Chapter 20

RECERTIFICATION: A TECHNIQUE TO IMPROVE SERVICES IN PKI

Ravi Mukkamala, Satyam Das and Mahantesh Halappanavar

Abstract Efficient and timely distribution of certificate revocation information is a major challenge currently facing the providers of Public-key Infrastructure (PKI). All of the current schemes, including the Certificate Revocation List (CRL) and its variants, place a considerable processing, communication, and storage overhead on the infrastructure elements (e.g., Certification Authorities (CAs) and its repositories) as well as the relying parties. In this paper, we describe schemes to improve the current situation using *recertification* concept. Here, a certificate needs to be recertified frequently after its initial issuance. As a consequence, the size of the CRLs get much shorter and subsequently it is possible to publish them more frequently. In addition, it provides opportunities to offer different types of services (with different QoS requirements) to a relying party. For example, it is possible for a relying party to completely place the burden of proof of a certificate non-revocation on the certificate-holder itself. Alternately, for high-valued transactions, it may verify itself as is done in current systems. In addition to the basic protocol, we describe an implementation scheme and the performance gains due to the recertification process. The proposed protocols work within the current PKI standards (e.g., X.509).

Keywords: Certificate freshness, certificate revocation, certificate revocation list, communication cost, public key infrastructure, quality of service, recertification

1. Introduction

With the ever-increasing growth in electronic messaging and electronic commerce, the need for an infrastructure to provide confidentiality, security, as well as confidence for such exchanges to take place is quite evident. Public-Key Infrastructure (PKI) is one such mechanism. It provides the needed infrastructure support for confidentiality and authentication through the generation, distribution, control and accounting of public key [1, 9].

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: 10.1007/978-0-387-35697-6_26

E. Gudes et al. (eds.), *Research Directions in Data and Applications Security*

When a PKI certificate is issued, it is expected to be in use for its entire validity period. However, various circumstances may cause a certificate to become invalid prior to the expiration of the validity period. Such circumstances include change of name, change of association between subject and CA (e.g., an employee terminates employment with an organization), and compromise or suspected compromise of the corresponding private key. Under such circumstances, the *CA needs to revoke the certificate* [1, 2, 4].

Several schemes have been proposed for certificate revocation and dissemination of the revocation information. The most commonly used method for dissemination of this information is the certificate revocation list (CRL)—a digitally signed list of revoked certificates. Under this scheme, each CA periodically issues a CRL [2, 4]. Several variants of CRL schemes have also been suggested [4].

However, PKI researchers and practitioners have identified several shortcomings of the CRL and its variants. First, they are expensive to distribute [4]. Second, they involve expensive storage and validation costs at the relying parties (e.g., service providers) [5, 10, 11]. Third, they provide only negative information (i.e., a certificate is not revoked) instead of positive confirmation [8, 11]. Fourth, they place a considerable burden on a relying party to verify a user's certificate [11]. Fifth, they contain substantial redundant information (e.g., consecutively published CRLs would have more than 99% of redundancy) [8].

In this paper, we overcome these shortcomings of the current revocation schemes by introducing the concept of *recertification*. While the need for such a concept has been well argued by several researchers (e.g., [11]), to our knowledge, no one has proposed efficient means to achieve it. In short, our recertification concept works as follows. First, a CA issues a digital certificate for the normal period of duration (say 1 or 2 years). Next, a recertification authority (RCA) recertifies the original certificate. The frequency of recertification depends on the uses to which a certificate is put to. A user also has an option to withhold recertification during certain period of inactivity. In addition, as discussed later, our schemes offer a variety of quality of service features to relying parties. Such features are completely absent in the current schemes. In summary, our schemes are completely novel, efficient, and flexible.

This paper is organized as follows. Section 2 discusses the concept of recertification as discussed in this paper. Section 3 describes a basic architecture for recertification. In addition, it discusses a scheme to implement recertification in the PKI context. In Section 4, we discuss the trust and security aspects of the proposed schemes. Section 5 has detailed discussions on performance implications of recertification on the PKI infrastructure, the users, and the relying parties. It considers communication cost, accuracy, and response time. The flexibility in quality of service offered to the relying parties due to our

schemes is discussed in Section 6. Finally, Section 7 summarizes the paper and discusses future plans for extension of this work.

