Certificate Revocation List Distribution in Vehicular Communication Systems

April 14, 2008

Abstract

A number of projects have been developing security architectures for Vehicular Communication (VC) systems, with consensus on utilizing public key cryptography to secure communications. In spite of their advanced status on many aspects, none of these projects, to the best of our knowledge, has investigated and addressed the problem of Certificate Revocation List (CRL) distribution. As the need to evict compromised, faulty, or illegitimate nodes from the VC system is commonly accepted, our contribution here is a solution tailored to the requirements and constraints of the VC systems. Our design is scalable and efficient, and can deliver seamlessly CRLs to all nodes within a region within tenths of minutes. More general, our analysis and simulation evaluation set the basis for the design of such CRL distribution systems, showing how to configure them to achieve more stringent requirements.

1 Introduction

Vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication will enhance transportation safety and efficiency. Recently, however, efforts on vehicular communication (VC) security have been undertaken, as it has been well-understood that VC systems are vulnerable to attacks and that the privacy of their users is at stake. For example, an attacker could inject messages with false information, or collect vehicle messages to track their locations and infer sensitive user data. Three major efforts to design security and privacy enhancing solutions for VC are currently underway, along with the efforts of the Car-to-Car Communication Consortium (C2C-CC): the NoW project [5], the IEEE 1609.2 working group [7], and the SeVeCom project [10].

A few basic ideas transcend all these VC security architectures: they all build on top of a currently accepted networking protocol stack, with primary security requirements being message authentication, integrity, and non-repudiation, as well as protection of private user information. They all rely on a *Certification Authority (CA)*, and public key cryptography to protect V2V and V2I messages, with each node (VC-equipped vehicle or RSU) are registered one CA and can participate in the network operation. Of course, it is also clear that a node equipped with a certificate is not necessarily complying with the implemented protocols, neither is it operating correctly (e.g., it may simply inject faulty data). A line of defense against faulty or compromised nodes is crucial for the trustworthiness of the VC system. A well-understood method is to defend the system is to evict the misbehaving nodes. In the context of current secure VC architectures, *revocation of the certificate* of such nodes is an appropriate approach that has been utilized in other types of systems. Moreover, revocation can be useful for other reasons; for example, credentials of stolen vehicles can also be revoked. Once revoked, messages from that nodes will be ignored by system nodes.

The problem of revocation in VC systems has hardly attracted any attention in the literature, as explained in Sec. 6: the need for CRL distribution is discussed in [5,7,10], and aspects of it are addressed in [10, 11]. In all these documents, the distribution of Certificate Revocation Lists (CRLs) has a central importance, it is understood that only the CA can revoke a node (include the node's certificate in a CRL), and that renewal of the CRL (with exclusion of prior and inclusion of new entries) is relatively infrequent. Distributed mechanisms that allow the identification of misbehaving nodes are proposed [11] to bridge this gap until a misbehaving node's certificate is revoked. Yet, in all works in the Vehicle Ad Hoc Networking (VANET) literature on revocation, the distribution of the CRL is the final and definitive line of defense.

However, none of the above-mentioned prominent efforts on security for VC, and to the best of our knowledge no other proposal in the literature, have been concerned with the fundamental problem of *how to distribute the CRL* across a large-scale and multi-domain system as the vehicular communication systems. More important, how to deliver in a timely manner the appropriate CRL to all vehicles without an omni-present fixed infrastructure, and how to do so as the number of equipped vehicles will grow? Furthermore, how to ensure that the CRL distribution protocol will incur low overhead, especially due to the limited bandwidth? In this paper, we address exactly these questions.

Our contribution in this paper the first investigation on the distribution of CRLs in vehicular ad hoc networks. As alluded in [5, 7, 10], leveraging on fixed infrastructure, albeit sparsely connected, is the appropriate choice, and for this we elect the use of RSUs, as they will in fact be deployed, e.g., in the US by public authorities, as enablers of

VC systems. But these RSUs will be assigned tasks that are more time-critical than the CRL distribution (e.g., junction warnings), and, often, in dense traffic conditions the wireless channel will be already congested.

To meet such constraints that are unique to the VC setting, we propose and evaluate here a scheme tailored to the specific constraints of VANET-enabled systems. We propose a collaboration scheme between regional CAs that allows CRLs to contain only regional revocation information; a low-rate, randomized method for RSUs to broadcast the CRL; the use of erasure codes to enhance the robustness and flexibility of the CRL distribution. Our scheme does not require any communication and cooperation between RSUs on the CRL distribution task, and minimizes the CA-RSU and vehicle-CA-RSU interactions. Our results show that allocating a bandwidth of few KBytes/s to the CRL distribution broadcast is sufficient for a very high percentage of (practically, all) vehicles to receive securely the complete CRL within minutes.

In the rest of the paper, we provide the system model in Sec. 2, followed by a more precise problem statement and solution overview in Sec. 3. The detailed description of our scheme components is provided in Sec. 4, and an analytical and simulations-based evaluation of our scheme is given in Sec. 5. We conclude the paper after a survey of related work (Sec. 6), and a discussion.

2 System Model

Existing administrative processes run by automotive authorities (e.g., departments of motor vehicles) offer a paradigm for the organization of the (large number of) Certification Authorities (CAs) that will be necessary for the deployment of secure vehicular communication systems. In accordance with the proposals of the IEEE 1609.2 [7] and the SeVeCom architecture [10], each CA is responsible for the identity management of all vehicles registered in its *region* (national territory, district, county, etc.). This results in a forest of hierarchical structures and *cross-certification* among high-level CAs. Vehicles registered with different CAs can thus communicate securely as soon as they validate the certificate of one CA_A on the public key of some CA_B. Various procedures for easily obtaining these cross-certificates can be implemented. Moreover, the deployment of secure vehicular communications could still be handled locally, to a great extent, especially as those systems' penetration increases gradually.

