Seminar

Database and Information Systems

Security Primitives and Methods

Yingyan Zhang
zhangyin@rhrk.uni-kl.de
Angewandte Informatik

30.06.2004

Betreuer: Boris Stumm

Contents

1 Introduction ……. 3
1.1 Definition …….. 3
1.2 Example ……. 4
1.3 Overview …….. 4

2 Information Security Primitives …… 5
2.1 Confidentiality ……..... 5
2.2 Integrity ……. 5
2.3 Authentication ……... 5
2.4 Non-repudiation …… 6
2.5 Availability ……... 6
2.6 Accountability …….. 6

3 Information Security Methods…..…….. 7
3.1 Symmetric Encryption (Private Key Encryption) …… 7
3.1.1 Stream Cipher: Vigenère ……. 8
3.1.2 Block Cipher: Playfair ……..……... 8
3.1.3 Block Cipher: AES ……..…… 9
3.1.4 Limits of the Symmetric Encryption ……... 10
3.2 Asymmetric Encryption (Public Key Encryption)…… 11
3.2.1 RSA …….. 11
3.2.2 ElGamal …… 12
3.2.3 Concerns of the Asymmetric Encryption ……. 13
3.3 One-Way Hash Function ……..…….. 13
3.4 Digital Signature …… 14

4. Information Security Protocols ……..……...…….. 16
4.1. Kerberos ……..…….. 16
4.2. PGP: Pretty Good Privacy ……..…… 17
4.3. VPN: Virtual Private Network …… 17

5. Conclusion ……..… 19

A. Bibliography ……..….. 20

1. Introduction

Although officially the title of this article is Security Primitives and Methods, considering the name of this seminar is: Database and Information Systems, I would like to confine the concept "security" to "information systems security", which fits better in this context. First of all, I would like to introduce some related definitions.

1.1 Definition

Security is a widely spread concept, which covers different fields in our life, such as finance, private life, national defence, telecommunication, etc. According to the definition given by the internet encyclopaedia [1]:

> ***Security*** *is being free from danger. In absolute sense this is hardly possible, it is a relative matter. The term can be used with reference to crime, accidents of all kinds, etc.*

However, such a general concept is irrelevant to this article. In this context I will focus on the introduction of information security and the techniques we implement to keep information secure.

> ***Information Security*** *deals with several different "trust" aspects of information. It is not only confined to computer systems, nor to information in an electronic or machine readable form. It applies to all aspects of safeguarding or protecting information or data, in whatever form* [2].

The unique nature of information forces information security to take many different forms of security such as physical (hardware) security, communications (routing) security, emission (wireless signal) security, computer security and network (switches) security, into account so as to offer the most complete protection of data and resources. Only when the whole information system[1] is secure, information security can be achieved.

Based on information security, The U.S. National Information Systems Security Glossary [3] defines information systems security as:

> ***Information Systems Security*** ***(INFOSEC)*** *is the protection of information systems against unauthorized access to or modification of information, whether in storage, processing or transit, and against the denial of service to authorized users or the provision of service to unauthorized users, including those measures necessary to detect, document, and counter such threats.*

Once the data is created, INFOSEC starts to work during data storage, data process and data transition until the data is deleted at the end. During the whole lifecycle of information, INFOSEC not only prevents information from being read or revised by "illegal" users, who do not have the authorization to the data but also ensures the availability of the information to "legal" users, who have the correct authorization. In addition, INFOSEC takes necessary measures to spot the potential threats, keeps record of them and if needed, deals with them in order to maintain the security of the information.

Establishment of reliable systems by implementing advanced algorithms and sophisticated protocols, which help to protect information, is nowadays one of the main concerns for economical facilities, political

[1] Any telecommunications and/or computer related equipment or interconnected system or subsystems of equipment that is used in the acquisition, storage, manipulation, management, movement, control, display, switching, interchange, transmission, or reception of voice and/or data, and includes software, firmware, and hardware.

organizations, social public infrastructures and various communication partners, etc. I would like to take a quick look at a concrete example so as to have a better understanding of the application and importance of information security in our daily life.

1.2 Example

Bank is an important economic mechanism, which provides services like wire transfer, checks or credit cards payment and online transactions. All these make our life much more convenient and comfortable: Companies pay the salary to their employees through bank transfer, simply from company account to employees' personal accounts; People do not need to exchange the currency or bring a lot of cash when travelling to some foreign country, all the expenses can be settled by the travel checks; It is no longer necessary for crippled people to go to the bank just to arrange their personal finance since with an internet connection accounts can be managed online... However, without the assurance of information security, the above mentioned conveniences would turn out to be terrible catastrophes and bring us a lot of trouble.

Over the last few years, many banks have acquired an internet presence, with a website and facilities for customers to manage their accounts online. They also issue credit cards that customer use to shop online, and they acquire the resulting transactions from merchants. Let us have a close look at the normal procedure for executing a deal through internet.

- Open the website of the right bank;
- Input the username and password as requested;
- If the password is proved to be correct, account operations will be conducted;
- If the password is wrong, the transaction will be cancelled.