2. Recertification: The Concept

The primary impetus for introducing the recertification concept comes from the following observations:

- *It is more efficient for a CA to issue certificates with long validity periods.* Since there is a considerable overhead involved in issuing a certificate, it is more economical to issue long-life certificates.

- *The information about a revoked certificate needs to be maintained and distributed until its expiration time.* In other words, longer the lifetime of a revoked certificate, longer is the period of maintaining its status by a CA or a repository. So a CRL, for example, keeps maintaining a revoked certificate on its list until it expires. A longer CRL is expensive (processing cost) to prepare (at CA), expensive (communication cost) to distribute to repositories, expensive (communication cost) for relying parties to copy from the repositories, and expensive (processing cost) for the relying parties to search when users submit requests.

The concept of recertification aims to combine the benefits of long-life certificates for an issuer with the benefits of short-lived certificates for revocation. The main idea is to initially issue a certificate for the normal period of duration (e.g., 1 or 2 years) and then require the certificate-holder (or user) to get the certificate recertified at certain intervals during its lifetime. A relying party not only looks for the lifetime of a certificate but also for its recertification at the time of verification. To reduce the load on the certificate issuer (e.g., CA), the recertification task is assigned to a different entity called the recertification authority (RCA). Certainly, RCA should have been delegated this authority by a CA, say by issuing an attribute certificate to this effect. As will be shown later, RCA does not have to be as trusted and secure as a CA. However, it should be certified by the CA so the relying parties can trust its actions.

In order to get an intuitive idea about the benefits of recertification toward revocation, consider a CA that revokes 10,000 certificates each week. Suppose the lifetime of each certificate is 1-year or 52-weeks. Let us assume that a certificate is equally likely to be revoked in any week during its lifetime. So the maximum number of certificate identifiers in a CRL is 52*10,000 or 520,000. The average size would be 260,000 certificates. Now, if we assume that a certificate needs to be recertified once in 4 weeks (that is 13 times during its lifetime) the maximum CRL size would only be 4*10,000 or 40,000. The average size would be 20,000 certificates. In other words, we can reduce the average CRL size from 260K to 20K. This is a significant reduction achieved