Each node, vehicle or RSU, registered with exactly one CA, has a unique identity V and a pair of *private* and *public* cryptographic keys, k_V and K_V, respectively, and obtains a certificate $Cert_{CA}\{V, K_V, A_V, T\}$, with A_V is a list of node attributes and T the certificate lifetime. The CA issues such certificates for all nodes upon registration, and upon expiration of the certificate. The CA is also responsible for evicting nodes from the system, if necessary, either for administrative or technical reasons. The interaction of nodes with the CA does not need to be continuous, and in fact it cannot be. This is so, because the road-side infrastructure that acts a gateway for the CA to the vehicular part of the network will not fully cover the vehicular network area. On the other hand, other infrastructure-based networks (e.g., cellular) that have already broad coverage, cannot be assumed to enable basic security functionality or substitute for V2I communication. This is primarily due to: the cost of the cellular links, and the absence of consensus (thus far, at least) between telecommunication providers and automakers on the convergence of the two industries.

In principle, credentials and cryptographic keys correspond to a long-term identity of the node. Nonetheless, all three currently available (under development) security architectures for VC [5, 7, 10] also propose the use of short-lived keys and credentials to secure communication. This is the concept of *pseudonymity* or *pseudonymous authentication*: each vehicle is equipped with multiple certified public keys (pseudonyms) that do not reveal the vehicle identity, and the vehicle uses those pseudonyms alternately, each for a short period of time, so that messages signed under different pseudonyms cannot be linked.

The exact nature and use of certified public keys is largely orthogonal to the CRL distribution per se. The only difference would be the size of the CRL: if a revoked vehicle is equipped with multiple pseudonyms, its eviction would require the inclusion of all its unexpired pseudonyms to the CRL. In contrast, the IEEE 1609.2 working group is proposing the use of a large pool pseudonyms that are shared among vehicles: each vehicle picks a small subset of those and utilizes them alternately. Typical values for the overall pool of pseudonyms (key pairs essentially) and chosen subset are $PN = 10000$ and $n = 5$ respectively. In the former case, since neither NoW nor SeVeCom have a recommendation on the rate of pseudonym change, it is hard to estimate the number of additional certificates that would need to be revoked. In the latter case though, each revocation would require the inclusion of a constant n entries in the CRL. Between these two extremes, here we simply vary our estimate of the CRL sizes for our systems, as explained in Sec 5, and work with the more pessimistic assumption that the CRL size is proportional to the number of nodes registered with the CA.

Regarding the communication, we consider the essentially globally recommended use of a variant of the IEEE 802.11 technology. As the IEEE 802.11p/DSRC radios [1] have not been extensively evaluated, we consider 802.11a, and channel models investigated in the literature [12] with typical communication ranges up to 200m, and the typical communication patterns, (e.g., 10 beacons per sec per vehicle). In terms of cryptographic primitives, we consider in this paper also EC-DSA [2].

3 Problem Statement and Scheme Overview

Certificate revocation is deemed necessary in VC to ensure their trustworthiness, essentially by evicting illegitimate or faulty nodes or simply preventing the use of compromised cryptographic material. We do not dwell on the policies that govern node eviction, and neither the process of collecting evidence that specific credentials must be revoked, as those can hinge on legal issues and are beyond the scope of this paper. Rather, our focus is to design a system that enables the timely distribution of a Certificate Revocation List (CRL) *after* its most recent and up-to-date version is created by the CA. The intended recipients are all nodes (vehicles) circulating within the region (domain) of a CA.

More precisely, we are interested in designing a system that upon generation of a new CRL at time t_0, the CA will have the CRL distributed within a delay Δ to some fraction x of all nodes that circulated in the domain of the CA for at least Δ seconds after t_0. The challenge is to design a system that can push Δ to low values, and thus timely delivery, and accordingly push x to 1.

Such values can be system-selectable parameters. but since this problem has never been addressed before, notably, in the light of its geographical dimension and the intermittent CA-to-node connectivity, the charting of the design space and the identification of all the trade-offs and the investigation of the achieved performance is the first step towards understanding, for example, which Δ values are achievable and for which cost. Nonetheless, there several challenges salient to the VC operational environment that should be addressed. Our design relies on the use of the roadside infrastructure to disseminate the CRLs, a practical approach as equipment that will be deployed for VC is used, without additional cost. The basic idea is to disseminate simultaneously the CRL across the entire CA region, but in a way that does not degrade VC for other more time-critical operations (such as safety applications). We outline next the basic components of our scheme towards achieving these objectives, and provide additional details in Sec. 4:

Regional CRLs Since the size of a CRL is proportional to the number of vehicles registered with a CA, it is straightforward to maintain a CRL that corresponds only to a specific regional CA. We denote the most recent, up-to-date CRL for a given CA as CRL_{CA}. The motivation, stemming from the hierarchical CA structure, is to keep the CRLs small in size.

CA Collaboration As vehicles travel across geographical boundaries of CA regions, regional revocation information should flow across those boundaries. Thus, CAs exchange over the wire-line Internet CRLs. However, they do not merge those and do not disseminate them in their regio. Rather, we propose that each CA issues and manages short-lived *Foreigner Certificates* (FCs) for visiting vehicles V registered with another CA. If such an FC is revoked, it is included in the CRL_{CA} as an own certificate. In principle, only a small fraction of vehicles registered with some CA_A will enter the domain of some CA_B, the increase of the CRL_{CA_B} size will be accordingly small and thus keep the CRL size low.

Multi-RSU CRL Distribution The challenge is to distribute the CRL_{CA} at any point of the CA region. The deployment of RSUs, often by the same organization that instantiates the CA, makes it natural for the CA to leverage on them for the CRL distribution. Vehicles are able to obtain CRL_{CA} from any such RSU, or when needed, complete the "download" of the CRL_{CA} with the help of multiple RSUs.