Behind the scenes after receiving customer's username and password, the bank information security system authenticates transactions based on the username and password, in such a way as to defend against both outside (hacker) and inside (bank staffs) attack. This seemingly simple process actually requires sophisticated information security technology. Between the customer's computer and the bank server, there are a lot of stations, which transfer the authentication and transaction information between them. If the authentication process is easy to be cracked, the hacker is able to obtain the important data of the customer from the media and later simulate a transaction to get enormous amount of money illegally. So, it is very important to keep the transaction information secret. Actually not only bank but also those individuals, enterprises and organizations, who make frequent correspondences with their partners, would like to keep the content of their communication private and secret. As we can see, the security of information is a top priority to any systems.

1.3 Overview

What kind of information system is secure? Basically a secure information system should ensure the confidentiality, integrity, availability, authentication, non-repudiaton and accountability of the information it stores, processes or transacts. They are the cornerstones of INFOSEC. This article will help to understand these attributes and explain miscellaneous methods developed on the base of cryptography[2] to realize them.

[2] The study of ways to convert information from its normal, comprehensible form into an obscured guise. The information becomes unreadable if without special knowledge.

2. Information Security Primitives

The rapid development of electronic data processing led to the creation of large computer databases that were routinely communicated or made remotely accessible via resource-sharing networks. Public and private sectors alike soon recognized the need to safeguard both the privacy and the integrity of the data. Generally speaking, a secure information system should keep the confidentiality, integrity, authentication, non-repudiation, availability and accountability of the information.

2.1 Confidentiality

Confidentiality has been defined by the International Standard Organization (ISO) as 'ensuring that information is accessible only to those authorized to have access.' Confidentiality is one of the design goals for many crypto systems, made possible in practice by the techniques of modern cryptography [4].

For example: Suppose Alice would like to draw 500€ from a bank through the Internet. Earlier she has already set up an account at the bank and also got the corresponding username and password for online transactions. During the process, a hacker taps in the information transmission and copies the login and account data from Alice. He is now able to do anything to Alice's account as he wishes. But what if all the information is encrypted? The hacker succeeds in obtaining the necessary information but can not make use of it. Alice's account is safe though the information leaks out. Therefore, cryptography is widely accepted to ensure the confidentiality of the information.

2.2 Integrity

The word "integrity" stems from Latin adjective integer, which means whole, complete. In this context, the integrity of the information implies the wholeness and entireness of data, that is:

- Data is not change since it is created and has not been accidentally or intentionally revised or destroyed
- Data remains always the same during any data processing like transference, storage and retrieval.

2.3 Authentication

Authentication is the process by which a computer, computer program, or another user attempts to confirm that the computer, computer program, or user from whom the second party has received some communication is, or is not, the claimed first party [5].

A similar but not the same concept is authorization. Some consider authorization is identical to authentication, however, they are different. Authorization grants the access without proving identity, like keys and tickets, while authentication requires proof of identity.

Access control is an example: A computer system supposed to be used only by those authorized must attempt to detect and exclude the unauthorized. In our university, those who participate in a practical course will get a

username and password after registration. With this username and password, students are able to log into a certain server and get resources only open to them. Access to these resources is usually controlled by going through an authentication procedure (in this case is: type in the username and password) before access is granted.

The classic methods by which a human can authenticate himself are summarised into three categories:

- Something he is: fingerprint, DNA sequence, voice pattern...
- Something he has: Student card, ID card...
- Something he knows: password, expiration date on credit card...

2.4 Non-Repudiation

Non-repudiation is the concept of ensuring that a contract, especially one agreed to via the Internet, cannot later be denied by one of the parties involved. In today's global economy, where face-to-face agreements are often not possible, non-repudiation is becoming extremely important to commerce. Non-repudiation with proof of origin provides the recipient of data with evidence that proves the origin of the data [6]. Non-repudiation with proof of receipt provides the originator of data with evidence that proves the data was received as addressed. Non-repudiation is consequently an essential element of trust in e-business [7].

For an e-Commerce deal, the two parties of a transaction must be confident that the transaction is secure, that means, first of all, that the parties are who they say they are (authentication); if the two parties do not want to make the deal public it is kept private and secret (confidentiality); the transcation should be executed exactly without wrong message (integrity); when the transaction is verified as final, systems must ensure that a party cannot subsequently deny a transaction (non-repudiation).

To protect and ensure non-repudiation, the systems may employ Digital Signatures, which will not only validate the sender, but will also 'time stamp' the transaction, so later it cannot be claimed that the transaction was not authorised or not valid etc.

2.5 Availability

Availability allows information to be functional to the end user. This is done by ensuring the information will be available whenever it is needed. Availability counteracts DoS[3] attacks, meaning that systems are always up and the communication channels are always clear [8].

2.6 Accountability

Accountability does not stop attacks by itself. Instead it is used in combination with the other primitives to make them more efficient. In particular, confidentiality and integrity efforts would fail if not for accountability. Simply put, accountability adds responsibility and redundancy of certain duties, actions, and processes into the planning and implementation of security policies [8].

[3] DoS attacks occur when information, applications, or services cannot be accessed when they are needed.

3. Information Security Methods

In the last chapter, I have introduced the definitions of information security primitives. How to implement them in reality? What kinds of techniques are widely used? What are the advantages and disadvantages of these methods? This is the main content I will deal with in this chapter.