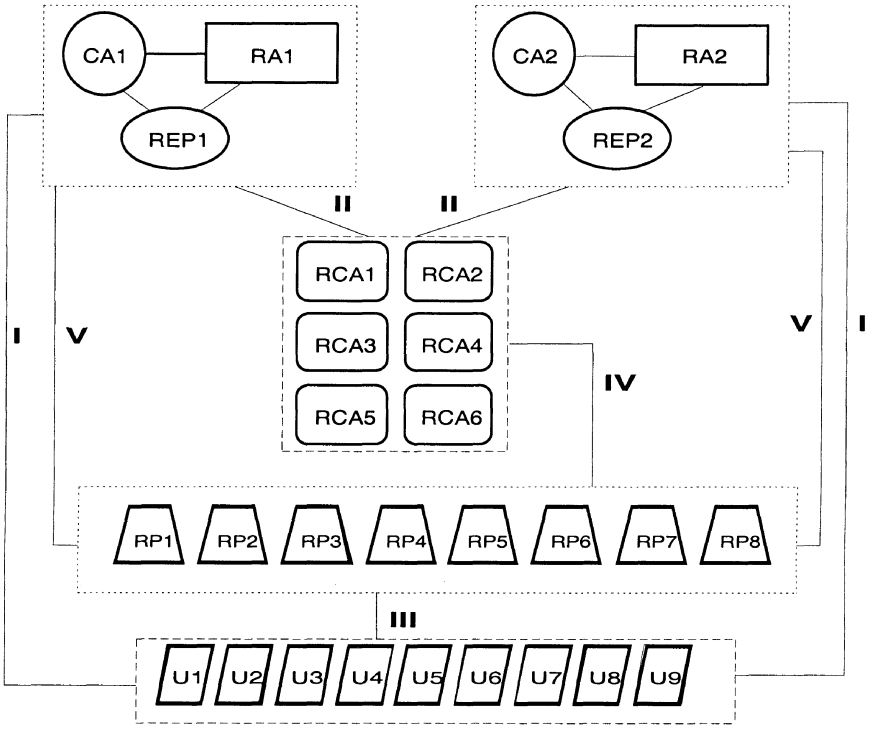

Figure 1. Proposed PKI architecture with recertification.

by requiring recertifying a certificate 13 times during its lifetime. In general, "p" periods of recertification would reduce the CRL to $1/p$ of its original size. The reduction is possible since the CRL published under this scheme contains only those that have been revoked with valid recertification dates. The additional load of recertifying each certificate p times during its lifetime is taken by an RCA with no involvement of the CA.

In summary, CA issues the certificates; RCA manages recertification and the publication of CRLs. The relying parties have the option of looking at not only the expiration time but also the status of recertification of a certificate.

3. Implementation Scheme

The basic architecture for the proposed recertification scheme is shown in Figure 1. Here, we see four major subsystems:

- *Subsystem-1*: The certification authority (CA), its registration authority (RA), and the repository (REP) together represent the traditional PKI infrastructure.

- *Subsystem-2*: The set of recertification authorities (RCA) represent the additional subsystem introduced by the proposed recertification.

- *Subsystem-3*: The relying parties (RP) are the primary users of the infrastructure for validation of certificates.

- *Substystem-4*: The users (U) or the certificate holders are contained in this subsystem. They request for issue of certificates, for recertification, for revocation, and for services (from the relying parties).

The proposed architecture can be implemented by several means. Here, we describe a scheme that uses attribute certificates. An alternate scheme requiring an RCA (instead of a CA) to sign the original digital certificate is described in [7].

3.1 Implementation using Attribute-Certificates

This scheme requires the least amount of change in the current X.509 format. The steps are as follows.

Step 1. User requests a certificate from the CA via the registration authority (RA).

Step 2. The RA verifies the user-supplied information. The CA then assigns a serial number (or certificate number), an RCA number (indicated in the certificate), etc., and digitally signs the certificate with its private key. In addition, CA also assigns a recertification frequency to the certificate for use by the RCA. The period of validity in this certificate would be the long-life as in conventional schemes. It is also important to note that by specifying the RCA's identification in the digitally signed certificate, the CA is enhancing the trust that a relying party may place on an RCA.

Step 3. CA hands over the user's request and its certificate to the assigned RCA. The RCA now creates an attribute certificate with the original certificate number and a new short validity period. The validity period of the attribute certificate would be determined by the original validity period and the recertification frequency specified in the CA's certificate. For example, if the original validity period is for a year with recertification frequency of 6, then the attribute certificate is issued for 2 months. The format of the attribute certificate is specified in X.509 standard. It is digitally signed by the RCA.

Step 4. RCA retains a copy of the original CA certificate. It sends the attribute certificate and CA's certificate to the user (as indicated by the CA).

Step 5. The user treats the CA's and RCA's certificates together as a single entity and submits the new entity to a relying party when needing a service.

Step 6. When a relying party receives a user's request with the certificate pair, it first validates the digital signatures of the original certificate and of the attribute certificate. It now has several options—from looking at the validity period in the certificate (the cheapest option) to looking at the CRL issued by a CA (the most expensive option).

Step 7. When an attribute certificate expires (or about to expire), the user sends a recertification request to the corresponding RCA. If the certificate's original validity has not expired and if it has not been revoked, the RCA would reissue the attribute certificate with a new expiration date. This is a light process since the public-key/private-key are not regenerated and the typical checks done prior to the issue of an original certificate are avoided.

Step 8. When a user needs to revoke a certificate, he notifies the CA. The CA revokes the certificate, and notifies the corresponding RCA about the revocation. The RCA also revokes the certificate.

