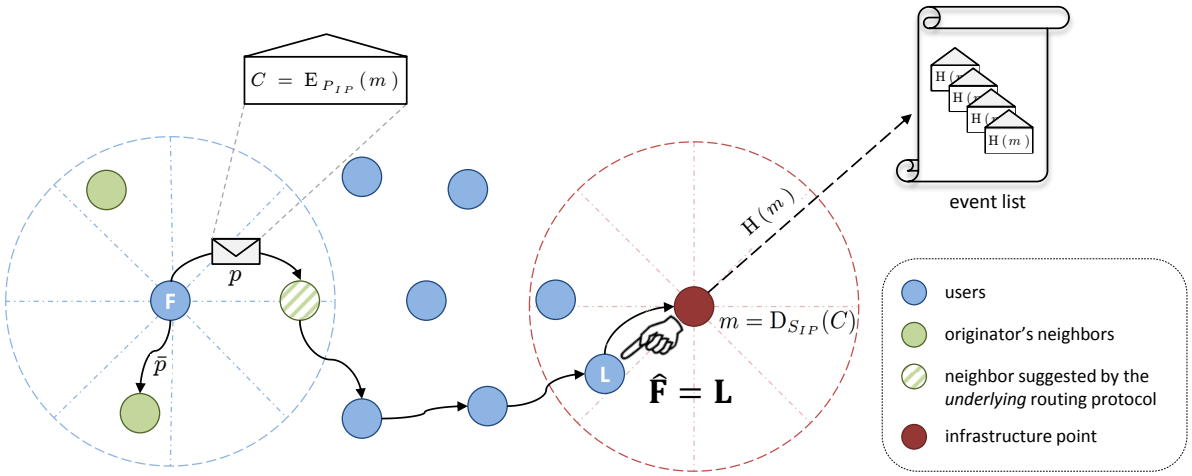

Figure 2: Operation of the anonymous-reporting protocol for vehicular ad hoc networks. In this figure, a user wants to report a traffic violation to the infrastructure point in a manner that their privacy is not compromised. For this purpose, the user first generates a message m including information about the offender, time and location; and then, encrypts it with the infrastructure point's public key P_{IP}. After this, the user must decide who will be the next user in the route towards the infrastructure point. Unlike the original Crowds, the user selects the next forwarding user among their neighbors. Specifically, the user flips a biased coin and depending on the outcome chooses between these two options—either to submit the ciphered message $C = \mathrm{E}_{P_{IP}}(m)$ to the neighbor suggested by the routing protocol, or to forward it to another user. The former option is the best in terms of QoS requirements, but the worst in terms of privacy. The contrary happens when the user adopts the latter strategy.

- Lastly, in our protocol users forwarding messages need not store information about their predecessors. As we shall see later, users will confirm the reception of their messages by downloading a list of encrypted messages, released by the infrastructure point.

Having examined the differences between Crowds and our approach, next we shall set forth the details of our anonymous-reporting protocol. For this purpose, consider the scenario described in Sec. 3.1 and a multi-hop routing protocol used by every user in said scenario. Under these premises, suppose that a user witnesses a traffic offense and decides to report it to the infrastructure point. To this end, the user creates a message containing the following three elements—first, identifying information of the complainant; secondly, a time interval around the instant when the infraction was detected; and finally, a perturbed version of the location where the incident was committed. We note that the accuracy of the spatio-temporal information given by the user will depend, on the one hand, on the type of infraction, and on the other, on their privacy requirements.

These three elements form the message m to be conveyed to the infrastructure point. However, since the users involved in the protocol are not fully trusted, our approach contemplates that the originator of the message encrypts it with the public key of the infrastructure point, P_{IP}; the result of this is the ciphered message $C = \mathrm{E}_{P_{IP}}(m)$. After encrypting the message, the user finds out which users are within their coverage range. Among these neighboring users, the routing protocol chooses a candidate to forward the message and communicates it to the user. We would like to note that the choice of this candidate will vary depending on the specific routing protocol assumed. With this information, the user decides either to send the message to that candidate, or to submit it to another neighboring user chosen uniformly at random. Similarly to the original Crowds, the user opts for the former strategy with probability p, and adopts the latter with probability $1 - p$. Clearly, the former (latter) strategy is the best (worst) in terms of QoS guarantees, but the worst (best) in terms of anonymity. For compactness, we write $1 - p$ as $\bar{p}$.

Upon receipt of the message by this neighboring user, they repeat the same forwarding technique. That is, the user chooses probabilistically whether to submit the message to the neighboring user recommended

by the routing protocol, or to send it to another user within their coverage range. This process goes on until the message reaches its destination. When this is the case, the infrastructure point cannot ascertain whether the user who last forwarded the message was actually its originator or was merely forwarding it on behalf of another user. This is due to the probabilistic nature of the forwarding process, to the perturbation of the spatio-temporal information included in the message, and to the fact that our approach works at the application layer.

