Information Technology
Inside and Outside
- David Cyganski & John A. Orr

VII. Networks and the Internets

21. Electronic Commerce and Information Security

Hoon-Jae Lee
http://cg.dongseo.ac.kr/~hjlee
E-mail: hjlee@dongseo.ac.kr
http://cg.dongseo.ac.kr/~hjlee

21. Electronic Commerce and Information Security

- Objectives:
  - the four fundamentally different **threats to data security**;
  - the **six basic types of security services** which may be provided;
  - **basic principles of cryptography** (creating secure codes) and **cryptanalysis** (the attempt to break the codes);
  - some examples of **classic encryption/decryption** schemes;
  - means of determining the relative level of security of a given coding system;
  - the concept of a ``**key,**'' and **public and private key crypto-systems**; and
  - concepts of **digital signatures**, **digital cash**, and other necessary components of **electronic commerce**.

21.1 Introduction

- In 1999, **Amazon.com, Inc.** claimed to be the **world's largest bookstore**, offering over **4.7 million book titles** in addition to music CDs and videos, all accessible over the Internet.
  - In 1997, net sales increased by 838% to 147.8 million dollars.
  - In 1998 those sales climbed again, to 1 billion dollars.
  - By the end of 1999 its sales had reached 1.64 billion dollars.
- **Cisco Systems, Inc.,** is **a computer networking giant** that has also observed increased profits due to conducting business over the Internet,
  - reporting sales of$4.1 billion in 1996,
  - $6.4 billion in 1997, and
  - $8.5 billion in 1998.
- How does one conduct private and sensitive financial transactions on such a system without betraying enough information to become **susceptible to theft and fraud**?
  - The key is the use of **information security protocols**; that is, **secret codes**.

21.2 Threats to Information Security

- Four potential security ***threats***
  - ***Data disclosure;***
  - ***Fraud;***
  - ***Data insertion, removal, and modification; and***
  - ***Denial of service.***
- ***Data disclosure:*** the threat of someone being able to read and comprehend our data.
  - For example, if we are transmitting our **credit card number** over the Internet, we would hope that no one could see it because they could then use it to make purchases in our name.
  - **Eavesdropping** is an example of an attack resulting in data disclosure.
- ***Fraud:*** the ability to misrepresent identities
  - In order for a consumer to feel comfortable providing an electronic storefront with information such as credit card numbers, the consumer must be able to positively **authenticate the identity of the commercial enterprise** at the other end of the connection. Likewise, the commercial enterprise will not want to commit a transaction unless **the consumer can be identified**.

21.2 Threats to Information Security (2)

- **Data insertion, removal, and/or modification**
  - Consider an **electronic banking system**. If it is possible to **modify** the data during transit, then it is possible to **alter** the financial transactions.
  - The attacker need not even be able to understand the data to cause havoc in this situation.
- **Denial-of-service(DOS) attack**
  - **preventing you from accessing data or service**
  - **by confusing or overloading the computers or networking equipment** involved in your transaction.
  - To establishing a TCP connection, **synchronization messages(SYN)** are exchanged between the two communicating computers.
  - With the TCP SYN flooding attack, a malicious user attempts to open so many TCP connections with a server that additional incoming connections will be denied because **all server resources used** to store and track these connections have been used up.
  - The **TCP SYN flooding attack** was used to bring down an ISP in New York City on September 6, 1996.[2] Further complicating matters, interactive Web technologies such as Java, JavaScript, and ActiveX increase the difficulty of preventing denial-of-service attacks.

21.3 Security Services

- Several **security services** have been categorized
  - ***Privacy***: preventing others from being able to comprehend our data.
  - ***Authentication***: positively identifying an object or identity.
  - ***Access Control***: restricting access to data or a service to privileged identities only.
  - ***Integrity***: ensuring that the data has not been altered since its creation.
  - ***Nonrepudiation***: ensuring that the originator cannot deny being the source of the data, and that the recipient cannot deny that the data was received.
  - ***Replay Prevention***: ensures that data previously deemed valid cannot be resent by an attacker and mistakenly validated by a system a second time.
- A ***privacy*** service is one that allows us to hide information in plain sight.
  - **Secret codes** are routinely used on the Internet to do exactly this.
  - In fact, **the commercial success of the Internet** would not have been possible without major advances in the theory and practice of **secret coding technology**.

