

Contents

- 1 Cryptography 2**
 - 1.1 What is cryptography 2
 - 1.2 Usage of cryptography 3
 - 1.3 Types of cryptography 9
 - 1.3.1 Symmetric Key Cryptography: 9
 - 1.3.2 Hash Functions: 9
 - 1.3.3 Asymmetric Key Cryptography: 9
 - 1.4 History of cryptography 9
 - 1.5 The Math Behind Cryptography 14

Chapter 1

Cryptography

1.1 What is cryptography

Cryptography is the science of secret communications. It is the study of secure communications techniques that allow only the sender and intended recipient of a message to view its contents. The term is derived from the Greek word *kryptos*, which means hidden. It is closely associated to encryption, which is the act of scrambling ordinary text into what's known as ciphertext and then back again upon arrival. In addition, cryptography also covers the obfuscation of information in images using techniques such as microdots or merging. Ancient Egyptians were known to use these methods in complex hieroglyphics, and Roman Emperor Julius Caesar is credited with using one of the first modern ciphers.

Long before the information age, cryptography was used only to ensure secrecy of information. Encryption was used to ensure confidentiality in communications by spies, military leaders and diplomats. The Egyptian hieroglyphs, the scytale transposition cipher used by the Spartans of Greece, waxed seals and different physical devices to assist with ciphers were used throughout history right up to modern times. These devices underwent further changes when computers and electronics came into the picture, immensely helping in cryptanalysis. Cryptography has become more mathematical now and also finds applications in day-to-day security. It helps you safely transfer or withdraw money electronically and you'd be hard-pressed to come

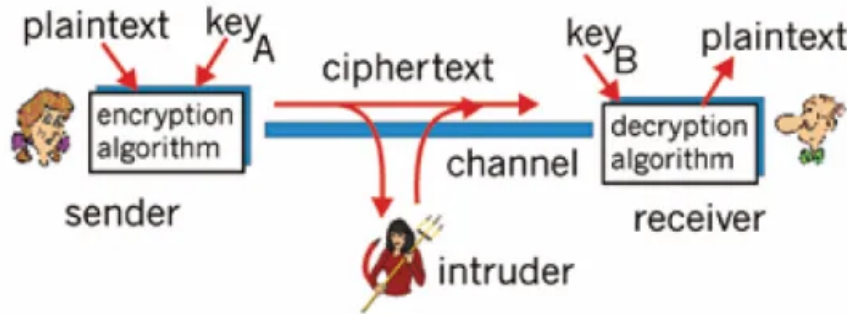
across an individual without a credit or debit card. The public-key encryption system introduced the concept of digital signatures and electronic credentials. Cryptography has a definitive existence in our lives today and the whole system will crumble in its absence. Let's now discuss the varied uses of cryptography in modern times and its intersection with computer science.



1.2 Usage of cryptography

Secrecy in transmission

The major goal of cryptography is to prevent data from being read by any third party. Most transmission systems use a private-key cryptosystem. This system uses a secret key to encrypt and decrypt data which is shared between the sender and receiver. The private keys are distributed and destroyed periodically. One must secure the key from unauthorized access, because any party that has the key can decrypt the encrypted information.



Alternately a key-generating-key, called a master key, can be used to electronically generate a one-time session-key for every transaction. The secrecy of the master-key should be maintained by all parties privy to the information. The disadvantage of this method is there's too much hope riding on the master-key, which if cracked, collapses the entire system.

A better method is to use a public-key cryptosystem. In this system, data can be encrypted by anyone with the public-key, but it can be decrypted only by using the private-key, and data that is signed with the private key can be verified only with the public key. With the development of publickey systems, secrecy can be maintained without having to keep track of a large number of keys or sharing a common master-key. If, say, Alex wants to communicate with Neil, she first generates her public/private key pair and sends the public key to Neil over a non-secure channel. Neil to encrypt information and sends it back to Alex. Only Alex has the private key with which she can decrypt the information. Anyone who intercepts the public key or the encrypted data can't decrypt the message due to the protocols followed during information transfer.

Secrecy in storage

Storage encryption refers to the application of cryptographic techniques on data, both during transit and while on storage media. Storage encryption is gaining popularity among enterprises that use storage area networks (SANs). Secrecy in storage is maintained by storing

data in encrypted form. The user has to provide the key to the computer only at the beginning of a session to access the data and it then takes care of encryption and decryption throughout the course of normal use. Hardware devices can also be used for PCs to automatically encrypt all information stored on disk. When the computer is turned on, the user must supply a key to the encryption hardware. The information is plain gibberish without its key thus preventing misuse if the disk is stolen.