As mentioned before, most methods to ensure information system security are developed with cryptography. The word cryptography comes from the Greek words kryptos (hidden or secret) and graphos (writing). The basic function of cryptography is to prevent the information between two parties from being read by a third party.

One kind of cryptography works on characters by shuffling and reordering the characters according to a certain rule. Some classic methods like Caesar Encryption vary the order of the text so that other people can not understand the text; another kind of cryptography bases on representing information as numbers and mathematically manipulating those numbers. Some advanced cryptography methods like AES, DES belong to the second category. Cryptography working on numbers is more flexible and advanced. Besides confidentiality, it can provide other services, such as integrity checking and authentication.

In cryptography, a normal message is called plaintext. When cryptography is applied to this message, the transformed message is known as ciphertext. The process for producing ciphertext from plaintext is defined as encryption. The reverse of encryption is called decryption. Encryption and decryption can not be implemented without key[4]. The graph illustrates briefly the whole process of cryptography:

3.1 Symmetric Encryption (Private Key Encryption)

Symmetric Encryption involves the use of a single key. Encryption and decryption use the same key.

Symmetric encryption has two fundamental classes: One is the stream cipher; the other is the block cipher. The former makes the encryption rule depend on a plaintext symbol's position in the stream of plaintext symbols, while the latter encrypts several plaintext symbols at once in a block.

[4] A key is a relatively small amount of information that is used by an algorithm to customize the transformation of plaintext into ciphertext (during encryption) or vice versa (during decryption).

3.1.1 Stream Cipher: Vigenère

The Vigenère is an early stream cipher, which was developed by the Frenchman Blaise de Vegenère, a diplomat who served King Charles IX.

The Vigenère works by adding a key repeatedly into the plaintext using the convention that A = 0, B = 1, … , Z = 25; and addition is carried out modulo 26, that is, if the result is greater than 25, we subtract as many multiples of 26 as are needed to bring us into the range [0, … .25], corresponding to [A, … , Z]. The formula is mathematically represented as:

Ciphertext = Plaintext + Key modulo 26

Because the key is possibly composed of more than one character (corresponds to a number), the same letter in plaintext may appear in the cipher first time as B, second time as C. Therefore, an eavesdropper can not guess the vowels or the frequently used words through the number of its appearance in the cipher. The longer the keys are, the more difficult it is to decipher the message.

However, this method is not absolutely secure. Given a long enough piece of the ciphertext, repeated patterns will appear at multiples of the keyword length. Here is an example [9]:

Plain:	tob eor not tob eth ati sth equ est ion
Key:	run run run run run run run run run run
Cipher:	kio vie eig kio vnu rnv jnu vkh vmg zia

The word “kiov” repeats after 9 letters and “nu” emerges again after 6 letters. A reasonable guess is: the key may have 3 letters. It follows that ciphertext letters one, four, seven, and so on all enciphered under the same key letter; so frequency analysis techniques can be applied to guess the most likely values of this letter, then repeat the process for the second and third letters of the key.

3.1.2 Block Cipher: Playfair

Besides stream cipher, block cipher is the other main kind of symmetric encryption. In block cipher the plaintext is first divided into blocks then encrypted, it is relatively difficult for intruders to get the plaintext through the frequency of the appearance of letters. One of the best known block ciphers is the Playfair system, which was invented in 1854 by Sir Charles Wheastone.

Playfair uses a 5 by 5 grid, in which the alphabet is placed, permuted by the keyword, and omitting the letter *J*. The plaintext is first conditioned by replacing *J* with *I* wherever it occurs, then dividing it into letter pairs, preventing double letters occurring in a pair by separating them with an “x”, and finally adding a “z” if necessary to complete the last letter pair. It is then enciphered two letters at a time using the following rules:

P	A	L	M	E
R	S	T	O	N
B	C	D	F	G
H	I	K	Q	U
V	W	X	Y	Z

If two letters are in the same row or column, they are replaced by the succeeding letters. For example, “so” enciphers to “tn”. Otherwise, the two letters stand at two of the corners of a rectangle in the table, and we replace them with the letters at the other two corners of this rectangle. For example, “dq” enciphers to “fk”.

The breakpoint of attacking Playfair encryption is: It has the property that if a single letter of a plaintext pair is changed, then often only a single letter of the cipher text will change. Thus, using the key in the table, it enciphers ‘rd’ to ‘tb’ while ‘rf’ enciphers of ‘ob’ and ‘rg’ enciphers to ‘nb’. The consequence is that, given enough cipher text or a few probable words, the table or an equivalent one can be reconstructed.

From the above introduced encryption methods, we may conclude three factors that are critical to block cipher:

- The block of the cipher should be big enough;
- The encipher key / method / function should be carefully selected;
- There should have enough encipher rounds.

If the block length is too short, once the hacker gets enough paired plaintext and ciphertext, he could construct a table to be used for decryption; If the key length is too short, the cryptographic scheme would not be secure because it would be too easy to search through all possible keys; The security of a block cipher can be greatly improved by using more encipher rounds. The encryption method AES involves more than one round of encryption.

3.1.3 Block Cipher: AES

Before going into AES, I would like to introduce two concepts: confusion and diffusion. In cryptography, confusion and diffusion are two essential properties of a cipher which are necessary to make it resistant to cryptanalysis[5].