After the reception of the message, the infrastructure point decrypts it with its secret key S_{IP}, and finds out about the traffic violation reported, $m = D_{S_{IP}}(C)$. Then, it generates the hash value of the message, $\mathrm{H}(m)$, which is incorporated into a list of (encrypted) traffic offenses. Afterwards, this list is made available to users, what ultimately allows them check whether their messages have been received. Fig. 2 illustrates the operation of our protocol.

In the end, we would not like to finish without emphasizing the fact that our approach strongly depends on two factors—first, the forwarding probability p; and secondly, the concrete multi-hop routing protocol integrated into our protocol, which decides the next forwarding user in the route towards the infrastructure point. The remainder of this work is entirely devoted to the evaluation of the performance of our proposal, both in terms of QoS and anonymity, under the consideration of two underlying routing protocols and different values of p.

4. Experimental Results

This section presents a number of experimental results that will allow us to evaluate our anonymous-reporting protocol in terms of anonymity protection on the one hand, and QoS requirements on the other. With this purpose, we first describe the simulation environment in Sec. 4.1. Then, we assess our approach under two scenarios, namely a residential area and another representing a downtown district.

4.1. Simulation environment

In this subsection, we present the simulation environment that we shall use to assess the performance of our anonymous-reporting protocol. As we described in Sec. 3, our approach is determined, first, by a multi-hop routing protocol assumed to be installed in the vehicle of each user; and secondly, by the forwarding probability p. Recall that p is the probability that the next forwarding step is that dictated by the underlying routing protocol, for best QoS, rather than a random neighbor choice for better anonymity. Throughout this section, we shall use the terms *vehicle* and *user* interchangeably.

In our series of experiments, we evaluate our proposal under two standard routing protocols, namely AODV and GPSR. In both cases, we consider $p = 5\%, 25\%, 50\%, 75\%, 100\%$. Note that a forwarding probability $p = 5\%$ implies that almost all messages are forwarded at random, which intuitively enables users to achieve a high level of privacy protection, clearly at the cost of high penalties in terms of QoS. On the contrary, $p = 100\%$ may be interpreted as if our anonymous-reporting protocol were *deactivated*, meaning that users always forward messages according to their routing protocols. The intuition behind this parameter p is illustrated conceptually in Fig. 3, where we show the protocol architecture assumed in our experiments.

An important aspect of our empirical evaluation is the consideration of two simulation scenarios. Both scenarios represent a vehicular ad hoc network in an urban context. The former does not include any building, what may resemble a residential neighborhood with landscaped areas and single-family housing. The latter does consider the presence of buildings. This may model a downtown area or a central business district. We shall refer to these two environments as *residential* scenario and *downtown* scenario, respectively.

On the other hand, the simulation platform chosen to conduct our experiments was NCTUns 6.0 [24], an open-source software widely used by the research community. In order to assess our approach under the aforementioned routing protocols, we had to implement, in a first stage, the protocol GPSR. This implementation is available to other researchers at `https://sertel.upc.edu/~maguilar/simulators.html`. In a second stage, we set the simulation parameters. In particular, the propagation model selected was the two-ray ground, which considers both the direct path and a ground-reflection path. We assumed a data rate of

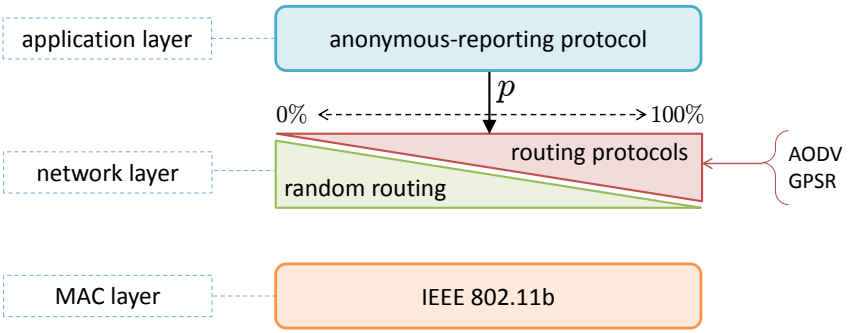

Figure 3: Protocol architecture of our approach. Our proposal is an anonymous-reporting protocol that, on the one hand, operates at the application layer, and on the other, builds on top of a multi-hop routing protocol. In our experiments, we assume the IEEE 802.11b as the MAC protocol, and consider two standard routing protocols in conjunction with our approach.

11 Mbit/s. Concretely, we contemplate that each user transmits constant bit rate (CBR) flows at a rate of 1 packet every 2 seconds, and that each flow is composed of 1000-byte packets. On the other hand, since the standard IEEE 802.11p is not completely implemented in NCTUns, we decided to use the IEEE 802.11b. The reason is that both standards do have the same MAC operation. The only significant difference is that IEEE 802.11p improves the physical layer for high speeds, so that we could expect even better results for this standard.

In our simulation scenario, vehicles move according to the mobility generator CityMob [25]. This mobility generator allows us to create realistic urban mobility scenarios where vehicles drive on streets and respect traffic signals. We selected the Manhattan mobility model. Our scenarios have 60 nodes randomly positioned on the streets. Loosely inspired by a regular area in the city of Barcelona, Spain, these scenarios are based on a grid of 8 x 8 streets , 40 m wide, outlining 7 x 7 square blocks with a side length of 100 m. There are traffic lights at each intersection. At intersections, vehicles randomly can turn right, left or continue in the same direction. In the end, consistently with the adversarial model described in Sec. 3.2, we consider a single, fixed infrastructure point (IP) to which users submit their messages. Fig. 4 depicts the simulation scenario described above.