Step 9. Both CA and RCA publish their CRLs periodically. Since RCA's CRL is short (due to the short lifetimes), it can publish it more frequently.

Clearly, this scheme has the advantage that it conforms to the X.509 certificate format without using any extensions or options. But it generates an attribute certificate for each of the original certificates. The publication of CA's CRL is purely optional and is mainly to satisfy those relying parties that trust no one else besides the CA for revocation verification.

4. Trust and Security

One of the primary security concerns in the proposed schemes is the introduction of Recertification CA (RCA). Clearly, both relying parties and users trust CAs either directly or through a chain of trust. Of course, one can argue that introducing RCA simply increases the trust chain by one more link. However, there are some major differences.

In the proosed scheme, CA alone is responsible for issuing the primary certificate. Hence, the primary certificate has the same trust as in a conventional scheme. RCA generates the attribute certificate. If an RCA was to be compromised, it can issue attribute certificates with valid periods for expired or revoked certificates. However, since an attribute certificate has no value without the primary certificate, an expired primary certificate automatically invalidates the attribute certificate. On the other hand, if CA revoked the primary

certificate and the relying party does not contact the CA, then there would be a problem of using an invalid certificate. However, we expect applications with very high validity requirements to contact the CA under this scheme.

Another danger under this scheme is the denial-of-service. If an RCA was to be hacked and it includes all its attribute certificates in its revocation list, then there would be denial of service under options (iii) and (iv). However, when such a situation is suspected by a relying party, it could contact the CA. Alternately, there could be replica RCAs. So when a relying party suspects the behavior of an RCA, it could contact a replica RCA.

5. Performance

We now look at the performance issues in the proposed recertifying schemes. Mainly we concentrate on communication cost, accuracy, and response time as the primary performance metrics.

In issuing certificates, the load on a CA remains unchanged due to recertification. The work of the registration authority (RA) is also unchanged since it is still responsible for verifying user-supplied data.

The situation is different for managing the revocation information. First, the size of the CRLs maintained by a CA remains unchanged. In other words, it does not reduce due to recertification. However, since most relying parties run applications that are satisfied with the short-lived certificates, and some others may only need the CRLs at an RCA, an insignificant portion only need access to CA's CRL. Thus, it does not have to be as widely published as before. This implies that the cost of distributing CA's CRLs to repositories is greatly reduced. It also means that only an insignificant number of relying parties get copies of the CRLs from the repositories. Thus, both the processing cost and communication cost due to the distribution and searching of CA's CRLs is greatly reduced.

However, there is an additional cost due to the CRLs at the RCAs. But, due to the short-life of the certificates handled by the RCAs, these CRLs are much shorter. In addition, most relying parties may simply rely on the validity of the short-lived certificates itself instead of verifying the RCA's CRLs.

In the following subsections, we discuss, in detail, the effect of recertification on the selected performance measures.

5.1 Communication Cost

Let us now look at different components of communication cost in this system.

If we assume that a CA revokes c certificates per year on the average, and t is the average lifetime (in years) of a certificate, then the average length of a CRL published by a CA is given by $L_c = c * t$.

If r is the frequency of renewals of a certificate, t/r is its recertification period. If a CA's certificates are partioned among mRCAs, then each RCA is assigned $1/m$ of the certificates issued by a CA. Since an RCA publishes its CRL with all recertified certificates assigned to it and revoked by CA, its length is given by L_r = *(c/m)*(t/r)*.

When we analyze (for details refer [7]) the communication costs due to the conventional scheme and the current scheme, we find that both costs are dominated by the cost of obtaining CRLs from the repositories by the relying parties. So for simplicity let us only look at this part of the communication cost. In both cases, it is $k * f * L_c * L_{ID}$ bytes. However, as noted above, the value of$f$for the recertification scheme is expected to be a fraction of that for a conventional scheme. Let it be $p * f$, $0 \leq p \leq 1$. Here, $p = 0$ implies that a relying party never needs to check with CA's CRLs and $p = 1$ means that it always needs to obtain the CA's latest CRL. So the percentage cost of savings in communication cost due to recertification is $(1 - p) * 100$.