21.3 Security Services(2)

- ***Authentication*** is used to prove that each end of a transaction is who it claims to be.
  - As you will see in the next section, **secret codes** play an important role in this too, but ultimately some party, somewhere, must be trusted in order to build a chain of trust down to the end parties.
  - Companies have appeared on the Internet (known as **Certificate Authorities**) that provide that trusted-party role for digital transactions.
- ***Access control*** can be enforced in as simple a fashion as **locking the door** to a computer with important information and only **giving a key to trusted individuals**.
  - Access control for computers connected to the Internet is often provided through the use of special equipment or software that is called a **``Firewall''** or **``Proxy Server.''**

21.3 Security Services(3)

- The operation of a ***Proxy Server***.
  - **prevent someone from finding some vulnerability of your computer** (like a weak password checking system) which might lead to a break-in by the attacker and loss of important data.
  - **connect just one computer**, **the Proxy Server**, to the Internet and place no confidential data on it.
  - executes **proxy server software** that allows your other computers to ask it for information from the Internet.
  - Thus, the computers inside ``the firewall" have access to the outside thanks to the handling of those requests by the Proxy Server, but only the proxy Server is threatened by the outside world.
- **Access control** can also involve using transmission cables that cannot be *tapped*, or requiring authentication prior to allowing a computer to divulge the content of certain data files.

21.3 Security Services(4)

- ***Integrity*** services provide means to determine whether or not data has been altered regardless of the care and resources of the counterfeiter.
  - This may seem impossible to do, but in the next section you will discover that a special application of secret code technology allows the creation of ***digital signatures*** that provide exactly that capability.
- ***Non-repudiation*** services disallow a different kind of fraud than that considered above.
  - Suppose we **purchase an expensive watch** and then call the credit card company and report our card as having been stolen just before that purchase.
  - Using techniques for non repudiation in an Internet transaction allows the credit card company to prove that it was indeed you that made the purchase.

21.4 Data Security and Cryptosystems

- **Cryptology** is composed of two fundamental fields: **cryptography** and **cryptanalysis**.
- The primary objective of **cryptography** is to allow two or more users to communicate securely over an insecure medium, such as the Internet.
  - The information to be transmitted, referred to as the plain text, is encrypted using a predetermined key to generate the cipher text.
  - The cipher text is transmitted over the insecure medium to the receiver, who recovers the plain text using a cryptographic key and algorithm.
  - Cryptography is the science of applying secret codes.
- **Cryptanalysis** is the process of recovering the encryption key from the cipher text; that is, breaking the code.
  - Cryptanalysis teams work hand in hand with cryptography teams because we cannot be sure how good the cryptanalysis practiced by attackers will be.
  - Our only hope for secure transactions is to vigorously attempt to break our own coding systems.

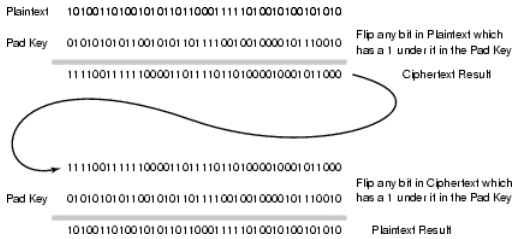
21.4 Data Security and Cryptosystems(2)

- The **strength of a crypto system** (a specific secret code system)
  - **computationally secure**, and
  - **unconditionally secure**.
- A crypto system is said to be **computationally secure** if the best known attack requires an amount of computational resources that is far too excessive to be a threat in practice.
- A cryptosystem is said to be **unconditionally secure** if the cryptosystem is secure against an attack with an infinite amount of resources available.
- Cryptosystems can further be divided into two categories: **symmetric key cryptography** and **public key cryptography**.
- **Working together**, **symmetric key and public key cryptography systems** possess the necessary characteristics to achieve information security for a wide variety of systems, including **secure electronic commerce, e-mail and WWW interactivity**.

21.4 Data Security and Cryptosystems(3)

- An example of an **unconditionally secure cryptographic algorithm** is the ***one-time pad*** code.
  - Historically, one-time pad Cryptosystems have been used by diplomats when exchanging sensitive information. In the original application of this type of system, the code by which information is encrypted is contained on the pages of a pad of paper. Two copies of the pad are then made, one to be used to encrypt the plain text and the second to be used to decrypt the cipher text.

Figure 21.1: The operation of a one-time pad cryptosystem.