Segmented, Erasure-Coded Protected, Secure CRLs With an error-prone wireless channel and relatively low communication ranges in high density situations, and often low vehicle-RSU contact times due to mobility, it is highly likely that a CRL download is never be completed.* To address this challenge, we propose that:

- The CRL_{CA} is segmented in M *CRL pieces*, which we denote by $p_1, p_2, \ldots, p_M$.
- The segmented CRL_{CA} is then encoded by an *erasure code*, which adds limited redundancy and produces $N > M$ pieces, such that for any M out of N pieces received the original CRL_{CA} can be reconstructed. Each of the N pieces $p_1, p_2, \ldots, p_N$ is transmitted by the RSUs across the wireless medium.
- The values parameters M and N are chosen to be large, and more specifically $N >> M$. Note that erasure codes are applicable with small N, M and simply $N > M$; for example, [**?**] can be used to encode a CRL_{CA} so that any $M = 2$ out of $N = 3$ pieces suffice for reconstruction. However, large M, and thus small piece size, along with $N >> M$ increase resilience to errors (packet loss) and short RSU connection time, as well as repetitive reception of the same piece by different RSUs.
- Each CRL_{CA} carries the CRL version identifier and timestamp, the CA and identifier, the piece sequence number, and the digital signature σ_{CA} of the CA. As a result, each piece can be validated individually, and an adversary cannot inject forged or outdated CRL pieces. As soon as any M out of N pieces, with their integrity, freshness, and authenticity validated, are obtained, the CRL_{CA} is constructed.

*Maintaining a file pointer and resuming the download when connection to another RSU is established is conceivable. But this would entail complexity and overhead, due to RSU-to-CA or RSU-to-RSU communication.

Minimal RSU-CA and No RSU-RSU Interactions The CA constructs the CRL_{CA} and forwards all its N pieces to all the RSUs within its region. This is an one-time operation within the lifetime of the CRL_{CA}. The RSUs will distribute the CRL pieces to the wireless part of the secure VC system, without any modification. Moreover, they do not request from the CA any service and neither do they forward any requests from vehicles to the CA related to the CRL distribution. Similarly, no communication between RSUs takes place with respect to the CRL distribution.

Randomized, Low-Rate Broadcast Distribution Each RSU randomly shuffles the N pieces of the encoded CRL_{CA}, and commences their transmission across the wireless data link. The transmission is broadcast, without any acknowledgement from the receiving vehicles. The rate of broadcast is r_B pieces/sec, and it is chosen so that in bytes/sec the bandwidth (data rate) consumed by CRL pieces $r_B << C$, with C the bandwidth the data link can support.

The uncoordinated broadcast is independent of the individual vehicle trajectories. With N,M selected as explained above, it is unlikely that pieces previously lost (due to mobility or channel errors) or pieces previously received and validated will be repeatedly lost or received during successive RSU encounters. The broadcast transmission is appropriate as CRLs are necessary to any vehicle, it allows the RSU to avoid keeping track of vehicles within range and their communications, and transfer the state to other RSUs. The low rate, r_B, is chosen so that the CRL distribution does not interfere with the rest of the VC operation. The lower the r_B the more bandwidth available for safety and traffic efficiency applications traffic, thus the more reliable the reception of those time-critical messages.

Overall, a major concern is *scalability*, achieved by keeping the CRL size low. The simplicity in design, with minimal RSU-CA and no RSU-RSU interactions, is a second critical factor also contributes to scalability: the number of RSUs can increase to very high numbers, with at most a linear increase in the CA-RSU communication, the infrequent, once per CRL version "push" of the N pieces from the CA to each RSU across the *wireline* network. Moreover, there is no dependence of our CRL distribution scheme on the number of vehicles in the network, which they will be in significantly larger in number than RSUs.

4 Scheme Components

4.1 CA Collaboration

Regional CAs collaboration is mainly based on the issuing and management of *foreigner certificates* (FCs), whose issuance takes place before a node V's entering the host CA_F's geographical area. Once in the new area, V utilizes only the foreigner certificate, which is readily recognizable as FC by a specific field and thus cannot be used by V to obtain yet another FC with another CA. We define a *foreigner certificate* $\{h, K_h, fr\}_F$ as a short lived certificate that CA_F issues to a valid vehicle h belonging to region H and visiting region F. For CA_F to consider h as valid, and to issue it a *foreigner certificate*, $\{h, K_h\}_H$ should not be present in the CRL of CA_H. h uses it *foreigner certificate* to communicate with other vehicles during its journey in F. In case CA_F detects a misbehavior of h, CA_F inserts $\{h, K_h, fr\}_F$ in its CRL and notifies CA_H which inserts $\{h, K_h\}_H$ in its CRL. Therefore, the CRL of any CA_X only contains revoked certificates that CA_X issued and *foreigner certificates* of vehicles from other regions that misbehaved during a journey in X.

H, F, G	Regions in the system
CA_X	CA responsible of region X
CA_{Root}	The root CA
h	A vehicle belonging to region H
$\{h, K_h\}_H$	A regular certificate issued by CA_H to h
$\{h, K_h, fr\}_F$	A foreigner certificate issued by CA_F to h
CRL_X	CRL of region X

4.1.1 Foreigner Certificate Delivery Protocol

As a vehicle h enters a foreign region F, it initiates a *foreigner certificate* delivery protocol with CA_F in order to obtain a *foreigner certificate*. Communication between the vehicle h and CA_F is performed over some RSU in region F. The vehicle h would also need to initiate the protocol if it is still in region F when its *foreigner certificate* expires. The FC is delivered by the protocol below, for which the time-stamp in the request is intended to defend against replay attacks.

$$\begin{array}{lll} vehicle\, h & & CA_F \\ & \xrightarrow{\{h, current\, time\}_{k_h}\, , \{h, K_h\}_{CA_H}} & \\ & \xleftarrow{\{h, K_h, fr\}_{CA_F}\, , \{F, K_F\}_{CA_{root}}} & check\, \{h, K_h\} \notin CRL_H \\ & \xrightarrow{\{h, current\, time\}_{k_h}} & \end{array}$$

4.1.2 Revocation of foreigner certificates

Foreign vehicles can misbehave while in the host region, and there must be a way in this case to revoke their certificates and to inform other CAs in the system. Typically, foreign vehicles stay a short period in a host region. Therefore, their certificates need only to have a relatively short lifetime, which makes their revocation less challenging.