- Confusion refers to removing the correlation between the occurrence of a symbol in the plaintext and its occurrence in the ciphertext.
- Diffusion refers to the property that a change of a single bit in the plaintext changes a large number of bits in the ciphertext [10].

In AES, various methods like permutation, add round keys, S-box (contains encryption function) are used to realize text confusion and diffusion.

In 1997 in America, the National Institute of Standards and Technology (NIST) initiated a process to select a symmetric-key encryption algorithm to be used to protect sensitive Federal information in furtherance of NIST’s statutory responsibilities. At a minimum, the algorithm should be faster and better as Triple DES[6], easy to apply and test, it would have to implement symmetric key cryptography as a block cipher and support a block size of 128 bits and key size of 128, 192 and 256 bits. On end, Rijndael is selected as Advanced Encryption Standard (AES), which was invented by Vincent Rijmen and Joan Daemen [11].

[5] Cryptanalysis is the study of methods and techniques for recovering information from encrypted material (produced by ciphers or codes), without knowledge of the key or codebook, or more generally, any attack on a cryptographic scheme that attempts to bypass a security measure.

[6] Triple DES is derived from DES (Data Encryption Standard). DES was published in 1977 by the National Bureau of Standards in USA. It was based on an algorithm known as the Lucifer cipher designed by IBM. DES uses a 56-bit key, 64-bit block and 16 rounds of encryption.

The block cipher Rijndael is designed to use only simple whole-byte operations. Also, it provides extra flexibility over that required of an AES candidate, in that both the key size and the block size may be chosen to be any of 128, 192, or 256 bits.

The general sectors in AES include: Add Round Key (ARK); Byte Substitution Byte (BSB); Shift Row (SR); Mix Column (MC).

ARK: The key material is added byte by byte after BSB, SR, MC. This means that 16 bytes of key material are needed per round; they are derived from the user supplied key material by means of recurrence relation. The round keys are derived from the cipher key by means of a key schedule.

BSB: AES uses a single S-box which acts on a byte input to give a byte output. For implementation purposes it can be regarded simply as a lookup table of 256 bytes. (Confusion)

SR: The shift is based on arranging the 16 bytes of the value being enciphered in a square and then doing byteweise shuffling and mixing operations. The top row of four bytes is left unchanged, while the second row is shifted one place to the left, the third row by two places and the third row by three places.

MC: The 4 byte standing in the same column will be combined. The effect of this combination is that a change in the input to the cipher can potentially affect all of the output after just two rounds. (Diffusion)

The general procedure of AES is:

ARK; {BSB, SR, MC, ARK}^{n};BSB, SR, ARK

Here *n* represents the number of rounds needed.

To encipher a block of data in AES, one first performs an Add Round Key step (XORing a subkey with the block) by itself; the regular rounds noted above, the final round with the Mix Column step is omitted.

3.1.4 Limits of the Symmetric Encryption

Symmetric Encryption uses the same key to encrypt a plaintext and decrypt a ciphertext. Therefore, inevitably it has two main drawbacks:

- One key is used between every two communication partners. With the growth of the number of the communication partners, the number of needed keys increases tremendously. For 3 people, 3 keys are needed; for 4 people, 6 keys; 10 people, 45 keys…
- The key between two partners must be secret to others. It is not difficult to realise if two partners can meet face to face. What if they are far away from each other? How to keep the communication tunnel secure from being eavesdropped? It is not impossible such a problem to solve, but it requires more cost and time, the solution is not completely secure, which endangers the security of the ciphertext.

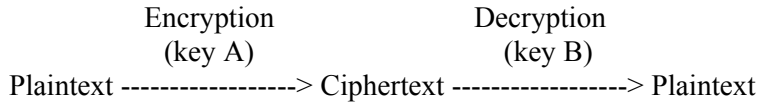
A possible solution for the bulk key problem is to use a trusted node known as a Key Distribution Center (KDC). The KDC knows keys for all users and resources. If a new node is installed in the network, only the new node and the KDC need to be configured with a key for that node. I will introduce the usage of KDC later in Kerberos. KDC makes key distribution much more convenient but it must be dependable, secure and efficient, else it will cause big trouble once the KDC crashes or breaks down when being attacked.

The only kind of system that offers perfect secrecy was first proposed by Gilbert Vernam during World War I: To make a stream cipher against attacks is for the key sequence to be as long as the plaintext, and to never repeat. The restriction of One-Time Pad is: First of all, it needs keys as many as the plaintexts. Therefore it will be very expensive for most applications. Nowadays it is still used for high-level diplomatic and

intelligence traffic. Secondly, how to transfer the secret key safely from the sender to the receiver if they live far from each other? Actually, this is a problem to all symmetric encryption, which will be explained later.

3.2 Asymmetric Encryption (Public Key Encryption)

Asymmetric cryptography was invented in 1975. Unlike symmetric cryptography, keys are not shared. Instead, each individual has two keys: a private key that must not be revealed to anyone, and a public key that is known to the entire world.



Here I would like to introduce two asymmetric encryption methods: RSA and ElGamal. The technique of asymmetric cryptography is to make the security of the cipher depend on the difficulty of solving a certain mathematical problem. The two problems used in most fielded systems are factorisation (RSA) and discrete logarithms (ElGamal).