21.4 Data Security and Cryptosystems(4)

Table 21.1: A table for generating cipher text from plain text given key consisting of either numerical shifts or English text.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

21.4 Data Security and Cryptosystems(5)

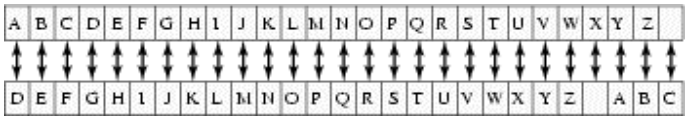
- An example of an **unconditionally secure cryptographic algorithm** is the ***one-time pad*** code.
  - Historically, one-time pad Cryptosystems have been used by **diplomats** when exchanging sensitive information. In the original application of this type of system, the code by which information is encrypted is contained on the pages of a pad of paper. Two copies of the pad are then made, one to be used to encrypt the plain text and the second to be used to decrypt the cipher text.
- The top row indicates the number of spaces in the alphabet that a character in the first column is to be shifted. Below the numerical shifts in the top row are corresponding letters from the one-time pad: a space in the pad determines a shift of zero, an A determines a shift of 1, and so on.
- **After generating the cipher text, the page is destroyed, never to be used again.** Only the possessor of the second pad can then decrypt the information. One-time pad code key is simply a random string of zeroes and ones. The transformation of plain text to cipher text simply involves lining up the binary string to be coded with the code key, and then flipping every plain text bit that lines up with a 1 in the code key and not flipping those that line up with a 0.

21. 5 Computationally Secure Symmetric Key Cryptography

- In symmetric key (sometimes called private key) cryptography, a single key is used for both encryption and decryption

Figure 21.2:The operation of a private key cryptosystem.

Figure 21.3:An example of a shift cipher: the letters in the to prow are transformed into letters in the bottom row by a shift of all letters by three, and wrap around at the end of the alphabet.



21. 5 Computationally Secure Symmetric Key Cryptography(2)

- ***Caesar Cipher:*** special case of the shift cipher
  - Plaintext=message=``CRYPTOGRAPHY''→ m_i
  - Ciphertext= ``FUASWRJUDSKA" → c_i
  - Key k=3
  - $c_i=(m_i + k)$; $m_i=(c_i - k)$
  - Exhaustive Key Search = 25
  - By reversing the transformation, the clear text can easily be obtained from the cipher text.

21.5.1 Commonly Encountered Symmetric Codes

- **In 1977**, the National Institute of Standards and Technology (**NIST**; formerly the National Bureau of Standards) published the **Data Encryption standard (DES).** The standard requires that NIST review the security of the algorithm **every 5 years**. DES was last **reviewed in 1993** and was approved for unclassified applications until 1998.

21. 5 Computationally Secure Symmetric Key Cryptography(3)

- While **DES** remains in widespread use, it is **no longer secure** and must be replaced with a more robust approach.
  - DES is a block cipher capable of encrypting **64 bits of data** at a time using a **56 bit key**. The resulting cipher text is 64 bits in length. The cipher text can be decrypted to obtain the original bit sequence using the same key used to encrypt the data.
- The Advanced Encryption Standard (**AES**) was approved **in the United States in 2001**.

Table 21.2:Common Private Key Algorithms

Algorithm	Mode	Block Size (bits)	Key Size (bits)
DES	Block Cipher	64	56
Triple-DES	Block Cipher	64	112
RC2	Block Cipher	64	Variable
RC4	Stream Cipher	Variable	Variable
RC5	Block Cipher	Variable	Variable
IDEA	Block Cipher	64	128

21. 5 Computationally Secure Symmetric Key Cryptography(4)

- Because **the DES key is only 56 bits in length**, the security of the algorithm is often questioned, as modern computational resources make the algorithm **vulnerable to attacks** based on an **exhaustive search of the key space**.
- For this reason, **various modes of operation** believed to strengthen the security of DES have been introduced.
- The **triple-DES** algorithm effectively doubles the security of DES by encrypting the data three times with multiple keys. DES and triple-DES are commonly used private key algorithms in both the government and commercial sectors.
- Popular algorithms in use today on the Internet are **RC2**, **RC4**, and **RC5**.
  - Both **RC2 and RC5 are block ciphers** and
  - **RC4 is a stream cipher**.

21. 5 Computationally Secure Symmetric Key Cryptography(5)

- **RC2, RC4 & RC5** were developed by **Ronald Rivest of RSA Data Security**.
  - **not patented**,
  - **trademarks are held on the names of the algorithms**.
  - In 1996 the source codes for both the RC2 and RC4 algorithms were **anonymously posted on the Internet**. All three algorithms use **variable length keys**.
  - RSA Data Security markets an implementation permitting key sizes between 1 and 2048 bits in length.
  - **40 bit RC2, RC4, and RC5** are considered relatively **insecure** and **international exportation part**.
  - widespread usage in many **WWW browsers**.
  - **When using key lengths greater than 64 bits**, the algorithms are believed to be relatively **secure** for most applications on the Internet today.