Implicit revocation One method is implicit revocation. Here, the lifetime of *foreigners certificates* is very short e.g. in the order of one day. If a vehicle is still in a foreign region when its *foreigner certificate* expires, it requests a new one. After running the protocol described above, the vehicle is issued a new *foreigner certificate*. If CA_F detects that h is misbehaving, it notifies CA_H, which inserts $\{h, K_h\}$ in CRL_H. This way, it is always enough for CA_F, or any other CA_X to check CRL_H in order to know about the

status of a vehicle belonging to region H. CA_F could also insert $\{h, K_h\}$ in CRL_F or simply keep it in a local black list for its own records. This approach is particularly good if foreign vehicles generally only transit by the host region.

Explicit revocation Another method is explicit revocation. Here *foreigners certificates* lifetime is longer than in the previous scheme, but still shorter than for normal certificates. If CA_F detects a misbehavior of h, the same steps are performed as for implicit revocation. Here, however, CA_F will necessarily insert $\{h, K_h\}$ in CRL_F. Moreover, CA_F needs to periodically check that $\{h, K_h\} \notin \mathrm{CRL}_H$ for the entire lifetime of the *foreign certificate* of h. This is to catch a situation in which h goes to its home region H, or another region G and performs some misbehavior while in there, and then comes back to region F wishing to use a still valid *foreigner certificate* $\{h, K_h, fr\}_{CA_F}$.

4.2 CRL Construction

CRL Encoding with Erasure Codes The M pieces of the CRL are encoded into $N > M$ pieces using an Erasure code. This allows to decrease the probability of reception of repeating pieces. In Erasure codes, some redundancy is added to the message, then the message and the redundancy are divided into pieces. The encoding of a message first segments the original message of length FS into L sequence of characters, each of length M, with padding if necessary. The segments of the original message are arranged as columns of a $M-$by$-L$ array B, to which a linear transformation is applied to get a $N-$by$-L$ array W. The linear transformation is such that it is possible to reconstruct the original message, when having any number of rows of W that is slightly larger than M. The CA encodes the M pieces of the CRL into N pieces using an Erasure code, and sends the encoded pieces to the RSUs in its region. A large $\frac{N}{M}$ would make the probability of reception of a duplicate piece smaller, but would increase the coding and decoding complexity.

CRL Encoding with Fountain Codes A fountain code produces for a given set of k input symbols $(x_1, \cdots, x_k)$ a potentially limitless stream of output symbols $z_1, z_2, \cdots$. Each output symbol is the sum of a randomly and independently chosen subset of the input symbols. It is possible to recover the input symbols with high probability from any subset of the output symbols that is slightly larger than k.

Universal Raptor codes are a special class of fountain codes that has liner time encoding and decoding. For a given set of k input symbols $(x_1, \cdots, x_k)$ and any real $\epsilon > 0$ a potentially limitless stream of symbols $z_1, z_2, \cdots$ such that any subset of symbols of size $k(1+\epsilon)$ is sufficient to recover the original k symbols with high probability. Each output symbol is generated using $O(log(\frac{1}{\epsilon})$ operations, and the original symbols are recovered from the collected ones with $O(k\, log(\frac{1}{\epsilon}))$ operations.

5 Performance Evaluation

In this section, we evaluate the performance of the different CRL distribution schemes introduced in the paper. We evaluate the resources that need to be allocated for the CRL distribution for different system parameters in order for vehicles to receive the entire CRL within a reasonable amount of time. We mainly consider the size of the CRL, the range and the density of RSUs, and the penetration ratio of VANET capabilities. Our results show that in most situations, when using Erasure codes or Fountain codes, allocating a bandwidth in the order of 1 *kbps* is sufficient.

5.1 Analytical evaluation

We first evaluate the expected CRL sizes, and then we compute T the time after which a vehicle completes the reception of a newly issued CRL while varying the range of RSUs R, the distance between them D, and the bandwidth allocated to the transmission.

5.1.1 CRL size

Let N_{eq} be the total number of equipped vehicles in the region that the CRL needs to cover, p_r the average proportion of certificates revoked per time period e.g. day ,and L_f the life time of a certificate in times periods. Let N_{CRL} be the number of certificates in the CRL. N_{CRL} is the number of non expired certificates that were revoked.

The average number of revoked certificate each time period is $N_{eq} * p_r$. We assume that a revoked certificate has en equal probability to become revoked at any time period of its life time $i \in \{1, \ldots, L_f\}$. When a certificate is revoked at time period i of its life time, it will stay in the CRL for $L_f - i$ time periods. Thus, the expected time a revoked certificate stays in the CRL is $E(L_f - i) = \frac{L_f}{2}$. From the above two equations, the expectation of N_{CRL} is $E(N_{CRL}) = (N_{eq} * p_r) * \frac{L_f}{2}$. We consider the CRL to consist of the identifiers of all the revoked certificates plus a security overhead of 500 Bytes. We expect that 4 Byte identifiers should be sufficient for vehicles, as this the current size of the IP space nowadays. Therefore, the expected CRL size is $E(S_{CRL}) = E(N_{CRL}) * 4\, Bytes + 1 KBytes$. The table in Fig. 1 from the National Insurance Crime Bureau gives numbers about the motor vehicle theft in the top ten U.S. metropolitan areas in 2005, ranked by the rate of vehicle theft reported per 100,000 people based on the 2000 Census, and the size of the CRL that would result from including the identifiers of all these vehicles in the CRL. In the early deployment phase, only a ratio $r < 100\%$ of vehicles will be equipped, and thus the size of the CRL would be $r\,.\,E(S_{CRL})$. According to FBI's Uniform Crime Reports, $1,237,114$ motor vehicles were reported stolen. Inserting all the identifiers of these vehicles would result of a CRL of $5 MBytes$. The coordination between regional CAs makes it possible to distribute regional CRLs, and in this paper we do not try to distribute a national CRL.