Multiple ciphers can be used for individual files and folders. The ciphers and keys should be changed frequently to ensure security of data. However, if the user forgets a key, all of the information encrypted with it makes no sense and is rendered useless. This is why backups of encrypted information are advised to be stored in plaintext. The data is only encrypted while in storage, not when in use. This leaves a loophole for the attackers. The system is vulnerable to a security breach if the encryption and decryption are done in software, or if the key is stored somewhere in the system.

Integrity in transmission

We can use cryptography to provide a means to ensure that data is not altered during transmission, i.e. its integrity is preserved. In electronic funds transfer, it is very important that integrity be maintained. A bank can lose millions if a transaction is illicitly intercepted.

Cryptographic techniques are employed to prevent accidental or intentional modification of data during transmission, leading to erroneous actions. One of the ways to ensure integrity is to perform a checksum on the information being transmitted and to transmit the checksum in an encrypted form as well.

The information is received on the other end and again checksummed. The transmitted checksum is decrypted and compared with the previous checksum. If the checksums agree, the information is most likely unaltered. The problem with this scheme is that the checksum of the original message can be known and another message with the same checksum can be generated and sent instead of the original one. This problem can be overcome by using a public-key cryptosystem. After generating the public-key/private-key pair, if we throw away the private-key and use only the public-key to encrypt the checksum, the checksum becomes impossible to decrypt. In order to verify the checksum, we generate a new checksum for the received information, encrypt it using the public-key and match it with the encrypted checksum. This is also known as a one-way function as it is hard to invert.

Integrity in storage

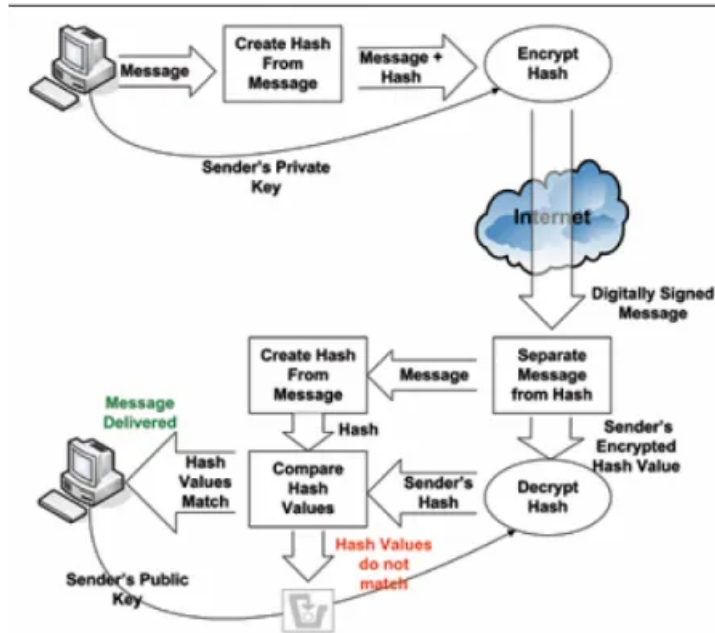
Integrity in storage had been ensured by access control systems with lock and keys and other guards to prevent unauthorized access to stored data. The existence of computer viruses has changed the scenario and the need of integrity against intentional attack has become a problem of epic proportions. Cryptographic checksums to ascertain validity of stored data are of help here. As in the case of transmission, a cryptographic checksum is produced and compared to the expected value. However, storage media are more vulnerable to attacks than transmission channels due to longer exposure and larger volumes of information.

Authentication of identity

Authentication is the process of verifying if the user has enough authority for data access. Simple passwords are used to identify someone. You must also have seen in classic gangster movies, the exchange of keywords to prove identity. Cryptography is similar to the practice of providing passwords for identity authentication. Modern systems use cryptographic transforms in conjunction with other characteristics of individuals to provide more reliable and efficient authentication of identity. Many systems allow passwords to be stored in an encrypted form, with read access available to all programs which may use them. Since passwords are not stored as plaintext, an accidental of data doesn't compromise the system's security. Passwords are analogous to the key in a cryptosystem that allows encryption and decryption of anything the password has access to. The principal element of this system is the password selection process. And that's a whole other subject that we can't cover here. But in a nutshell, the longer the password, the more random it will be and the harder it is to guess. So if you think it's easy for you to remember, you should know that it will be all the easier to crack.