21. 5 Computationally Secure Symmetric Key Cryptography(6)

- **IDEA**(International Data Encryption Algorithm)
  - Published **in 1990** by Dr. **X. Lai** and Prof. **J. Massey** and
  - **patented by Ascom-Tech AG**,
  - a **symmetric block cipher**
  - **high security** while being **easily implemented in software and hardware**.
  - **64 bit blocks of data using a 128 bit key**.
  - Although it is generally believed that IDEA provides greater security than DES at much higher data rates, IDEA has yet to see widespread usage as a result of patent issues.

21. 6 Public Key Cryptography

- The **public key cryptography** = **asymmetric cryptography**
  - published by **W. Diffie and M.E. Hellman in 1976**.[2]
  - **In private key Cryptosystems**, the same key is used for both encryption and decryption and must be distributed in a secure fashion.
  - **With public key cryptography**, the algorithm is chosen such that it is computationally infeasible to determine the decryption key given the encryption key and cipher text.

Figure 21.4: The operation of a public key cryptosystem.

21. 6 Public Key Cryptography(2)

- An algorithm based on **the difficulty** associated **with factoring large composite numbers into prime numbers** is an example of a **computationally secure system**.
- What does it mean to factor into prime numbers? If a number can be formed by multiplying two other numbers (**77 = 7 x 11**) then we say it is composite.
  - If it cannot be formed by multiplying two numbers, we say it is prime (7 or 11).
- If a number is sufficiently large, say **422982642794760484348001694649073**, then it requires an enormous amount of computer time to determine what numbers must be multiplied to form the composite number.
- The number in this example took more than 2 minutes to factor on a high-speed workstation, and the amount of time grows rapidly with the number of digits in the number.

21. 6 Public Key Cryptography(3)

- The **security of RSA** relies **on the difficulties associated with factoring large numbers into primes**, and to date there are no publicly known attacks capable of efficiently compromising the security of RSA for **key sizes greater than 1024 bits**.
- **public key cryptography**
  - quite **complex**,
  - the **computational cost** of encrypting and encrypting data using public key algorithms is usually much greater than that of private key algorithms.
- **private key algorithms**
  - much **faster** when applied to bulk data transfer.
- **private key hybrid schemes** are often employed.
  - **confidentially** is achieved **by first encrypting the private key** of the block or stream cipher (referred to as the symmetric key)
  - using an asymmetric algorithm and the corresponding **public key of the recipient**.

21. 6 Public Key Cryptography(4)

Figure 21.5: The steps involved in symmetric key distribution using public key cryptography.

- **21.6.1 Public Key Standards**
- The **Diffie-Hellman key exchange algorithm**, based on the mathematical properties of the discrete logarithm, is a system for distributing symmetric keys over an untrusted medium in a secure
- The **RSA algorithm** is based on the complexity of factoring large primes and can be used for both creating digital signatures and securely exchanging keys.
- The **ElGamal public key algorithm** can be used for both key exchange and digital signatures.

21.7 Digital Signatures

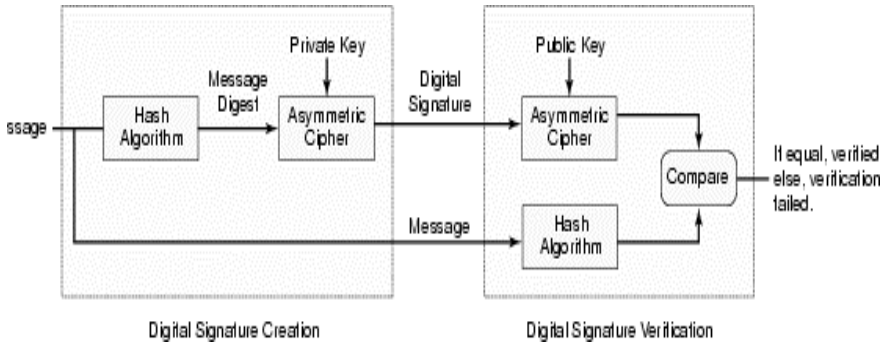
- *message digest* or *hash* functions
- Secure message digest functions possess three essential mathematical properties:
  - 1. Every input bit influences every output bit.
  - 2. If a single input bit is changed, every output bit has a 50% chance of changing.
  - 3. Given an input and corresponding message digest, it should be computationally infeasible to find another input with the same message digest.
- Common message digest functions include MD2, MD4, MD5, SHA, and SHA-1.
  - Developed **by Ronald Rivest(RSA Inc.)**, **MD2, MD4, and MD5** produce 128 bit digests
  - **MD2**: secure and computationally expensive.
  - **MD4 and MD5**: Some flaws, collisions (where two input streams result in the same digest)