As a final remark, we would like to point out that we have conducted 5 simulations for each scenario. Accordingly, the figures in Secs. 4.2 and 4.3 reflect confidence intervals (CI) of 99% obtained from 5 repetitions per point. Table 1 shows the most representative simulation parameters.

4.2. Residential Scenario

Our first scenario is a residential area, which we characterize in our simulator by the absence of buildings. This special case has the advantage that radio signals cannot be blocked by walls. This is in contrast to the case tackled in Sec. 4.3, where the presence of buildings models a downtown district.

Our experimental results confirm the intuition that anonymity level, packet loss, delay and average number of hops decrease with p. GPSR yields a higher anonymity level than AODV for p under 75%, as Fig. 5 indicates. This is because with GPSR, more packets require a higher number of hops to reach the IP. The reason lies in the hop-by-hop routing scheme used by GPSR, in contrast with AODV's need for end-to-end paths. Packets arriving from paths with more than one hop have a higher anonymity level, because of the increased difficulty for the IP in guessing the actual sender.

Both GPSR and AODV exhibit the same tendency with p in terms of average number of hops. From Figs. 6 and 7, we observe that, as p grows, the number of hops decreases and so the percentage of packet losses. Also, we notice from Fig. 6 that AODV presents with a lower number of hops than GPSR, specially for low values of p. This is because after several hops, the end-to-end path might be interrupted and, if so, the routing protocol would need to find a new path, probably shorter. However, AODV loses a higher number of packets during the recovery time of the path.

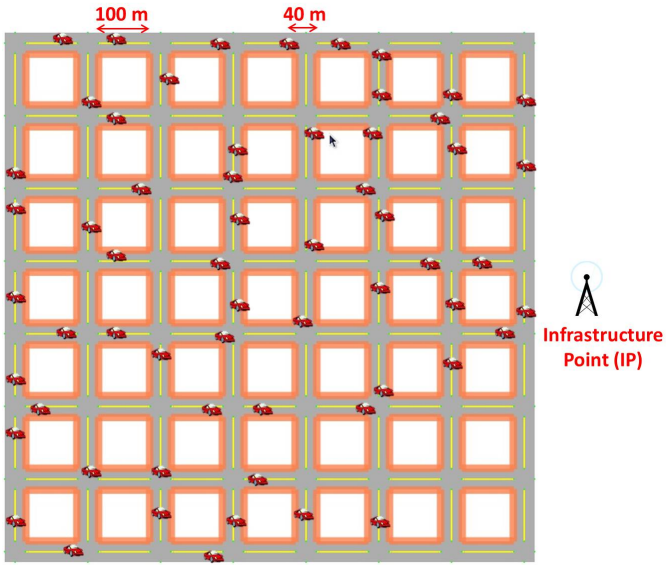

Figure 4: In our experimental analysis, we contemplate two simulation scenarios, with the same grid geometry shown here, but differing on whether the relevance of wall blocking is modeled.

Table 1: Simulation settings

Parameter	**Value**
Medium capacity	11 Mbps
Packet size	1000 bytes
Traffic source	CBR
Transmission range (T_x)	250 m
Carrier sensing range (S_x)	300 m
Simulation time	1000 sec
MAC specification	IEEE 802.11b
Area	1020 m x 1020 m
Maximum speed	50 km/h (14 m/sec)
Number of nodes	60
Propagation channel model	Two Ray Ground
Radio propagation	Rician
Mobility generator	CityMob [25]
Mobility model	Manhattan
Routing protocol	AODV and GPSR
Forwarding probability p	5%, 25%, 50% 75% and 100%

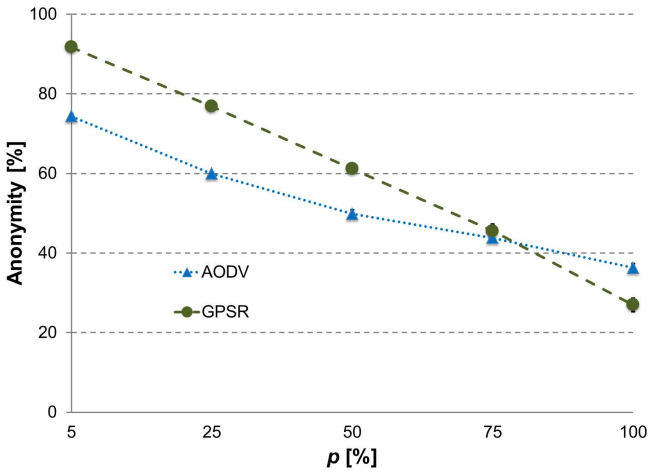

Figure 5: Anonymity level provided by our anonymous-reporting protocol in the residential scenario described in Sec. 4.2. We measure anonymity as the attacker's probability of error when guessing the identity of the originator of a message. Our approach is evaluated in the case when the underlying multi-hop routing protocols are AODV and GPSR, and for different values of the forwarding probability p.