3.2.1 RSA

RSA is a method based on factoring and commonly used to do public key encryption and digital signatures. The inventors of RSA are Ron Rivest, Adi Shamir, and Len Adleman. As we know, prime numbers are positive whole numbers, which can only be divided by 1 and the number itself. These numbers are: 2, 3, 5, 7, 11, 13, 17… It is relatively easy to find prime numbers and multiply them together to give a composite number, but much harder to solve a composite number into its factors. It is believed that a similar number of 1024-bits-length could not be factored without an advance in mathematics. Two mathematic theories supporting RSA are as follows:

1. Fermat`s Little Theorem
   For all primes *p* not dividing *a*, $a^{p-1} \equiv 1$ modulo p
   The detailed proof can be found in [12].

2. Euler`s function ø(N)
   ø(N) is the number of positive integers less than *n* with which it has no divisor in common: so if *n* is the product of two primes *p*, *q* then ø(N) = (p-1) (q-1).
   The detailed proof can be found in [13].

The encryption key is a modulus N which is hard to factor, *p* and *q* are two large randomly chosen primes. $N = p * q$ plus a public exponent *e* that has no common factors with either *p-1* or *q-1*. The private key consists of *p* and *q*, which are kept secret. Where *M* is the message and *C* is the cipher text, the formulas are defined as:

Encryption: $C \equiv M^e$ modulo N
Decryption: $M \equiv C^d$ modulo N

The two formulas are based on Fermat's Little Theorem:

$$C^d \equiv \{M^e\}^d \equiv M^{1 + k\,ø(N)} \equiv M.M^{\,k\,ø(N)} \equiv M * 1 \equiv M \text{ modulo N}$$

In RSA, {p, q} is the private key. {e, N} is the public key. *e* has no common factors with ø(N). To decipher the text, we need to calculate a number *d* out of *p* and *q*, which should be kept secret as well. *d* should be a number such that $d*e \equiv 1$ modulo ø(N).

Here is an example for RSA. In this case, *N* is relatively small. A much larger *N*, which has 512 bits or 1024 bits, is required in real security systems.

Message M:	422711082520
Public key N:	1073
Public exponent e:	17
Secret key p:	29
Secret key q:	37
Decipher key d:	573

First of all, we divide *M* into several blocks. Each block is smaller as *N*.

M = 422.711.082.520

$m_1 = 422$; $m_2 = 711$; $m_3 = 082$; $m_4 = 520$

Second step is encryption according to the definition $C \equiv M^e$ modulo N.

$c_1 = m_1^e$ modulo N = 422^{17} modulo 1073 = 587;
$c_2 = m_2^e$ modulo N = 711^{17} modulo 1073 = 134;
$c_3 = m_3^e$ modulo N = 082^{17} modulo 1073 = 948;
$c_4 = m_4^e$ modulo N = 520^{17} modulo 1073 = 240.

===> C = 587.134.948.240

Decipher uses the formula $M \equiv C^d$ modulo N.

$m_1 = c_1^d$ modulo N = 587^{593} modulo 1073 = 422;
$m_2 = c_2^d$ modulo N = 134^{593} modulo 1073 = 711;
$m_3 = c_3^d$ modulo N = 948^{593} modulo 1073 = 082;
$m_4 = c_4^d$ modulo N = 240^{593} modulo 1073 = 520.

===> M = 422.711.082.520

3.2.2 ElGamal

ElGamal is an encryption method using discrete logarithms. The basic theory behind ElGamal is: For each natural number *y*, there is a number *x*, which satisfies the equation: $y \equiv g^x$ modulo p. *p* is a random prime number, *y, g, x* are natural numbers smaller than *p*, that is, 1<y, g, x <= p-1. I will give an example which illustrates the relationship between these numbers, suppose *p* is 11 and *g* is 7.

y	x	because
1	10	7^{10} modulo 11 = 1
2	3	7^3 modulo 11 = 2
3	4	7^4 modulo 11 = 3
4	6	7^6 modulo 11 = 4
5	2	7^2 modulo 11 = 5
6	7	7^7 modulo 11 = 6
7	1	7^1 modulo 11 = 7
8	9	7^9 modulo 11 = 8
9	8	7^8 modulo 11 = 9
10	5	7^5 modulo 11 = 10

It is easy to find corresponding *x* for *y* by calculating so long as *p* is small. If the number is big it will be very difficult. ElGamal makes use of this difficulty and develops the encryption method. In ElGamal {*p, g, y*} is the public key, *x* is kept secret.

The encryption of Message m is: the sender takes a random number *k*, *k* is relatively prime to p and calculates *a* and *b* and send them to the receiver.

$a \equiv g^k$ modulo p
$b \equiv y^k$ m modulo p

After having the value of *a* and *b*, the receiver decrypts the message with this formula:

$(b / a^x) \equiv$ m modulo p

Let us look at this example. We choose *p* as 23, *x* as 3, *g* as 5, $y = g^x$ modulo p, that is, y = 10 and *k* as 7. The plain text *m* is 13.

$a = 5^7$ modulo p → a = 17
$b = 10^7$ m modulo p → b = 130,000,000

When the receiver gets the value of *a* and *b*, he uses the secret key *x* to decipher the cipher text.