Digital signatures

A digital signature is a mechanism by which a message is authenticated i.e. proving that a message is coming from a given sender, much like a signature on a paper document. To be as effective as a signature on paper, digital signatures must be hard to forge and accepted in a court of law as binding upon all parties to the transaction. The need for digital signatures arises when the parties dealing in a transaction are not physically close, and the volume of paperwork is high, in other words big business dealings. Digital signatures can be created using a public key cryptosystem and hashing process.



Hashing produces a message digest that is a small and unique representation of the original message. Hashing is a one-way algorithm, i.e. the message can't be derived from the digest. Let's say that Alex is sending a message to Neil. Alex first hashes the message to produce a digest, and then encrypts the digest with her private-key to create her personal signature; the public-key and hash algorithm are appended to it. The whole message including the digest is then encrypted using a one-time symmetric-key which is known only to Alex and Neil. Neil decrypts the message using the symmetric-key. He then decrypts the message digest using the public-key. He would then hash the original message using the same hash algorithm (whose name was appended in the message) with which it was previously hashed. If the evaluated digest and decrypted digest match, then the signature has been verified and the recipient would be sure that the message integrity has been preserved.

Another aspect of this system is the non-repudiation of digital signatures. Since the private-key is only privy to the sender, he can't deny signing the message. Also, a digital signature can be verified by anyone using the sender's public-key which is usually included in the

digital signature format.

1.3 Types of cryptography

In general there are three types Of cryptography:

1.3.1 Symmetric Key Cryptography:

It is an encryption system where the sender and receiver of message use a single common key to encrypt and decrypt messages. Symmetric Key Systems are faster and simpler but the problem is that sender and receiver have to somehow exchange key in a secure manner. The most popular symmetric key cryptography system is Data Encryption System(DES).

1.3.2 Hash Functions:

There is no usage of any key in this algorithm. A hash value with fixed length is calculated as per the plain text which makes it impossible for contents of plain text to be recovered. Many operating systems use hash functions to encrypt passwords.

1.3.3 Asymmetric Key Cryptography:

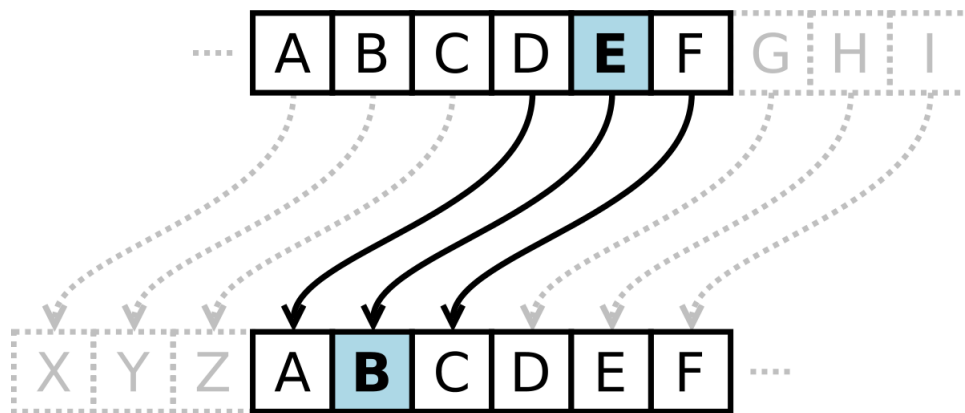
Under this system a pair of keys is used to encrypt and decrypt information. A public key is used for encryption and a private key is used for decryption. Public key and Private Key are different. Even if the public key is known by everyone the intended receiver can only decode it because he alone knows the private key.