In summary, the benefit of recertification is realized the most when a relying party does not depend on CA's CRL. There is no benefit if it always needs to check with the latest CA's CRL. The gain will also be evident if other revocation distribution schemes besides CRL are also used.

5.2 Accuracy

One of the primary concerns of relying parties in using PKI-based certificates is whether or not they are provided with accurate information. It is well known that conventional CRL-based schemes suffer from obsolescence of revocation information [6]. Let us discuss how the proposed recertification scheme helps in improving the up-to-dateness of this information.

Clearly, since the certificates are recertified frequently, a recently recertified certificate is more likely to be still active than the one issued long time ago. Thus, the short-life of the recertified certificates itself improves the confidence of the relying parties. (A relying party is more likely to trust a certificate recertified 1 hour ago than the one issued 1 year ago.)

Since RCA's CRLs are much shorter than the CA's CRLs, they will be more frequently published with very little cost. Thus, the relying parties have access to this accurate information, when needed.

Since a CA's certificates are divided among several RCAs, the length of RCA's CRL is much shorter and hence they can be frequently published and procured by relying parties.

Based on the above observations, it is easy to see that the up-to-dateness offered by the recertification scheme is much higher than that of the conventional CRL scheme without much additional cost. In fact, in incurs a smaller processing and communication costs.

5.3 Response Time

One of the other performance metrics of concern is the response time—the time when a relying party receives a user's request (along with its certificate) to the time it has verified the certificate and begins to offer the required service. So the time to verify a certificate is what is in question. Under the conventional CRL-based schemes, a relying party either has to search through the unexpired CRL it may already have or obtain a new CRL from a repository. Since the traditional CRLs are long, both steps consume considerable time. Hence, the response time is significantly high.

On the other hand, in the proposed scheme, due to frequent recertification, a relying party has several options (as discussed earlier). Accordingly, most of the time a relying party is satisfied with a certificate that has been recently recertified and needs no further validations. In a few cases, a relying party may have to search, or obtain and search, the CRL posted by the corresponding RCA. Since this CRL is much shorter, the process of obtaining such a CRL as well as searching it are significantly faster than the conventional schemes.

In summary, the proposed scheme offers much better response times to relying parties (and users) than the conventional schemes.

6. Flexibility in QoS

One of the main advantages of the proposed recertification technique is the flexibility in the offered quality of service (QoS) that may be offered to relying parties. While some of these aspects may have become apparent in the discussions of various options in the implementation, here we discuss them in more detail.

First, let us consider a conventional system that uses CRLs. Here, a relying party has two options: (i) Use the certificate expiration and other checks within the certificate itself to validate a certificate. (ii) Use the CRL provided by the CA at the directory service to confirm that a certificate has not been revoked. (Alternately, it could use OCSP-like protocol also.) Given that the lifetime of the conventional certificates is long, it is almost mandatory for a relying party to obtain the latest CRL and check for the submitted certificate. This is an expensive process.

Now, let us consider the proposed recertification scheme. Here, a relying party has four options.

- Check the original certification expiration time
- Check the short-term expiration time or the expiration of recertification)
- Check the CRL of RCA
- Check the CRL of CA

In addition to these general options, the following possibilities exist.

Degree of freshness based on relying party. A relying party can dictate the degree of freshness of a certificate. In other words, it can insist a certificate to have been renewed in the last 1-hour or in the last 24 hours. The user can go to RCA and get it renewed as per the relying party's requirements.

Renewal frequency. Depending on the type of applications for which a certificate will be used, a user can request a certain period of renewal frequency for a certificate. Thus, a certificate that is used in highly secure or high-valued transactions can have more frequent renewals or shorter lifetimes.

Need-based renewal. More importantly, a user has the option of renewing a certificate only when needed (e.g., just-in-time).