$(130{,}000{,}000 / 17^3) \equiv$ m modulo 23 → m = 23-10 = 13

3.2.3 Concerns of the Asymmetric Encryption

Same with symmetric encryption, asymmetric encryption also has the concern of how to distribute the public keys safely over the network. With public key cryptography, key distribution is easier. Each node is responsible for knowing its own private key, and all the public keys can be accessible in one place. If, for example, all the public keys are saved in a certain place, how to make sure the information is correct? How to keep the information from being tampered? An intruder, who might break into this place, can change the public key.

The typical solution is to have a trusted node known as a Certification Authority (CA) that generates certificates, which are signed messages specifying a name and the corresponding public key. A certificate might be described symbolically as:

$C_A = \text{Sig}_{Ks}(T_S, L, A, K_A, V_A)$

Here T_S is the certificate's starting date and time, L is the length of time for which it is valid, A is the user's name, K_A is his public encryption key, and V_A is his public signature verification key. In this way, only the administrator's public signature verification key: V_S needs to be communicated to all the users and resources over the network in a trustworthy manner.

3.3 One-Way Hash Function

Till now, the methods we have discussed help to prevent the message from being read by third parties. What if the third party gets the message on the way and changes the message or sends his own message but with a camouflaged identity? To ensure the security of the message, we should take not only the confidentiality of the message but also integrity and authentication into account.

During the transportation of a message, it may go through dozens of middle stations before it finally arrives at the destination. It is possible that the data packets are destroyed on its way to the target. Therefore sender programs will set a checksum for each packet, which examines whether bits in this packet are destroyed or changed. After receiving the packet, the receiver programs are able to verify the integrity of the packet with the help of the checksum. This kind of checksum-algorithm is called hash function.

A one-way hash function takes a variable-length input sequence of bytes and converts it into a fixed-length sequence. The fixed length is considerably shorter than the typical length of the input, and hence the function is a hash function. It can be used to generate a MIC (Message Integrity Code) to protect the integrity of messages transmitted over insecure media.

One-way-hash function should meet the following three conditions.

- It should be able to deliver a relatively small value as the checksum for any message no matter its length.
- The algorithm should be easy to calculate.
- Two different messages should not have the same hash value. Even though it is obvious that many different values of messages will be transformed to the same hash value because there are many more possible values of messages, it is computationally infeasible[7] to find two values that hash to the same thing.
- Each message has a hash value, but the correct message should not be deduced from the given hash value.

If we merely sent the message and used the hash of the message as MIC, this would not be secure. However, if the sender (S) concatenates a key (only sender and receiver know this key) to the data, computes the hash of the message (data and key) then sends the hash SH and the data to the receiver R. R can add the key to the data and calculate the hash value RH. If SH and RH are the same, R can have the confidence the message was sent by someone knowing the key. In this sense, one-way hash function works to some extent like a Digital Signature, detailed introduction of Digital Signature is in next sector.

One-way hash functions are also known as message digests (MD), fingerprints, or compression functions. The most popular one-way hash algorithms are MD4 and MD5 (both producing a 128-bit hash value), and SHA[8], also known as SHA1 (producing a 160-bit hash value) [14].

3.4 Digital Signature

Digital Signature is also widely used to authenticate communication partners. Digital Signature is the equivalent of the hand-written signature on a document. It is often useful to prove that a message was generated by a particular individual, especially if the individual is not necessarily around to be asked about authorship of the message. The basic idea of Digital Signature is public-key cryptosystem. Each user who wishes to sign messages has a private signing key *X* and a public signature verification key *Y*. It is different from one-way hash function mentioned before. One-way hash function using a secret key to verify the sender, the communication partners share the same key to authenticate while Digital Signature uses the public key. Bob's signature for a message *m* can only be generated by someone knowing Bob's private key. And the signature depends on the contents of *m*. If *m* is modified in any way, the signature no longer matches. Digital Signature provides two important functions: it proves who generated the information, and proves that the information has not been modified in any way by anyone since the message and matching signature were generated.

In addition to realization of authentication and integrity, Digital Signature is also an effective technique to implement non-repudiation. For example, Party A has ordered a flight ticket to Africa at a travel agency through the Internet. Later due to some emergencies, he is not able to make the travel any more. In order to

[7] Roughly speaking, computationally infeasible means that a certain computation that we are talking about takes way too long (hundreds of years) to compute using the fastest of (super) computers.

[8] SHA is the NIST (National Institute of Standards and Technology) proposed message digest function, which stands for secure hash algorithm.

avoid the penalty for withdrawing the confirmed order, he might think of denying that he has ever intentioned to purchase the ticket. If the system they use supports Digital Signature, then Party A will have great trouble for doing so because the travel agency has the order signed with A's private key. Only someone has knowledge of A's private key could have made this order.

Till now, I have elucidated on the definitions of confidentiality, integrity, authentication, non-repudiation, availability and accountability. Traditionally, confidentiality, integrity and availability are widely accepted as three most significant elements of information security. They can be remembered by the mnemonic "CIA". Authentication is closely connected to integrity, without authentication, the integrity of the information can not be checked. To e-commerce, non-repudiation is indispensable. Accountability is actually a combination of other primitives. It self does not prevent any attacks.

