INTERNET SECURITY: PUBLIC-KEY INFRASTRACTURES AND CERTIFICATION SYSTEMS

Ben Soh, Luke Sledziona
Department of Computer Science & Computer Enguneering, La Trobe University, Bundoora, VIC, Australia 3083

Keywords: Public-Key Infrastructures, Certification

Abstract: In the current business environment there is an ever growing view of the World Wide Web, commonly referred to as the Internet, as the new frontier for electronic commerce or e-commerce. As a result many businesses are developing applications and/or websites in order to conduct e-commerce on the Internet without properly considering the implications of the certification system that they are choosing to use, if they use one at all. The aim of this paper is to present work in the area of public key infrastructures and certification systems by discussing important topics pertaining to this area of research. The security needs of businesses will be initially discussed as an introduction to certification systems. This leads into the discussions of X.509 public key infrastructures and certificate revocation, where the associated problems will be discussed.

1 INTRODUCTION

In the current environment of electronic commerce, security is becoming an increasing concern for customers. As a result a properly implemented security system for a business can mean the difference between the success and failure of their products. A public key infrastructure (PKI) can assist businesses in conducting secure electronic commerce as well as securing their important customer information.

This paper attempts to provide discussions on public key infrastructures (PKIs) and certification systems. Central to this topic are the security needs of businesses, these are discussed in Section 2. Sections 3 and 4 introduce X.509 PKIs and certificate revocation, the two key areas to enabling a business to conduct effective e-commerce. Section 5 highlights some of the potential social and technical problems of X.509 PKIs.

2 BUSINESS SECURITY NEEDS

In the realm of online business, security infrastructures are no longer an option, they are a necessity. The author of [1] describes security in general as requiring three main approaches. These are enablement, intrusion detection and response, and perimeter control. PKIs and certification systems are classified as enablement, which is defined as implementing a security plan and having the infrastructure to support it. The business requirements for a security infrastructure consist of five aspects [1]: entity authentication, data confidentiality, data integrity, non-repudiation, and privilege management.

These business requirements should be kept in mind when considering any perspective security infrastructure. If a security infrastructure fails to meet any of these requirements it can not be considered an appropriate infrastructure for use with e-commerce on the Internet. The most prevalent types of security infrastructure on the Internet that meet these requirements are X.509 public key infrastructures.

3 X.509 PUBLIC KEY INFRASTRUCTURES

One of the proposals to the problem of business security needs was introduced by Whitfield Diffie and Martin Hellman in their paper “New Directions

Soh B. and Sledziona L. (2004).
INTERNET SECURITY: PUBLIC-KEY INFRASTRACTURES AND CERTIFICATION SYSTEMS.
In *Proceedings of the First International Conference on E-Business and Telecommunication Networks*, pages 93-98
DOI: 10.5220/0001387900930098

in Cryptography" in 1976 [2]. This infrastructure was designed to enable secure, convenient, and efficient discovery of public keys. This architecture also provided additional functionality in the form of digital certification which was designed to assure that communication had taken place and had not been altered in any way. There are various models of PKIs each differing in information required, trust rules, and flexibility. The X.509 model [3], developed by the ITU Telecommunication Standardization Sector (ITU-T), will be the main focus of this paper. In order to fully understand the context of PKIs, their uses, operations, components, architectures, and responsibilities must first be investigated.

3.1 Addressing Business Security Needs

The primary applications of PKIs are electronic commerce and electronic service delivery [4]. These applications exist in the environments of business-to-business (B2B), single business (B), business-to-customer (B2C), and individual (I) [5]. PKIs address all business requirements, as listed previously, required from a security infrastructure in the following manner [1].

3.2 Cryptographic Techniques Used

There are two forms of cryptography widely used on the Internet, secret key "symmetric" and public key "asymmetric". SKC involves using a single key to encrypt and decrypt data as opposed to public key cryptography (PKC) which uses a pair of keys, one public and one private. PKC is the more flexible of the two forms since an entity only needs one key pair to communicate with everybody as opposed to SKC where an entity must have a unique key for each other entity it wishes to communicate with. X.509 PKIs utilise the best of both of the above mentioned cryptographic forms. When used in PKIs, PKC is used for digital signatures, for example if a message is encrypted with an entity's private key any other entity can decrypt it with the corresponding public key. Likewise, when used in PKIs SKC is used for the encryption of messages, for example if a message is encrypted with an entity's public key only the entity owning the corresponding private key can decrypt it. Furthermore, X.509 PKIs are capable of using two authentication methods, simple authentication via passwords and strong authentication via cryptography techniques. It is strong authentication this paper will focus on.