Temporary suspension. It is also possible to temporarily suspend a certificate. For example, if a manager is going to be on leave of absence for 4 weeks, his recertification which is say done everyday, could be stopped during this time. This automatically suspends the certificate without involving any other entity. We find this option to be one of the key advantages of the recertification process not offered by conventional certificates.

Many other options for flexible operation of the system are possible with the proposed recertification scheme [7].

7. Conclusions and Future Work

The cost of distributing and processing the bulky certificate revocation lists has been plaguing the current PKI systems. Especially, since digital certificates and PKI are expected to be used in almost all aspects of E-business and E-transactions in the future, the problem is expected to become worse.

In this paper, we have proposed one way to solve this problem—by using recertification. In our scheme, an original certificate is issued only once by a CA and it is then recertified by an RCA. The short-life of a certificate certainly increases the trust a relying party may have on an unexpired certificate. In addition, we provide opportunities for the revocation of certificates as in a conventional scheme. We discussed several types of qualities-of-service that may be offered to relying parties that are otherwise not possible in the conventional schemes.

We discussed a scheme to implement the recertification concept. Using these schemes, we showed how the communication cost is reduced, and the availability is increased. While the trust is improved, the security is unaffected.

In future, we propose to conduct a rigorous performance analysis of the two schemes. In addition, we will look into other ways of implementing the recertification concept. Clearly, we expect this paper to pave the way for many other researchers to follow this or similar mechanisms to improve the PKI implementations.

Acknowledgments

The work is supported in part by a grant from the Commonwealth Information Security Center (CISC) at James Madison University, Harrisonburg, Virginia.

References

[1] Adams, C. and S. Farell, IETF RFC 2510 Internet X.509 public key infrastructure certificate management protocols, www.ietf.org/rfc/rfc2510.txt, 1999.

[2] Adams, C., S. Lloyd and S. Kent, *Understanding Public-Key Infrastructure: Concepts, Standards and Deployment Considerations*. Macmillan Technical Publishing, 1999.

[3] Chadwick, D.W., An X.509 Role-based privilege management infrastructure, *Business Briefing: Global Infosecurity 2002*, www.technology-briefing.com/businessbriefing/ pdf/Infosec2002/, 2002.

[4] Denker, G., J. Millen and Y. Miyake, PKI and revocation survey, *SRI Technical Report*, SRI-CSL-2000-01, 2000.

[5] Kocher, P.C., On certificate revocation and validation, *Proceedings of the Second International Conference on Financial Cryptography (FC'98)*, Anguilla, British West Indies, *Lecture Notes in Computer Science (LNCS 1465)*, Springer-Verlag, pp. 172-177, 1998.

[6] Mukkamala, R. and S. Jajodia, A novel approach to certificate revocation management, *Proceedings of the IFIP WG 11.3 Working Conference on Database and Applications Security*, Niagara-on-the-lake, Ontario, pp. 223-238, 2001.

[7] Mukkamala, R., S.K. Das and M. Halappanavar, Recertification: A technique to improve services in public-key infrastructure, *Proceedings of the IFIP WG 11.3 Working Conference on Database and Applications Security*, King's College, Cambridge, England, pp. 277-294, 2002.

[8] Myers, M., Revocation: Options and challenges, *Proceedings of the Second International Conference on Financial Cryptography (FC'98)*, Anguilla, British West Indies, *Lecture Notes in Computer Science (LNCS 1465)*, Springer-Verlag, pp. 165-171, 1998.

[9] Nash, A., W. Duane, C. Joseph and D. Brink, *PKI: Implementing and Managing Security*, Osborne/Mc-Graw Hill, 2001.

[10] Naur, M. and K. Nissim, Certificate revocation and certificate update, *IEEE Journal on Selected Areas in Communications*, 18(4), pp. 561-570, 2000.

[11] Rivest, R.L., Can we eliminate certificate revocation lists? *Proceedings of the Second International Conference on Financial Cryptography (FC'98)*, Anguilla, British West Indies, *Lecture Notes in Computer Science (LNCS 1465)*, Springer-Verlag, pp. 178-183, 1998.