Metropolitan Area	Vehicles stolen	Rate	CRL size (KBytes
Modesto, CA	7,071	1,418.80	30
Las Vegas/Paradise, NV	22,465	1,360	91
Stockton, CA	7,586	1,167.30	32
Phoenix/Mesa/Scottsdale, AZ	41,000	1,103	165
Visali/Porterville, CA	4,257	1,060	18
Seattle/Tacoma/Bellevue, WA	33,494	1,057	135
Sacramento/Arden-Arcade/Roseville, CA	20,268	82	
San Diego/Carlsbad/San Marcos, CA	28,845	983	116
Fresno, CA	8,478	978.11	35
Yakima, WA	2,212	965.54	10

Figure 1: Motor vehicle theft, top ten U.S. metropolitan areas, 2005

5.1.2 Total number of CRL pieces

Let M be the number of pieces in which a CRL is cut. A vehicle receives a number of pieces P from each RSU it encounters. It completes the reception of the CRL with high probability after it collects a total of P_{tot} pieces from n consecutive RSUs it encounters. P_{tot} is the main parameters that changes between schemes.

Scheme without encoding Let random variable X represent the number of pieces that allowed a vehicle h to recover its CRL. A 99.99% confidence interval for X is $[E(X) - 3.9\sqrt{var(X)}, E(X) + 3.9\sqrt{var(X)}]$, and therefore $P_{tot} = E(X)+3.9\sqrt{var(X)}$. For this scheme, counting X is equivalent to counting the the number of balls that need to be tossed in order to fill M bins where each ball is equally likely to fall into any of the M bins and that the tosses are independent.

Let random variable X_j represent the number of tosses required to have a ball land in a $j+1th$ bin once j bins contain a ball for $j = 0, 1, \cdots, M$. We have $X = \sum_{j=1}^{M} X_j$.

After j bins contain a at least a ball each, the probability that the next ball tossed will fall into a new bin is $\frac{M-j}{M}$. Therefore, the random variable X_j has a geometric distribution

$$P(X_j = k) = \left(\frac{j}{M}\right)^{k-1} . \frac{M-j}{M}$$

The expectation of X is

$$\begin{aligned} E(X) &= \sum_{j=1}^{M} E(X_j) \\ &= \sum_{j=1}^{M} \frac{M}{M-j} \\ &= M \sum_{j=1}^{M} \frac{1}{j} \end{aligned}$$

The variance of X is

$$\begin{aligned} Var(X) &= \sum_{j=1}^{M} \frac{\frac{j}{M}}{\left(\frac{M-j}{M}\right)^2} \\ &= M \sum_{j=1}^{M} \frac{j}{(M-j)^2} \end{aligned}$$

Scheme using Erasure codes The analysis is very similar to the the one for the scheme without coding, except that here we count the number of balls that need to be tossed in order to fill the first M bins out of N bins, where each ball is equally likely to fall into any of the N bins and that the tosses are independent. We have

$$P(X_j = k) = \left(\frac{j}{N}\right)^{k-1} . \frac{N-j}{N}$$

The expectation of X is

$$\begin{aligned} E(X) &= \sum_{j=1}^{M} E(X_j) \\ &= \sum_{j=1}^{M} \frac{N}{N-j} \end{aligned}$$

The variance of X is

$$\begin{aligned} Var(X) &= \sum_{j=1}^{M} \frac{\frac{j}{N}}{\left(\frac{N-j}{N}\right)^2} \\ &= N \sum_{j=1}^{M} \frac{j}{(N-j)^2} \end{aligned}$$

Scheme using a Fountain code For the broadcast using Fountain codes, we simply have $P_{tot} = (1+\epsilon)M$.

Comparison When no encoding is used, $P_{tot} >> M$ causing a considerable waste of bandwidth. In figure 2, P_{tot}