3.3 Responsibilities

Certificate authorities (CAs) are the main component of PKIs. As such, they are responsible for the services they provide as well as the quality of the services they provide to entities. Entities expect CAs to be reliable, have integrity and be liable for any unauthorised misuse of any of their products. However, legally CAs are not liable for the misuse of their products and only have to adhere to two criteria. These criteria involve proving an entity's public key has a working private key counterpart and that an entity's distinguishing name (DN) is unique to that CA [11]. Functionally, however, CAs need to handle additional management responsibilities such as registration, initialisation, certification, key pair recovery, key pair updates, revocation requests, and cross-certification [9].

4 CERTIFICATE REVOCATION

Certificate revocation is a very important issue for PKIs due to certificate's secure nature and their use for identifying entities. As a result certificate revocation needs to be clearly defined, fast, efficient, and timely. In general, if a CA wishes to revoke a certificate it sends a revocation notice to a key server which updates a CRL (Certificate Revocation List). It is then the distribution of this revocation information that has the potential to be the most costly part of running a PKI [12]. Reasons for revoking certificates can include: key compromise, change of affiliation, superseded information, cessation of operation, algorithm compromise, revocation of superordinate certificate, lost or defective security token, change of key usage, or change of security policy [13].

4.1 Popular Methods

There have been many methods for certificate revocation that have been proposed. The four most popular methods are CRLs and Delta-CRLs, online certificate status protocol (OCSP), the certificate revocation system (CRS), and certificate revocation trees (CRTs). These methods can be classified by certain attributes such as their method of checking (online/offline), the type of lists they use (black/white), their way of providing evidence (direct/indirect), and their way of distributing information (push/pull mechanism) [13].

The most commonly used certificate revocation method is that of the CRL which was introduced in 1988 by the ITU-T. The certificate revocation list is

a file that contains a list of all the invalid certificates that a CA has previously issued. As part of the authentication process this file is checked prior to the confirmation of any certificate to ensure that the certificate that is being validated is indeed valid.

The advantages of CRLs are that they are straightforward and easy to understand whilst the disadvantages are that they can grow very large and take a long time to transmit if validity periods are long. Delta certificate revocation lists are the same as CRLs but are used between CRL releases and contain only the updates since the last CRL release. Their purpose is to provide updated information to entities without them having to download a completely new CRL. CRLs are classified as checking offline, using black lists, and providing indirect evidence.

OCSP was developed by the Internet Engineering Task Force (IETF) and works differently to the traditional method of checking certificate validity though the use of CRLs. OCSP works by checking the validity of the receiver's certificate at the time of sending instead of at the time of receiving. However, this introduces a problem in that if the certificate of the receiver is revoked whilst the message is in transit they can still read the message once it has been received [14]. The main advantage of OCSP is that it provides more timely status information than any other revocation method. OCSP is classified as using online checking and using black lists.

4.2 Reduced Request Rate Revocation

One of the responsibilities of CAs is to periodically issue CRLs to a repository. Each CRL that is issued by a CA contains a next update field with the date and time of when the following CRL will be released. This information is then used by entities in determining when they will next update their CRL information. Requests for new CRLs are typically made by entities the first time they wish to verify another entity after their current CRL has expired. In the traditional method all CRLs expire at the same time resulting in a very high peak rate of accesses to CRL repositories [12]. After this initial peak period there is an exponential decline in the requests for CRLs and after a longer yet period of time the repository will become under utilised. Ideally, requests for CRLs should be spread out over an extended period of time instead of producing peak loads after a new release.

One method of reducing peak loads can be achieved by making CRLs expire at different times. This process which involves issuing a new CRL before the previous one has expired is known as over-issuing and results in reduced peak request rates and greater repository utilisation. Another method of reducing peak loads can be achieved by splitting CRLs into segments to reduce their size. However, whilst this does not reduce the peak request rate it will allow the repository to service requests quicker, this is known as segmented CRLs. Segmented CRLs can be combined with over-issuing, but as segmentation increases the benefits resulting from over-issuing decrease.

When attempting to reduce the request rate for CRLs from a repository the validity period of CRLs needs to be taken into consideration. If CRLs are going to be valid for very short periods of time then the best method to implement is segmented CRLs. Conversely, for CRLs that are going to be valid for a long period of time the best method to implement depends on the expected number of revoked certificates. If very few certificates are expected to be revoked then over issuing should be used, whilst if many certificates are expected to be revoked then segmented CRLs should be used. Finally, if the entity operates offline then over-issued should be used regardless of other factors [12].

4.3 Fast Revocation and Security Capabilities

The biggest problem with current revocation methods is that they are not fast revocation methods. If a certificate is revoked it can take anywhere from one week to a month before it appears in a CRL. As a result this can cause a serious breach of security in organisations where certificates are linked to access privileges. The use of a security mediator (SEM) in conjunction with a variant of the threshold based Rivest-Shamir-Adleman (RSA) cryptosystem [16], known as mediated RSA, offers a number of practical advantages. These include simplifying validation of digital signatures, enabling certificate revocation with legacy systems, and providing immediate revocation capabilities [14].