21.7 Digital Signatures (2)

- The Secure Hash Algorithm (**SHA**) was developed by the NSA and produces a **160 bit message digest**.
  - Shortly after the publication of **SHA**, the NSA announced that the algorithm is **insecure** without a minor change, and thus revised SHA to create SHA-1.
  - **SHA-1** accepts a variable length input and produces a 160 message digest.
- Digital Signature Standard (**DSS**)
  - Developed by the **NSA** and adopted as a standard by the National Institute for Standards and Technology(NIST),
  - a **commonly used** system for digital signatures.

21.7 Digital Signatures (3)

Figure 21.6: Process of digital signature creation and verification.



21. 8 Digital Certificates

- **Digital certificates** provide a mechanism to connect an identity (that is, a person or machine) to a public key in a way that can be trusted and verified.
- With digital certificates, a certain entity (person, company, computer) called a ***trusted party*** is responsible for verifying a set of credentials according to a predefined policy.
- Digital certificates are issued by entities called **Certification authorities (CAs).**
- **CA's policy**
  - photo ID, a social security number, birth certificate, and background check
  - a unique e-mail address

21. 8 Digital Certificates (2)

- **Public Key Infrastructure (PKI),**
  - **Verisign, GTE, Xcert, the U.S. Postal Service, and others**.
  - PKI is a term that encapsulates the certification hierarchy and includes various levels of CAs,
  - a **root CA** authenticates and issues certificates to subordinate CAs, based on geographical region where certificates are to be issued.
  - By verifying the validity of a certificate and thus linking a public key to an identity, we can authenticate the origin of a message because only the corresponding private key could have created the signature.

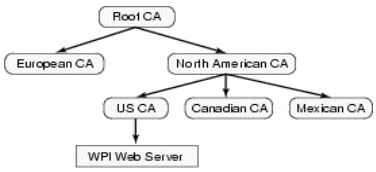


Figure 21.7: An example of public key infrastructure.

21. 8 Digital Certificates (3)

- **Public Key Infrastructure (PKI),**
  - **Verisign, GTE, Xcert, the U.S. Postal Service, and others**.
  - PKI is a term that encapsulates the certification hierarchy and includes various levels of CAs,
  - a **root CA** authenticates and issues certificates to subordinate CAs, based on geographical region where certificates are to be issued.
  - By verifying the validity of a certificate and thus linking a public key to an identity, we can authenticate the origin of a message because only the corresponding private key could have created the signature.

21. 9 Electronic Commerce

- **Three schemes for accepting credit card payments**
  - off line,
  - online without encryption, and
  - online with encryption
- Secure Sockets Layer (**SSL**)
  - Developed **by Netscape** Communications Company,
  - **a standard** developed to encrypt data **between WWW browsers and servers.**
  - When you are using a Netscape Web browser, for instance, the small key symbol on the browser window indicates whether or not your current communications are SSL secure or not.

Figure 21.8: The Netscape browser symbolically indicates with a broken key (left)or a full key (right) whether or not a Secure Socket Layer connectionis being used at any given time.

21. 9 Electronic Commerce (2)

- Secure electronic Transactions (**SET**) payment protocol
- SET is an open industry standard for the secure transmission of payment information over communication networks including the Internet.
- SET employs public and private key cryptography, digital signatures, and digital certificates. Several industry giants are involved with the SET initiative, including Master card, Visa, GTE, IBM, RSA Data Security, Teresa, Verifone, and Verisign.
- With E-Cash, both the consumer and merchant must open an account with a bank that issues E-Cash. Once an account is established, the merchant can create a WWW-based electronic to specify. The E-Cash issuing bank provides the client with special consumer software. Prior to making a purchase, the consumer obtains digital tokens, called digital coins, from an E-Cash issuing bank. E-Cash offers the consumer both unconditional and conditional anonymity. However, the consumer always knows the identity of the merchant.