The methods we employ to aid in the confidentiality, integrity, authentication and non-repudiation of information are cryptography-based. But generally we use non-crpyto-methods like backups, fail-over systems, as well as disaster recovery plans to achieve the availability of information.

4. Information Security Protocols

Based on the combination of various encryption methods and variants, protocols and systems are developed to keep a certain resource or a communication tunnel safe from being attacked. Here I would like to introduce three of them: Kerberos, VPN and PGP.

4.1 Kerberos

Kerberos is a network authentication protocol. It is designed to provide strong authentication for client/server applications by using secret key cryptography [15]. Authentication is very closely connected to Integrity. We can use symmetric encryption, asymmetric encryption and hash function to authenticate a message.

The name Kerberos comes from Greek mythology; it is the three-headed dog that guarded the entrance to Hades. It is said in computer security realm, three heads represent client, server and KDC (Key Distribution Centre).

Kerberos has an authentication server (AS) to which users log on, and ticket-granting server—KDC, that gives them tickets allowing access to various resources such as files. This enables more scalable access management. After logging into the system by giving his username and password to the AS, the user will probably need to access remote resources. There remote resources need to authenticate his identity. AS will perform the authentication protocol on user's behalf.

For example, user Alice logs on to the authentication server using a password. The client software in her PC fetches a ticket from this server, which is encrypted under her password and which contains a session key K_{AS}. Assuming she gets the password right, she now controls K_{AS}; to get access to a resource B, controlled by the ticket-granting server S, the following protocol takes place. Its outcome is a key, K_{AS}, with timestamp Ts and lifetime L, which will be used to authenticate Alice's subsequent traffic with that resource [10]:

1. Alice	->	Server:	A, B
2. Server	->	Alice:	$\{T_S, L, K_{AB}, B, \{T_S, L, K_{AB}, A\} K_{BS}\} K_{AS}$
3. Alice	->	Resource B:	$\{T_S, L, K_{AB}, A\} K_{BS}$, $\{A, T_A\} K_{AB}$
4. B	->	Alice:	$\{T_A+1\} K_{AB}$

The KDC shares a secret key with each user and resource.

(1) & (2): When Alice informs the KDC that she wants to talk to resource B, the KDC invents a session key K_{AB} for Alice and B to share, encrypts K_{AB} with Alice's secret key for Alice, encrypts K_{AB} with B's secret key for B and returns all this information to Alice.
The message containing the session key K_{AB} ciphered with B's secret key is known as a ticket to B. Alice can't read what's inside the ticket, because it's encrypted with B's secret key.

(3) & (4): After receiving the message from Alice, B decrypts the ticket and discovers K_{AB} and Alice's name. Based on the ticket, B knows that anyone else who knows K_{AB} is acting on Alice's behalf. B sends the reply back with a timestamp using K_{AB}.

So Alice and resource B can authenticate each other, and optionally encrypt or integrity-protect their entire conversation, based on the shared key K_{AB}.

4.2 PGP: Pretty Good Privacy

PGP stands for Pretty Good Privacy, is a free secure mail protocol. Written by Phil Zimmermann and released in 1991, PGP works on virtually every platform and has become very popular.

The article “Why do you need PGP?” from Phil Zimmermann is fun to read [16]. The author states although we honest people do not have anything to hide, we should all start encrypting our email so that in case someone needs privacy, he won't draw attention by being the only one encrypting mail. Analog is the snail mail, if everyone writes the letter on the postcard, those whoe use an envelope will arouse suspicion.

PGP regards mail is the same as ordinary files. Someone who wishes to send a secure mail message could first transform the file to be mailed using PGP, and then mail the transformed file using a traditional mailer. Similarly, if one were to receive a PGP-encrypted mail message, one could treat the received message as a file and feed it to PGP to process.

PGP has now a number of versions but in its most basic form, it serves a dual role. It can be used to encrypt, it can also be used to authenticate a package. Each user generates a private/public keypair. To protect a message, you sign a hash of it using your private key, encrypt both the message and the signature with a randomly chosen session key, then encrypt the session key using the public key of each of the intended recipients. Thus, if Alice wants to send an encrypted email M to Bob and Charlie, she forms the message [10]:

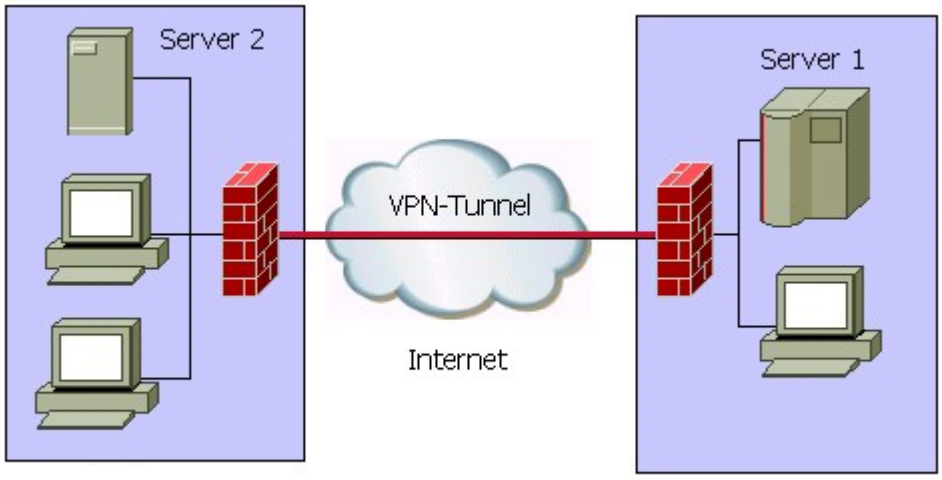
$$\{K_S\}_{KB}, \{K_S\}_{KC}, \{M, sig_{KA}\{h(M)\}\}_{KS}$$