The SEM architecture is implemented using a background process running on a server and involves splitting an RSA private key into two parts and giving half to the entity and half to the SEM. In order for an entity to sign or decrypt messages a message specific token must first be obtained from the SEM. Without the token a user's private key can not be used. For example, for a user to decrypt an email message they need to send the message header to the SEM, whilst to sign an email message they need to send a hash of the message to the SEM. Each time the SEM will check its list of revoked users and respond with a security token if the user is still valid.

With the SEM architecture there is no need to check the validity of certificates. Once an entity's certificate is revoked the SEM no longer allows the generation of valid signatures. Hence, the existence of a signature verifies the certificate was valid at the time of signing. Also the SEM uses CA based key generation with key escrow. Meaning for example, if a person is fired from a company the company can access the person's files by obtaining their private key from the CA that originally issued it.

Whilst using a SEM may appear to be the silver bullet of certificate revocation, replicating the SEM makes it easier to expose the SEM's key to malicious entities. If this key was to be compromised by an attacker they could un-revoke revoked certificates or block valid ones. As a security precaution against this the SEM architecture is recommended only for medium sized organisations and not for wide scale Internet usage.

5 X.509 PUBLIC KEY INFRASTRUCTURE PROBLEMS CERTIFICATE REVOCATION

The primary concern for businesses involved in e-commerce is the potential loss of assets due to security breaches of transactions and computer systems. E-commerce is of a digital nature and as such is open to substitution, modification, and replication. Most security issues can be classified into one of three categories, social, technical, and appropriate usage. With social issues many of the problems stem from a lack of understanding of the technology, while with technical issues, problems stem from technical flaws with the technology. With appropriate usage issues, problems stem from the incorrect usage or implementation of the technology.

5.1 Social

Many problems that are normally attributed to PKIs are social problems that are caused by users. Social problems mainly consist of privacy, social engineering, and trust issues. Likewise, risks associated with usage of PKIs by people can include but are not limited to: sabotage, vandalism, loss of data integrity, theft, fraud, and breaches of privacy [1]. Many experts see privacy as being the major hurdle to the public accepting PKI technology in earnest. If people see something as intrusive or invasive into their personal lives it will decrease their confidence in the product. However, these people are often the most susceptible to social engineering, which is the use of psychological tricks by a person in order to gain confidential information. Social engineering can take many forms but predominately involves a person convincing a receiver that a fake key is a real key of another user, in this scenario the receiver will be fooled into thinking they are receiving authentic documents. Likewise, a person can convince a sender that a fake key is the public key of another user, in this scenario that user can then intercept communications intended for the receiver.

There are social problems associated with the PKI architectures that were previously discussed in this paper as well. With the single CA model there is no global organisation that is completely trusted by all countries, education institutions, or businesses to undertake this role. This model also presents problems for people in obtaining certificates since it would be inconvenient, insecure, and expensive to obtain a certificate from a distant organisation. Also if a single organisation were to control the issuing of all certificates they would have a monopoly on the market and could charge exorbitant prices for their services. The oligarchy of CAs model is no better since it suffers from the same accessibility problems as the single CA model. In the anarchy model however the main problem is one of trust. A user can trust another user but they do not know how trustworthy the users that they trust are.

There is also a lack of security awareness in users of PKI technology. Many people believe that PKI technology is the silver bullet of Internet security and access to it results in the entire Internet being secured. However, most electronic communications are not private or secured unless explicitly stated. Users are often quiet unaware that their practices can cause security problems. For example, private keys are normally stored on personal and sometimes public PCs that are susceptible to attack by viruses and other malicious code. To confuse the user even more most CAs have CPSs that are very hard to understand and often state in fineprint that they have zero liabilities for the damage that their product could cause in the hands of a malicious user.

5.2 Technical

Even though PKIs have some social problems they are normally not of a serious concern to an organisation investing in the technology. This is due to the fact that they can normally be avoided with prior training of the potential users. The real worry to the potential investor is if the technology is technically sound itself. Papers produced by the academic community present a multitude of problems with the technical aspects of PKIs and its

use of PKC ranging from very minor problems to quiet serious problems. In regards to the responsibilities of CAs there are problems that exist in certificate acquisition, recognition, revocation, distribution, re-distribution, validation, key-binding to an identifier, and key-attribution to a real-world entity [18].

Some of the technical problems stem from the use of PKC by PKIs. The main problems with PKC is that it is vulnerable to "man-in-the-middle" and "chosen ciphertext" attacks which need to be specifically countered due to PKCs use of public exponent two [19]. Another potential problem is the fact that PKC relies on simple mathematics using very large numbers. A breakthrough in computational number theory could make all keys easily breakable overnight [19]. Two further criticisms of PKC are that it requires computational intensive algorithms and that it has a default password for private keys when they are first issued. The intensive algorithms are no longer an issue with the computational power of today's computers, but the default password for private keys possesses a security risk if the private key was sent to the customer electronically via email instead of sent to them inside of a physical security device.