1.4 History of cryptography

Cryptology was well established in ancient times, with both Greeks and Romans practicing different forms of cryptography. With the fall of the Roman Empire, cryptology was largely lost in the West until the Renaissance, but it flourished in the Arabic world. The Arabs

invented the first reliable tool for cryptanalysis, frequency analysis. With the end of the Middle Ages and the increase in commerce and diplomacy, cryptology enjoyed a Renaissance of its own in the West

Julius Caesar, probably the greatest of all Roman generals, was no stranger to cryptology. In his famous *Commentary on the Gallic Wars*, Caesar himself describes using a form of a cipher to hide a message. If he had anything confidential to say, he wrote it in cipher, that is, by so changing the order of the letters of the alphabet, that not a word could be made out. If anyone wishes to decipher these, and get at their meaning, he must substitute the fourth letter of the alphabet, namely D, for A, and so with the others.” (Seutonius 1957, Ch. 56) This is the first written description of the modern monoalphabetic substitution cipher using a shifted standard alphabet. Using Caesar’s cipher, the cipher alphabet looks like



For 900 years the monoalphabetic substitution cipher was the strongest cipher system in the Western world. The Romans used it regularly to protect their far-flung lines of communication. But after the fall of the Western Roman Empire in 476 C.E. the knowledge of cryptology vanished from the West and wasn’t to return until the Renaissance. Indeed, with the decline of literacy and scholarship in Europe during the Dark Ages following the fall of Rome cryptology turned from a useful technique for keeping communications secret into a dark art that bordered on magic. But interest in cryptology was not dead. In the latter part of the first millennium, there was another place where intel-

lectual curiosity and scholarship flowered and where mathematics and cryptology saw their biggest advances since Caesar – the Arab world. And it was the Arab world from which the next big advance in cryptanalytic techniques would come. The period around the ninth century C.E. is considered to be the beginning of the Islamic Golden Age, when philosophy, science, literature, mathematics, and religious studies all flourished in what was then the peace and prosperity of the Abbasid Caliphate. Into this period was born Abu Yūsuf Ya-qūb ibn Isāq as-Sabbāh al-Kindi (801–873 C.E.), a polymath who was the philosopher of the age. Al-Kindi wrote books in many disciplines including astronomy, optics, philosophy, mathematics, medicine, and linguistics, but his book on secret messages for court secretaries, *A Manuscript on Deciphering Cryptographic Messages* is the most important to the history of cryptology. It is in this book that the technique of frequency analysis is first described.



Although cryptography has a long and complex history, it wasn't until the 19th century that it developed anything more than ad hoc approaches to either encryption or cryptanalysis (the science of finding weaknesses in crypto systems). Examples of the latter include Charles Babbage's Crimean War era work on mathematical cryptanalysis of polyalphabetic ciphers, redeveloped and published somewhat later by the Prussian Friedrich Kasiski. Understanding of cryptography at this time typically consisted of hard-won rules of thumb; see, for example, Auguste Kerckhoffs' cryptographic writings in the latter 19th century. Edgar Allan Poe used systematic methods to solve ci-

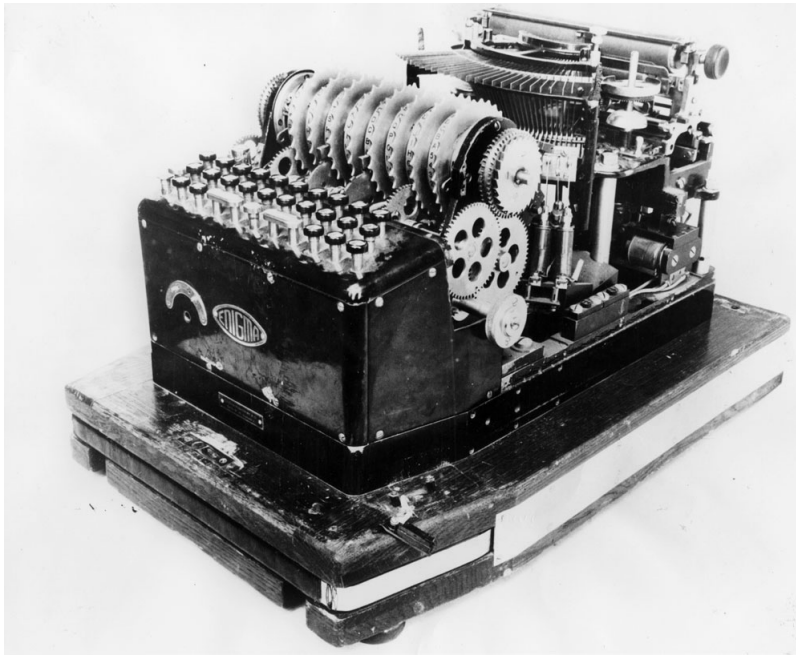
phers in the 1840s. In particular he placed a notice of his abilities in the Philadelphia paper Alexander's Weekly (Express) Messenger, inviting submissions of ciphers, of which he proceeded to solve almost all. His success created a public stir for some months.[24] He later wrote an essay on methods of cryptography which proved useful as an introduction for novice British cryptanalysts attempting to break German codes and ciphers during World War I, and a famous story, The Gold-Bug, in which cryptanalysis was a prominent element.