The message M is signed with Alice’s private key K_A. Then the message it self and the hash value are enciphered with a random session key K_S. Since Alice wants to send the message to Bob and Charlie, she encrypts K_S with Bob and Charlie’s public keys K_B and K_C. Only Bob and Charlie or anyone who knows their private keys can get the session key K_S, then checks the hash value and trusts the message if the hash value is proved to be correct.

4.3 VPN: Virtual Private Network

Virtual Private Network, or VPN, is a private communication network usually used within a company, or by several different companies or organisations, communicating over a public network. VPN message traffic is carried on public networking infrastructure (ie, the Internet) using standard (possibly insecure) protocols [17].

The idea behind VPN is: if an organisation has a number of branches, or a number of organizations that communicate with each other, they encrypt the traffic between internal sites at their firewall. This way, the Internet can connect their networks without their communication being exposed to eavesdropping.



Branch Headquarters

This picture well illustrates the VPN concept. Server 1 is the headquater of an organization and Server 2 is one of its branches. The communication between Server 1 and Server 2 using Internet network but encrypted. Therefore, the content will not be disclosed to other people.

The most obvious advantage of VPN comparing with WAN (Wide Area Network) is: It costs much less money to build VPN instead of WAN because VPN uses public network infrastructures (Internet networks) instead of private leased lines.

There are several types of VPN protocols catering to different requirements and the implementation of VPN requires an in-depth understanding of public network security issues and taking proper precautions in VPN deployment. It is said if correctly selected and implemented, VPN can really realize the security of the information.

5. Conclusion

Confidentiality, integrity, availability, authentication, non-repudiation and accountability are important principles for information systems security. Cryptography is the basic theory behind all the techniques we use to keep these principles.

Generally there are two main classes in cryptography, symmetric encryption and asymmetric encryption. The former uses a single key. The same key is used to encrypt and decrypt a message. This key has to be kept secret otherwise the scheme is compromised. The main problem is how to distribute the secret key and ensure that it remains secret. Asymmetric Encryption uses two keys, one is private key, the other public key. The private key is kept secret, the other key is made public. To communicate with the owner of the private key, a message is encrypted with the corresponding public key, this message can only be decrypted using the private key.

Derived from encryption methods, Digital Signatures are used to authenticate user's or resource's identity, check the integrity of a message and realise non-repudiation; Kerberos helps to authenticate users; VPN is developed to maintain the security of a private network; PGP is designed to keep electronic mail private.

Many other applications are developed to keep the Internet safe and the communication private. However, still the Internet is a place offers no absolute security. Just as Sir Arthur Conan Doyle, the writer of Sherlock Holmes said: "What one man can invent another can discover." The war between the cryptography and cryptanalysis will last.

Bibliography

[1] Internet Encyclopaedia: word iQ at http://www.wordiq.com/definition/Security

[2] Internet Encyclopaedia: word iQ at http://www.wordiq.com/definition/Information_security

[3] National Information Assurance (IA) Glossary, CNSS Instruction No. 4009, P33 at http://www.nstissc.gov/Assets/pdf/4009.pdf

[4] Internet Encyclopaedia: Wikipedia at http://en.wikipedia.org/wiki/Confidentiality

[5] Internet Encyclopaedia: Wikipedia at http://en.wikipedia.org/wiki/Authentication

[6] M. Vandenwauver. "Introduction to Cryptography" at http://www.esat.kuleuven.ac.be/cosic/intro/

[7] The Encyclopaedia of Computer Security at http://www.itsecurity.com/dictionary/nonrepud.htm

[8] University of North Texas Network Security Guide at http://www.unt.edu/security/awareness/netmanguide.html

[9] Ross. Anderson, "Security Engineering", John Wiley & Sons, Inc (2001), ISBN: 0-471-38922-6

[10] Internet Encyclopaedia: WorldHistroy.com at http://www.worldhistory.com/wiki/C/Confusion-and-diffusion.htm

[11] Daeman, Joan. Rijmen, Vincent, A Specification for Rijndael, the AES Algorithm at http://ece-classweb.ucsd.edu/ece111a/aesspec.v311.pdf

[12] http://www.fact-index.com/p/pr/proofs_of_fermat_s_little_theorem.html

[13] http://www.cut-the-knot.org/blue/Euler.shtml

[14] Charlie. Kaufman, Radia. Perlman, Mike. Speciner: "Network Security", Prentice Hall PTR (1995), ISBN 0-13-061466-1

[15] Website of Massachusetts Institute of Technology: http://web.mit.edu/kerberos/www/#what_is

[16] The International PGP Home Page: http://www.pgpi.org/doc/whypgp/en/

[17] Internet Encyclopaedia: Wikipedia at http://en.wikipedia.org/wiki/VPN