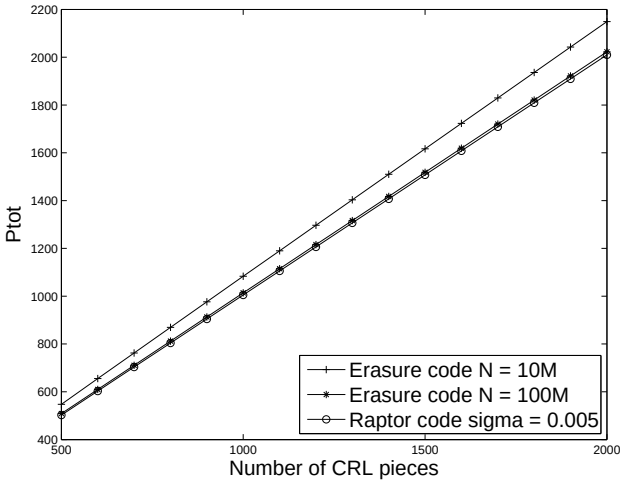

Figure 2: Number of pieces to be received vs. number of pieces in the CRL

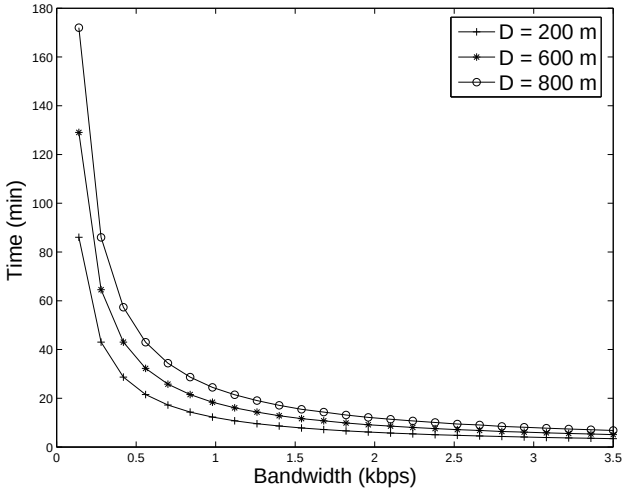

Figure 3: Time vs. bandwidth

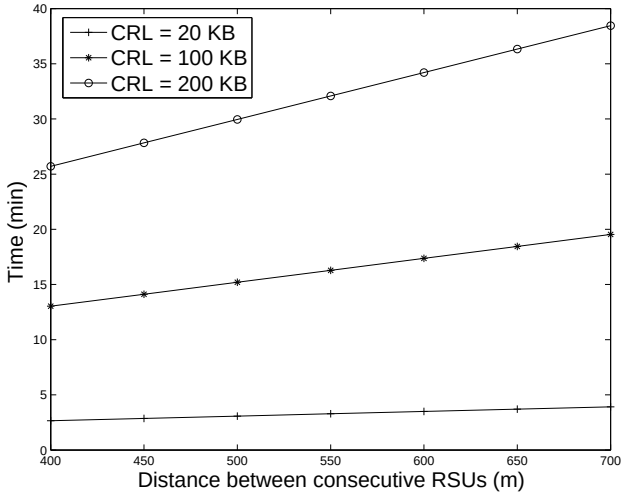

Figure 4: Time vs. distance between RSUs

is depicted for the case when Erasure codes and Fountain codes are used and for different values of M. Both codes perform equivalently well achieving $P_{tot} \approx M$. This is due to the fact that the probability of receiving duplicate pieces is small for an Erasure code with $N >> M$. The decision on which coding scheme to use should be based on the speed, computation, and memory complexity of the codes. The existence of patents are another important aspect to consider. We do not make a distinction between the coding schemes for further analysis.

5.1.3 Time to complete the CRL

In this section, we compute T the time for a vehicle to complete the reception of a newly issued CRL. The vehicle drives at velocity v, and encounters consecutive RSUs $1, 2, 3, \ldots, i, \ldots$ separated by distances $D_1, D_2, \ldots, D_{i-1}, \ldots$. RSUs have range R and broadcast CRLs at a bandwidth B. Let $pr(d)$ be the probability of reception of packets in function of the distance d of the vehicle from the RSU. If sz is the size of the CRL data contained in each broadcasted CRL packet and oh the size of overhead to the packet, then the number of pieces received from one RSU can be found as $P = \frac{B}{sz+oh} * \frac{R}{V}$. The number of RSUs n a vehicle needs to encounter to complete its CRL is $n = \frac{P_{tot}}{P}$. The total time to complete the CRL is therefore

$$T = \frac{1}{V}\left[\sum_{1}^{n-1} D_i + R\right]$$

5.2 Simulations

We simulate an urban scenario, in the shape of a grid, but without the road segments at the perimeter. Essentially, the grid we utilize has N vertical and N horizontal roads, and thus N^2 intersections, in the shape of a hash (#). Each road has two lanes in opposing directions. The distance between two consecutive horizontal (and two consecutive vertical) roads is L. The intersections we consider have:

- *No Traffic Lights*, with vehicles yielding to vehicles arriving from their right[†].
- *Synchronized Traffic Lights:* Intersections are equipped with synchronized traffic lights, so that the light is green in the horizontal direction at all intersections at the same time.

Vehicles enter the simulated area from all $4N$ end-points of the grid. We experiment with the rate of car entry, f, to increase the vehicle traffic congestion level and thus the mobility of vehicles. The entrance of new vehicles is randomized but the aggregate results in an in-flow of f vehicles/hour/lane. The vehicles select a destination when they

[†]The default rule in Europe, which differs from the "four way stop" rule used in the United States

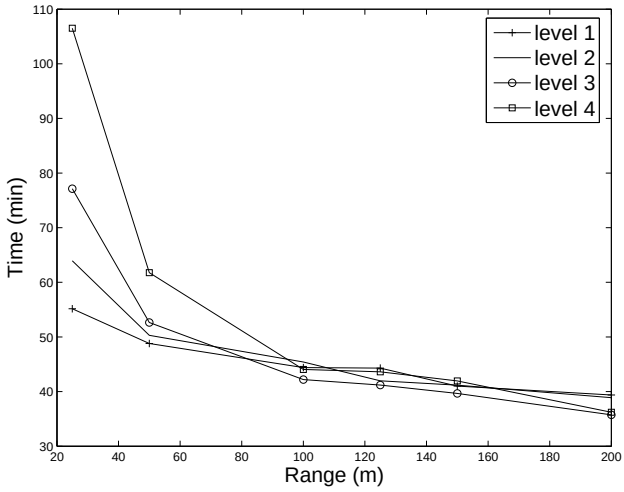

Figure 5: Time vs. range of RSUs, traffic lights

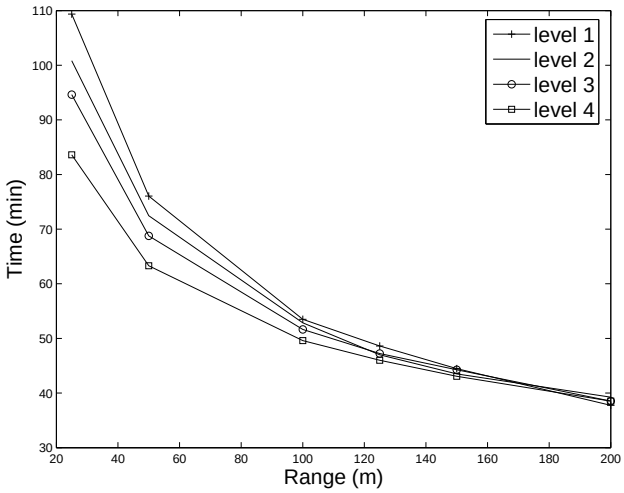

Figure 6: Time vs. range of RSUs, no traffic lights

enter the network, choosing uniformly randomly their destination among the N road end points at the opposite side of the road network. For realistic vehicle mobility, we use SUMO [4], a microscopic, continuous-space and discrete-time traffic simulator.[‡] The parameters we use are summarized in the Table 1. We clarify that for each value of f, as it increases, we get a different level of average car density, which we denote in our results presentation as "density level."

Table 1: Simulation parameters

Simulation duration	1000sec
Grid dimension (N)	5, 10
Number of lanes per road	2
Road segment length L	600 m
In-flow rate (f)	100..400 Vehicles/h/lane
Connectivity range r [m]	50 to 200

5.3 Results

5.3.1 Effect of the range of RSUs

The time to complete the CRL is inversely proportional to the range of RSUs (Figs. 5, 6). When the range of RSUs is large, a vehicle has a large contact time with each RSU, and thus gets a bigger chunk of the CRL from each RSU. The vehicle can then complete the entire CRL after meeting few RSUs. In the situation where there are traffic lights, larger ranges do not cause the total time to decrease further as the traffic lights already cause vehicles to spend considerable time within the range of each RSU.