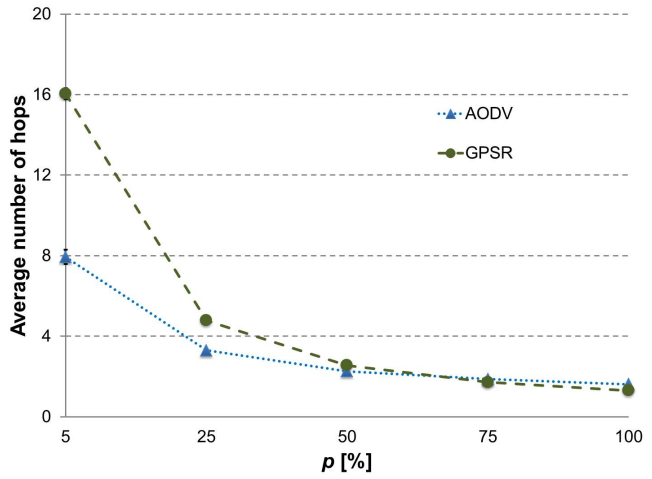

Figure 6: Average number of hops in the residential scenario. The forwarding probability p determines whether the next node in the route towards the infrastructure point will be chosen according to the routing protocols AODV and GPSR, or at random.

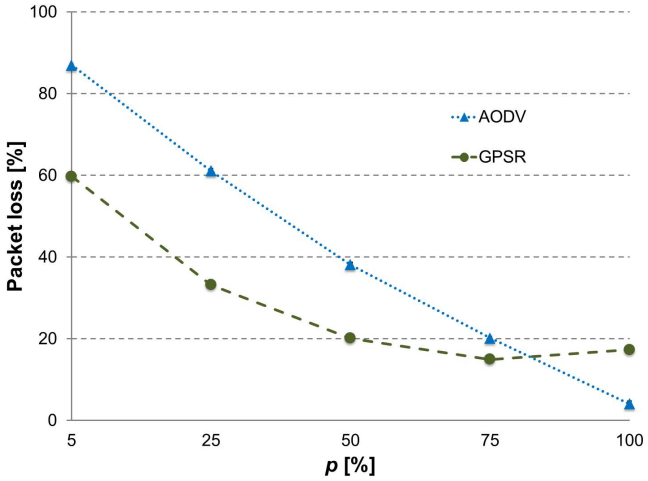

Figure 7: Percentage of packet loss in a residential scenario. The protocols AODV and GPSR are evaluated for different values of the forwarding probability p. As expected, the higher this probability, the lower the percentage of packet losses.

Fig. 8 plots the average packet delay. GPSR presents the highest delay because it delivers the highest number of packets that successfully arrived after many hops. On the other hand, AODV presents the lowest delay, except for a forwarding probability $p = 5\%$. The reason is that, for this or similar values of p, the next forwarding user is almost always chosen uniformly at random. This causes paths to break, and as a result, the protocol needs to resort more often to recovery processes —higher number of RREQ and RREP packets to find a new path—, which ultimately increases delay. AODV experiences less delay than GPSR because the computation of the average delay is based on packets that actually reach their destination, with a lower number of hops compared to GPSR.

As for packet losses, GPSR performs better than AODV for almost all values of p, as depicted in Fig. 7. As expected, we observe in both protocols that the percentage of packet losses is relatively high for low values of p. This is because packets wander around along longer paths, so that the chance of losing a packet is higher. AODV, on the other hand, provides a lower anonymity protection than GPSR because the former makes use of a buffer. The explanation lies in the high packet loss ratio, basically due to collisions. More precisely, nodes try to periodically send buffered packets waiting for a near forwarding node, resulting in a higher probability of collisions with packets from other nodes.

From Fig. 7, we may conclude that GPSR is the best routing protocol in terms of losses, except for $p \geq 80\%$ where AODV outperforms GPSR. This is because the latter does not implement any buffer. Particularly, for high values of p, GPSR is forced to drop a high number of packets if it does not find a next forwarding node. As expected, we observe that low values of p result in high losses and vice versa. That is, a low p means that the packet will follow a path mainly formed by nodes randomly chosen, producing a longer path which increases losses. Conversely, a high p means that the packet will be forwarded according to the original scheme of the protocol, consequently having shorter paths and lower losses. Lastly, note from Figs. 7 and 5 that, in the case of the GPSR protocol, a forwarding probability of 50% leads to a anonymity gain of 35%, at the cost of an increase of only 4% in packet losses. In other words, the anonymity gain is much greater than the packet loss incurred by the adoption of our anonymous-reporting protocol.