There are technical problems with some of the proposed architectures of PKIs as well. In the single CA, single CA plus RA, oligarchy of CAs, and configured plus delegated CAs models if a CA is compromised then there is the opportunity to revoke that CA and all the certificates that it has previously issued. The anarchy model however suffers from a unique problem in that as the number of entities increases the number of certificates stored increases exponentially. The flexible bottom up model also has technical problems in the form of which names are permitted. If all names are permitted it can form a similar structure to that of the anarchy model and in turn can result in an unscalable structure which defeats the advantage of using the flexible bottom up model.

The author of [20] lists some of the technical risks that are involved in the adoption of PKIs. A risk is involved if certificate verification uses one or more root public keys. This is due to the fact that if a malicious user can add their key to the list of valid keys they can issue certificates that will be treated as legitimate certificates. Another risk is involved in models that have RAs that are separated from CAs such as the single CA plus RA model. These models can be considered less secure because the CA has no idea what they are signing and relies on the word of the RA that all the information it is receiving has been previously validated. The last technical risk that the author mentions is in regards to how certificates associate public keys with a DN. Technically DNs must be unique to a CA but do not need to be unique to all CAs. As a result there could be many entities using the Internet with certificates that contain the same DN.

In regards to technical problems of certificate revocation traditional techniques do not provide immediate revocation. As such it can take a considerable amount of time before a revoked certificate appears in a CRL. In addition whilst CAs do produce CRLs and distribute them they are often done so via insecure online protocols. Legacy systems also present a problem since they do not have the means for checking CRLs when determining the validity of certificates [14], for example Netscape 3.0.

5.3 Appropriate Usage

Problems with PKIs can also stem from the improper usage and control of the technology. Some models of PKIs such as the oligarchy of CAs model hardcode their public keys into client software, for example Internet browsers. This results in the user being forced to accept these certificates even if they do not trust them. As a result, Internet browsers automatically trust certificates generated by CAs that they have hard-coded public keys for, all without the user's permission. In regards to Internet email clients they store certificates on a user's computer if they are received with an email message regardless of their validity or if the user deletes the message. A certificate that is stored from a deleted email becomes known as a virgin birth certificate and remains active on the user's PC until it expires [18]. The most blatant inappropriate control of the technology however has been suggested by some governments that wish to force the implementation of weak encryption and key escrow to provide a back door into certificates in order to make decryption a relatively easy matter. This action can not be allowed to occur since it places a serious security hole in the infrastructure.

6 CONCLUSION

The current interest in electronic commerce has resulted in a need for a security infrastructure to allow for secured trading on the Internet. Public key infrastructures have been theoretically available since 1976 but have only seen wide spread implementation and acceptance in the last five years. The author of [5] claims this is due to the significant amount of research that has been done in the areas of architecture models, certificate revocations methods, standards and compatibility issues thus making it the

right solution for any environment. Overall, PKIs fulfil all security requirements required by businesses for conducting electronic commerce, and enable businesses, organisations and individuals to confidently conduct commerce on the Internet.

The social problems listed in regards to PKIs can be solved through the use of user training and the appropriate usage problems are either controlled by software development companies, for example Microsoft, or can be solved through the implementation of policies at the CA level to prevent weakening of the infrastructure. This leaves technical problems of PKIs to be addressed.

REFERENCES

[1] Lareau P., Schulz B., "PKI Basics – A Business Perspective," PKI forum, PKI Note, April 2002, http://www.pkiforum.org/pdfs/PKI_Basics-A_business_perspective.pdf.

[2] Diffie W., Hellman M., "New Directions in Cryptography", IEEE Transactions on Information Theory 22, 1974, pp 644-654.

[3] ITU-T X.509 Recommendation, "Information Technology – Open Systems Interconnection – The Directory Public Key and Attribute Certificate Frameworks", ISO/IEC 9594-8 2000, June 2000

[4] Clarke R., "Public Key Infrastructure Position Statement," Australian National University, Department of Computer Science, Canberra, Australia, May 1998, http://www.anu.edu.au/people/Roger.Clarke/DV/PKIPosn.html

[5] Adams C. et all., "Which PKI (public key infrastructure) is the right one?," in Proceedings of the 7th ACM conference on Computer and communications security, Athens, Greece, 2000, pp 98-101.

[6] Adams C., Lloyd S., "Profiles and Protocols for the Internet Public-Key Infrastructure," in Proceedings of the Sixth IEEE Computer Society Workshop on Future Trends, October 1997, pp 220-224.