[‡]SUMO implements a "car following" model and dynamically identifies routes to take based on the route length and the average achievable velocity.

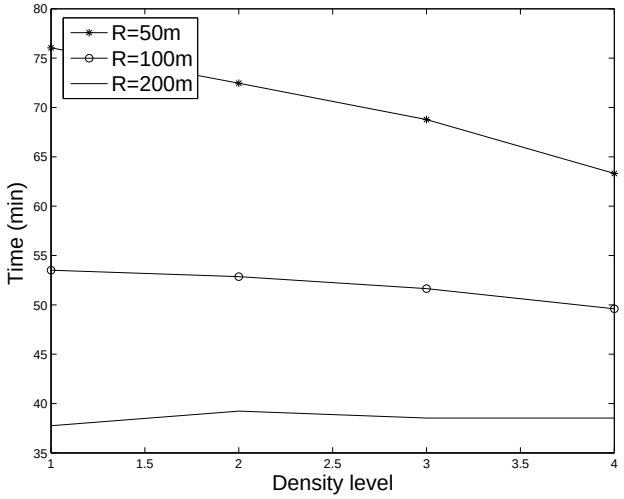

Figure 7: Time vs. density of vehicles, no traffic lights

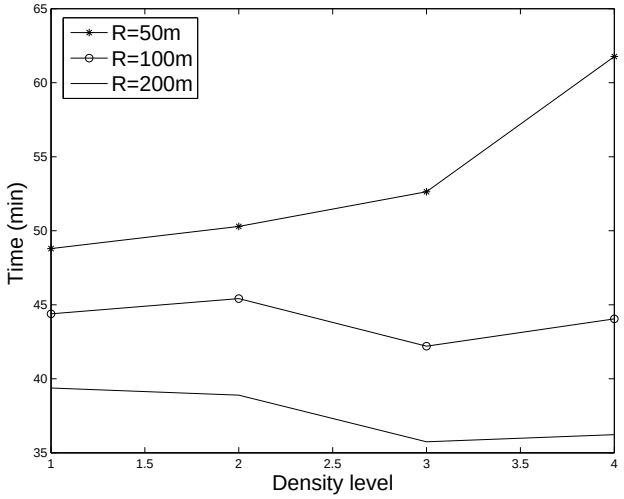

Figure 8: Time vs. density of vehicles, traffic lights

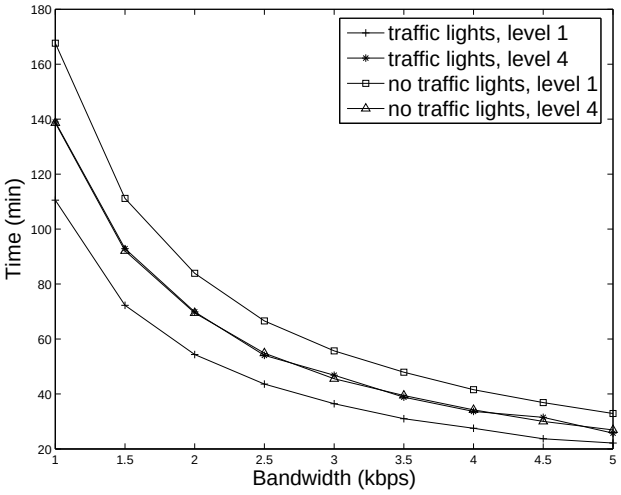

Figure 9: Time vs. broadcast bandwidth, R = 100 m

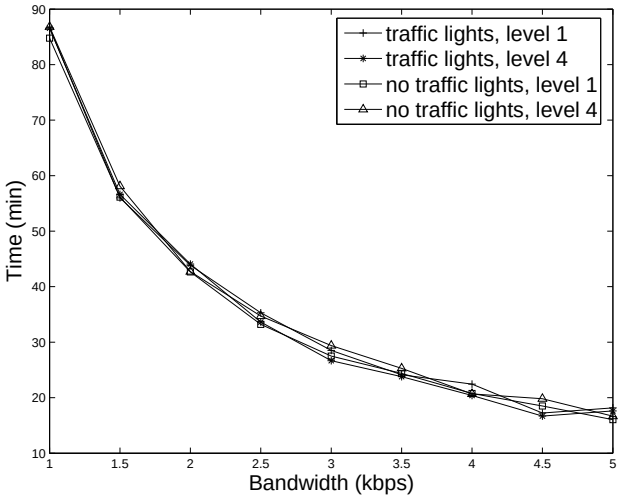

Figure 10: Time vs. broadcast bandwidth, R = 200 m

5.3.2 Effect of the distance between RSUs

The time to complete the CRL increases linearly with the distance D between RSUs (Fig. 4). As this distance increases, a vehicle needs more time until it encounters the number of RSUs required to complete the CRL. The distance D reflects the deployment density of RSUs. In the early phase of deployment, it is expected that RSUs would be sparse, and thus the distance D would be large. However, at that stage, the number of equipped vehicles would be small, and thus a small number of vehicles would need to be revoked, limiting the size of CRLs.

5.3.3 Effect of the density of vehicles

When the density of vehicles is large on the roads, congestion occurs and vehicles move slowly. This affects the total time to complete the CRL in two manners. In the first, a vehicle have a larger contact time with each RSU as it needs more time to traverse the range, and thus only needs to encounter fewer RSUs. In the second, now a vehicle needs more time to traverse the distance outside the range of RSUs, and that it takes it longer to encounter the required number of RSUs. From Figs. 7, 8), it can be seen that when the range of RSUs is large, the two effects cancel each others. However, congestion causes the total time to increase when the ranges are small.

5.3.4 Effect of the broadcast bandwidth

Allocating more bandwidth to the CRL broadcast significantly decreases the total time to complete the CRL (Figs. 3, 9, and 10). A vehicle only needs to encounter few RSUs, as it receives a big CRL chunk from each one of them. However, this effect saturates after some point, and to decrease the total time further, there is a need to change the other parameters, such as the range of RSUs, or the distance between them.