4.3. Downtown Scenario

This section contemplates a simulation environment modeling a downtown scenario. This scenario is characterized by the presence of buildings, which may block the transmission signal, in contrast to the case described in Sec. 4.2. We anticipate that the presence of these buildings will lead to a decrease in the number of neighbors to whom forward packets, since vehicles behind a wall will be discarded as possible

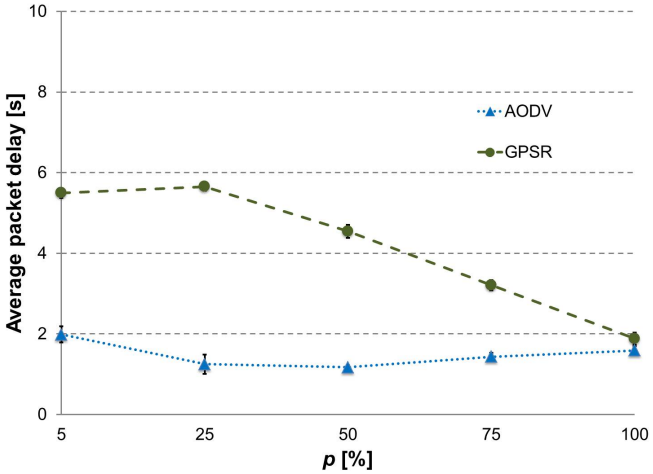

Figure 8: Average packet delay in the residential scenario contemplated in Sec. 4.2.

next hop. As a matter of fact, in this scenario we will observe that the percentage of packet loss will be much higher than in the residential scenario explored in Sec. 4.2.

One of the interesting findings is that, when our anonymous-reporting protocol is disabled, i.e., $p = 100\%$, AODV and GPSR do provide a high level of anonymity, as shown in Fig. 9. This is because not all the neighbors are in the same coverage range, so multi-path routes may be needed to reach the IP. This happens more often in the scenario with buildings, and naturally the anonymity level with $p = 100\%$ is higher compared to the residential scenario. In addition, the presence of buildings may lead to an increase in the path length.

The AODV protocol provides a higher anonymity level because of its buffer. The fact that packets may be stored for a certain period of time contributes to maintain longer paths and thus makes it more difficult for the IP to know who sent the packet. GPSR, however, does not have such a buffer. Consequently, the anonymity protection resulting from the integration of GPSR and our anonymous-reporting protocol is relatively lower. But this is true only for $p \geqslant 50\%$, when the choice dictated by the routing protocol prevails over the random one. Taking into account that a packet that hopped more than once can be considered as anonymous, the constant level of anonymity observed in Fig. 9 is due to the presence of buildings that often lead to a path length greater than one. Equation (1) computes the anonymity level of the sender of a given message, which will be given a high value in those paths longer than one hop. For instance, AODV compared to itself in the scenario without buildings maintains a high level of anonymity (see Figs. 5 and 9) because it establishes longer paths to achieve the IP (see Figs. 6 and 10).

Regarding the average number of hops, GPSR performs better than AODV for $p \geqslant 50\%$, as it is shown in Fig. 10. This behavior is due to the fact that as p increases the forwarding criterion of each protocol is used more often. AODV establishes end-to-end paths that are mostly longer, due to the high density of the scenario. Conversely, GPSR achieves destination using a hop-by-hop scheme, thus it selects the closest node each time using less hops than AODV.

Fig. 11 shows the average delay. AODV provides the lowest delay in all the points with $p \leq 75\%$, since most of the packets received were from paths with a low number of hops. When p=100% GPSR obtained the best result. This is because packets with a low number of hops reach destination when this protocol is used.

As for packet loss, we see in Fig. 12 that the AODV protocol presents an intuitive tendency: the higher the forwarding probability p, the lower the packet loss. In this figure we also observe that GPSR performs worse than AODV for $p >$75%. This is because the former protocol does not use a buffer and does not verify whether the next forwarding node is actually reachable or not (e.g., because it is behind a wall), which significantly increases packet loss.

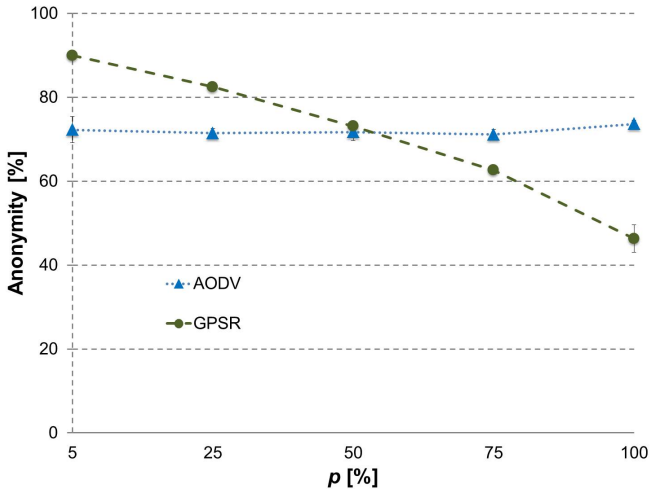

Figure 9: Anonymity level offered by our protocol in the special case of a downtown scenario, when our approach is integrated with two standard routing protocols. We measure anonymity as the probability of error of an adversary who strives to ascertain the identity of the originator of a given message.

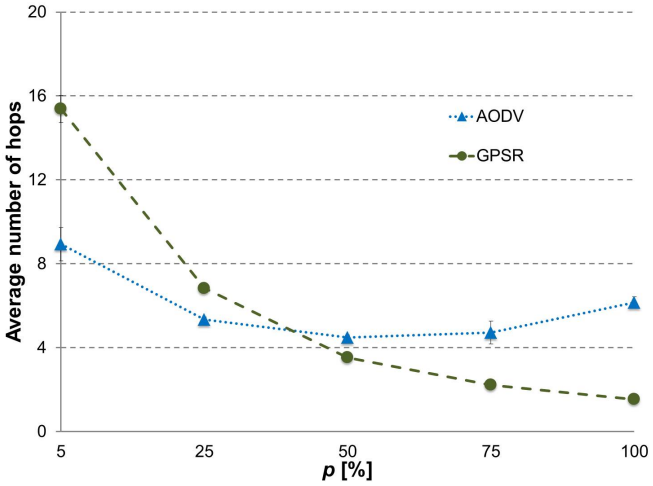

Figure 10: Average number of hops in an urban vehicular ad hoc network. The simulation environment corresponds to a downtown district.

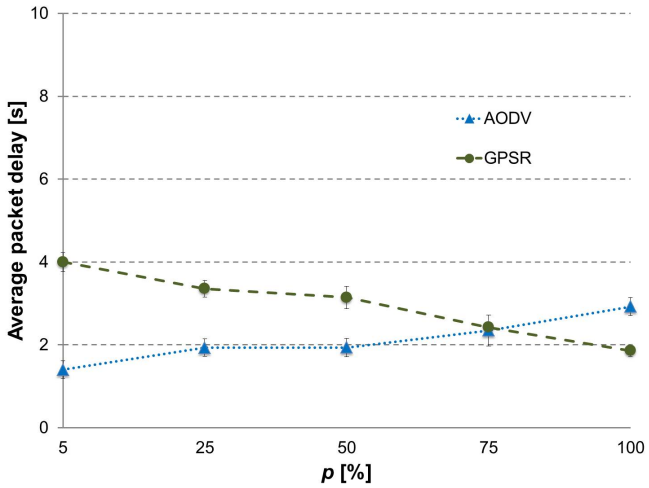

Figure 11: In this figure, we plot the average packet delay as a function of the forwarding probability. The scenario contemplated is a downtown district, characterized in our simulator by the presence of buildings.

On the other hand, GPSR yields lower losses when $p \leq 50\%$ because the forwarding path is hop-by-hop and the packet could arrive to destination after several hops. In addition, the buffer in the case of AODV produces more losses than it can save due to the generation of collisions. These collisions are produced when the node constantly tries to forward the packets stored in the buffer, increasing the use of the common wireless medium. This results in a high probability of collisions between the packets sent from the buffer and those forwarded by other nodes. Further, AODV outperforms GPSR because the buffer starts to avoid losses when the random forwarding is less prevalent, i.e., when p increases.

In the end, we summarize the major conclusions drawn from our experimental evaluation:

- First, in the special case when our anonymous-reporting protocol is deactivated, i.e., $p = 100\%$, we verify that the underlying routing protocols under study, AODV and GPSR, still provide a certain level of anonymity protection. This is due to the nature of the multi-hop behavior of the forwarding path. Accordingly, the IP will sometimes fail in guessing the actual sender of a packet from the last forwarding node who delivered the packet.

- Secondly, the results exhibit a strong dependance of our approach on the forwarding decision criteria of the aforementioned routing protocols. In the case of AODV, for example, the buffer helps to increase the anonymity level. This is because a packet sent from the buffer may hinder the IP in its efforts to ascertain the actual sender of a given packet. On the flip side, buffering leads to losses for small values of p, since the node constantly tries to forward the packets stored in the buffer, ultimately increasing the use of the medium and the number of collisions.

- In the downtown scenario, paths are commonly longer than one hop due to the presence of buildings. This increases the probability that the IP fails to unveil the identify of the originator of a packet, and therefore contributes to a higher level of anonymity protection. A packet is already anonymous when coming from a path of more than one hop, because it is unlikely for the IP to guess its actual source.

- Finally, our empirical analysis shows that GPSR outperforms AODV in the residential and downtown scenarios, both in terms of packet loss and anonymity level achieved.

5. Conclusions

In recent years, vehicular ad hoc networks have caught the attention of both industry and academia, since they are seen as a means of improving road safety in current and future transportation systems. Among the