6 Related Work

The concept of revocation, especially in connection with public key cryptography, and the use of certificate revocation lists was introduced in the context of the wire-line Internet. In fact, CRLs are the standard method for revocation in the Internet [6]. A CRL is issued each revocation period, e.g., once a month. $\Delta-$CRLs were proposed to avoid large size CRLs, providing information relative to the last issued CRL. This means that a $\Delta-$CRL is useful only if the full CRL was already received. But this is not a desirable option for the error-prone, highly volatile, often disconnected VC environment.

In the context of VC, there are only a few proposals on how to revoke certificates and evict nodes. Two protocols are proposed in [11] to evict nodes: The Revocation of the Trusted Component (RTC) and the Revocation with Compressed Certificate Revocation Lists (RC^{2}RL) protocols. RTC instructs, via the RSU infrastructure, the trusted component on-board the vehicle to "self-destruct," i.e., erase its own private key. But it requires the CA, however, to be able to geographically localize any vehicle in the system, while a sophisticated adversary could always control the communication between the radio and the on-board computing platform. On the other hand, RC^{2}RL is a CRL-based revocation that compresses traditional CRLs using Bloom filters, and thus limits the size of the CRL. Since Bloom filters have false positives, some legitimate certificates that are not part of the (compressed) CRL, that is, the filter, can get revoked as well. Nonetheless, legitimate users would be unhappy to have their credentials revoked and thus have their vehicles unfairly excluded from the network. Our approach does not face the restrictions of RTC and RC^{2}RL, and address the problem of CRL distribution.

Short-lived certificates is a technique that was proposed to allow for implicit revocation. A short-lived certificate is issued to every node e.g. every day. vehicles regularly acquire proofs that their credentials remain valid. Instead of

requiring them to download revocation information, vehicles download *verifiers* from the CA. These verifiers are then included when the certificate is presented to other nodes [9]. This approach however, incurs considerable certificate issuing costs and requires vehicles to maintain a rigid contact schedule with the CA.

Beyond VC systems, the idea of revocation appeared in the context of mobile ad hoc networks [13]. However, this scheme considers the instantiation of the CA and not the problem of revocation, especially in the highly mobile VC environment. On the other hand, impromptu coalitions of network nodes (e.g., [3, 8]) cannot be applicable in VC systems either. In fact, it would be unacceptable to allow to any small subset of adversarial nodes to maliciously accuse and evict legitimate nodes.

7 Discussion and Conclusion

We first discuss the resilience of our scheme. It turns out that the simplicity of the design leaves little space for abuse, even without authentication of the RSUs. CRL pieces are protected (singed by the CA), and the adversary cannot inject any forged pieces and delay the reconstruction of the CRL, or replay pieces of an older or foreign CRL, as those would be promptly detected and ignored. On the other hand, a rogue RSU cannot abuse the system either: it can at most avoid broadcasting CRL pieces or transmit them a high rate, causing interference and eventually jamming communications within its range. But this misbehavior would be possible independently of any CRL distribution scheme.

Elaboration of the scheme for special cases of users (and thus their vehicles), such as those that frequently traverse a border across which they work, or users and thus vehicles registered with a CA with a small, geographically, region, and detailed performance evaluation of the FC functionality would be an interesting direction of future work. Another interesting topic would be to investigate whether a separation of the CRL of a CA into a small number of sub-CRLs, one for a different class of vehicles or RSUs registered with the CA, could yield performance gains.

In conclusion, we investigated the problem of CRL distribution in VC systems, which to the best of our knowledge is the first work on this topic. We show how, with very low bandwidth used for CRL transmissions, and a simple and robust design, practically all vehicles can obtain the latest CRL within a delay of 30 or 40 minutes of drive, e.g., the duration of a commute. Our analysis and simulations show the trade-offs and how the system can be configured to reduce the delivery delay if needed.

References

[1] DSRC: Dedicated short range communications. http://grouper.ieee.org/groups/scc32/dsrc/index.html.

[2] IEEE 1363a 2004. IEEE standard specifications for public-key cryptography- amendment 1: Additional techniques, 2004.

[3] C. Crépeau and C. Davis. A certificate revocation scheme for wireless ad hoc networks. In *Proceedings of SASN'03*, 2003.

[4] Michael Bonert Daniel Krajzewicz and Peter Wagner. The open source traffic simulation package sumo. 2006.

[5] M. Gerlach, A. Festag, T. Leinmller, G. Goldacker, and C. Harsch. Security architecture for vehicular communication. In *WIT 2005*, Hamburg, Germany.

[6] R. Housley, W. Polk, W. Ford, and D. Solo. Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. RFC 3280, 2002.

[7] IEEE1609.2. IEEE trial-use standard for wireless access in vehicular environments - security services for applications and management messages, July 2006.

[8] J. Kong, H. Luo, K. Xu, D. Gu, M. Gerla, and S. Lu. Adaptive security for multilevel ad hoc networks. *Wireless Communications and Mobile Computing*, 2(5):533–547, 2002.

[9] S. Micali. Efficient certificate revocation. In *MIT Laboratory for Computer Science, Tech. Rep. TM-542b, Mar. 1996.*

[10] P. Papadimitratos, L. Buttyan, J-P. Hubaux, F. Kargl, A. Kung, and M. Raya. Architecture for secure and private vehicular communications. In *ITST'07*, Sophia Antipolis, France.

[11] M. Raya, P. Papadimitratos, I. Aad, D. Jungels, and J.-P. Hubaux. Eviction of misbehaving and faulty nodes in vehicular networks. *IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on Vehicular Networks*, 2007.

[12] M. Torrent-Moreno, P. Santi, and H. Hartenstein. Distributed fair transmit power adjustment for vehicular ad hoc networks. In *SECON '06*, pages 479–488, 2006.

[13] L. Zhou and Z. Haas. Securing ad hoc networks. *IEEE Network*, 13(6):24–30, 1999.