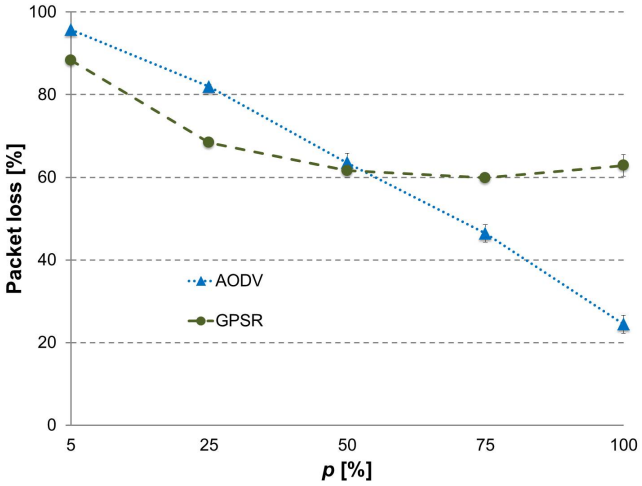

Figure 12: Percentage of packet losses in the downtown scenario considered in Sec. 4.3. An eye-opening finding is that for $p = 50\%$, the GPSR protocol experiences approximately the same packet loss ratio, but exhibits an anonymity gain of 24%, as Fig. 9 confirms.

potential safety applications enabled by such networks, one of the most promising is the reporting of traffic violations such as speeding or tailgating.

Allowing users to report traffic offenses through the VANET may contribute to improve road safety on the one hand, but on the other it may pose serious privacy threats; not only because of the risk of reindentification but also because of the risk of profiling. Current anonymous-communication systems based on TTPs are clearly an inappropriate approach to this problem, largely due to the ad hoc nature of these networks. Other approaches, not requiring infrastructure, rely on the principle of user collaboration. Although the literature abounds with examples of collaborative anonymous-communication systems, the fact of the matter is that none of them are suited to the specific requirements of the networks at hand.

Our main contribution is precisely a collaborative protocol that enables users to report traffic violations anonymously in multi-hop VANETs. Operating at the application layer, our approach is essentially inspired by the popular anonymous protocol Crowds. Our anonymous-reporting protocol depends on a forwarding probability that determines whether the next forwarding step in message routing is random, for better anonymity, or in accordance with the routing protocol on which our approach builds, for better QoS. In terms of implementation, our approach only requires that users collaborate in the same way as they do when they route messages of other peers. Put differently, the proposed protocol could, in principle, be easily deployed and integrated into any multi-hop routing protocol.

This contribution is also valuable from a methodological perspective, as we illustrate a systematic approach to the performance evaluation of usability and privacy of a privacy-enhancing technology. In keeping with [9], we measure anonymity as the probability of error of an adversary who endeavors to ascertain the identity of the sender of a given message. Together with message delay and end-to-end losses as standard measures of QoS, this anonymity measure enables us to quantitatively assess the multi-objective performance of the proposed mechanism.

Concordantly, we wrap up our contribution with an extensive empirical assessment of our proposal, in terms of both QoS guarantees and anonymity protection, under the consideration of two standard routing protocols, AODV and GPSR. Our approach is evaluated in two urban environments, namely a residential scenario and a downtown district, and parameterized by the probability p of direct compliance with the forwarding step dictated by the routing mechanism, versus random forwarding. Interestingly, the routing protocols per se, that is, for $p = 100\%$, when they are not integrated into our anonymous protocol, already provide a reasonable degree of anonymity. As p decreases, favoring a random choice in the next immediate forwarding step, our experimental results quantify the trade-off between anonymity and QoS expected.

According to these results and for both of the scenarios considered, GPSR outperforms AODV, in terms of packet loss and anonymity. For both scenarios, GPSR routing, and values of p in the 75-100% range, corresponding to high QoS, it must be pointed out that packet losses are not noticeably affected, and that the relative increment in anonymity is roughly of the same order as the relative increment in delay incurred. This empirically supports and quantifies the usefulness of Crowds-like collaborative strategies in multi-hop VANETs for better anonymity, when users are willing to pay an eminently reasonable price in QoS.

Acknowledgments

This work was partly supported by the Spanish Government through projects Consolider Ingenio 2010 CSD2007-00004 "ARES", TEC2010-20572-C02-02 "Consequence" and by the Government of Catalonia under grant 2009 SGR 1362. Carolina Tripp has a FI-AGAUR grant of the "Comissionat per a Universitats i Recerca del DIUE" from the Generalitat de Catalunya and the Social European Budget. She has also obtained a grant from the Autonomous University of Sinaloa, Mexico. Luis Urquiza is the recipient of a grant from the Secretaria Nacional de Educación Superior, Ciencia y Tecnología SENESCYT and the Escuela Politécnica Nacional (Ecuador). D. Rebollo-Monedero is the recipient of a Juan de la Cierva postdoctoral fellowship, JCI-2009-05259, from the Spanish Ministry of Science and Innovation.

References

[1] S. Olariu and M. Weigle, *Vehicular Networks: From Theory to Practice*, ser. Chapman & Hall/CRC computer and information science series. CRC Press, 2009.

[2] V. Benjumea, J. López, and J. M. T. Linero, "Specification of a framework for the anonymous use of privileges," *Telemat., Informat.*, vol. 23, no. 3, pp. 179–195, Aug. 2006.

[3] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Commun. ACM*, vol. 24, no. 2, pp. 84–88, 1981.

[4] "The Tor project, Tor: Overview." [Online]. Available: http://torproject.org/overview.html.en

[5] M. Reiter and A. D. Rubin, "Crowds: anonymity for Web transactions," in *ACM Trans. Inf. Syst. Secur.*, vol. I, 1998, pp. 66–92.

[6] D. Rebollo-Monedero, J. Forné, A. Solanas, and T. Martnez-Ballesté, "Private location-based information retrieval through user collaboration," *Comput. Commun.*, vol. 33, no. 6, pp. 762–774, 2010. [Online]. Available: http://dx.doi.org/10.1016/j.comcom.2009.11.024

[7] C. Perkins and E. Royer, "Ad-hoc on-demand distance vector routing," in *IEEE workshop on mobile computing systems and applications*, 1999, pp. 90–100.

[8] B. Karp and H. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th annual international conference on Mobile computing and networking*, ser. MobiCom. ACM, 2000, pp. 243–254.

[9] D. Rebollo-Monedero, J. Parra-Arnau, C. Diaz, and J. Forné, "On the measurement of privacy as an attacker's estimation error," *Int. J. Inform. Secur.*, 2012, to appear. [Online]. Available: http://arxiv.org/abs/1111.3567

[10] L. Cottrell, "Mixmaster and remailer attacks," 1994. [Online]. Available: http://obscura.com/~loki/remailer/remailer-essay.html

[11] G. Danezis, R. Dingledine, and N. Mathewson, "Mixminion: Design of a type III anonymous remailer protocol," in *Proc. IEEE Symp. Secur., Priv. (SP)*, Berkeley, CA, May 2003, pp. 2–15.

[12] A. Serjantov and R. E. Newman, "On the anonymity of timed pool mixes," in *Proc. Workshop Priv., Anon. Issues Netw., Distrib. Syst.* Kluwer, 2003, pp. 427–434.

[13] U. Möller, L. Cottrell, P. Palfrader, and L. Sassaman, "Mixmaster protocol – Version 2," Internet Eng. Task Force, Internet Draft, Jul. 2003. [Online]. Available: http://www.freehaven.net/anonbib/cache/mixmaster-spec.txt

[14] M. Rennhard and B. Plattner, "Practical anonymity for the masses with mix-networks," in *Proc. Int. Workshop Enabling Technol.: Infra. Col. Enterprises (WETICE).* IEEE Comput. Soc., 2003, pp. 255–260.

[15] G. Danezis, "Mix-networks with restricted routes," in *Proc. Workshop Priv. Enhanc. Technol. (PET).* Lecture Notes Comput. Sci. (LNCS), 2003, pp. 1–17.

[16] C. Chow, M. F. Mokbel, and X. Liu, "A peer-to-peer spatial cloaking algorithm for anonymous location-based services," in *Proc. ACM Int. Symp. Adv. Geogr. Inform. Syst. (GIS)*, Arlington, VA, Nov. 2006, pp. 171–178.

[17] S. Jaap, M. Bechler, and L. Wolf, "Evaluation of Routing Protocols for Vehicular Ad Hoc Networks in City Traffic Scenarios," in *Proc of the 11th EUNICE Open European Summer School on Networked Applications*, 2005, pp. 584–602.

[18] J. Härri, F. Filali, and C. Bonnet, "Performance comparison of AODV and OLSR in VANETs urban environments under realistic mobility patterns," in *Med-Hoc-Net 2006, 5th IFIP Mediterranean Ad-Hoc Networking Workshop*, Jun. 2006, pp. 324–336.

[19] V. Naumov, R. Baumann, and T. Gross, "An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces," in *Proceedings of the 7th ACM international symposium on mobile ad hoc networking and computing*, ser. MobiHoc. ACM, 2006, pp. 108–119.

[20] C. Tripp, M. Mateos, P. Regañas, A. Mezher, and M. Aguilar, "Smart city for VANETs using warning messages, traffic statistics and intelligent traffic lights," in *IEEE Intelligent Vehicles Symposium (IV'12)*, june 2012, pp. 902–907.

[21] P. Venkitasubramaniam, T. He, and L. Tong, "Anonymous networking amidst eavesdroppers," *IEEE Trans. Inform. Theory, Special Issue Inform.-Theor. Secur.*, vol. 54, no. 6, pp. 2770–2784, Jun. 2008.

[22] R. Shokri, J. Freudiger, M. Jadliwala, and J. P. Hubaux, "A distortion-based metric for location privacy," in *Proc. Workshop Priv. Electron. Society*, 2009, pp. 21–30.

[23] R. Shokri, G. Theodorakopoulos, J. Y. L. Boudec, and J. P. Hubaux, "Quantifying location privacy," in *Proc. IEEE Symp. Secur., Priv. (SP)*. Washington, DC, USA: IEEE Comput. Soc., 2011, pp. 247–262.

[24] S. Y. Wang, C. L. Chou, C. H. Huang, C. C. Hwang, Z. M. Yang, C. C. Chiou, and C. C. Lin, "The design and implementation of the NCTUns 1.0 network simulator," in *Computer Networks*, 2003, pp. 175–197.

[25] F. Martinez, J. Cano, C. Calafate, and P. Manzoni, "CityMob: A Mobility Model Pattern Generator for VANETs," in *ICC Workshops'08. IEEE International Conference on Communications*, may 2008, pp. 370–374.