Cryptography, and its misuse, were involved in the execution of Mata Hari and in Dreyfus' conviction and imprisonment, both in the early 20th century. Cryptographers were also involved in exposing the machinations which had led to the Dreyfus affair; Mata Hari, in contrast, was shot.

In World War I the Admiralty's Room 40 broke German naval codes and played an important role in several naval engagements during the war, notably in detecting major German sorties into the North Sea that led to the battles of Dogger Bank and Jutland as the British fleet was sent out to intercept them. However its most important contribution was probably in decrypting the Zimmermann Telegram, a cable from the German Foreign Office sent via Washington to its ambassador Heinrich von Eckardt in Mexico which played a major part in bringing the United States into the war.

In 1917, Gilbert Vernam proposed a teleprinter cipher in which a previously prepared key, kept on paper tape, is combined character by character with the plaintext message to produce the cyphertext. This led to the development of electromechanical devices as cipher machines, and to the only unbreakable cipher, the one time pad.

During the 1920s, Polish naval-officers assisted the Japanese military with code and cipher development.



Mathematical methods proliferated in the period prior to World War II (notably in William F. Friedman's application of statistical techniques to cryptanalysis and cipher development and in Marian Rejewski's initial break into the German Army's version of the Enigma system in 1932).



1.5 The Math Behind Cryptography

Secrecy systems based on modular arithmetic. The first of these had its origin with Julius Caesar; the newest systems that we will discuss were invented in the late 1970s. In all these systems, we start by translating letters into numbers. We take as our standard alphabet the letters of English and translate them into the integers from 0 to 25, as shown below

| Letter | 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 |
|----------------------|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Numerical Equivalent | 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 |

Of course, if we were sending messages in Russian, Greek, Hebrew, or any other language, we would use the appropriate alphabet and range of integers. Also, we may want to include all ASCII characters, including punctuation marks, a symbol to indicate blanks, and the digits for representing numbers as part of the message. However, for the sake of simplicity, we restrict ourselves to the letters of the English alphabet. The transformation of letters to numbered equivalents can be done in many other ways (including translation to bit strings). Here we have chosen a simple and easily understood transformation for simplicity. First, we discuss secrecy systems based on transforming each letter of the plaintext message into a different letter (or possibly the same) to produce the ciphertext. The encryption methods in these cryptosystems are called character, or monographic, ciphers, because each character is changed individually to another letter by a substitution. Altogether, there are $26!$ possible ways to produce a monographic transformation. We will discuss some particular monographic transformations based on modular arithmetic. Julius Caesar used a cipher based on the substitution in which each letter is replaced by the letter three further down the alphabet, with the last three letters shifted to the first three letters of the alphabet. To describe this cipher using modular arithmetic, let P be the numerical equivalent of a letter in the plaintext and C be the numerical equivalent of the

corresponding ciphertext letter. Then

$$C \equiv P + 3 \pmod{26}$$

The correspondence between plaintext and ciphertext is given below

| | 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 |
|------------|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Plaintext | 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 |
| Ciphertext | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 0 | 1 | 2 |
| | 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 |

To encrypt a message using this transformation, we first change it to its numerical equivalent, grouping letters in blocks of five. Then we transform each number. The grouping of letters into blocks helps to prevent successful cryptanalysis based on recognizing particular words.

Affine Transformation

The Caesar cipher is one of a family of similar ciphers described by a shift transformation.

$$C \equiv P + 3 \pmod{26}$$

where k is the key representing the size of the shift of letters in the alphabet. There are 26 different transformations of this type, including the case of $K \equiv 0 \pmod{26}$ Where letters are not altered, because in this case $C \equiv P \pmod{26}$ More generally, we will consider transformations of the type

$$C \equiv aP + b \pmod{26}$$

Where a and b are integers with $(a, 26) = 1$. These are called affine transformations. Shift transformations are affine transformations with $a = 1$. We require that $(a, 26) = 1$, so that as P runs through a complete system of residues modulo 26, C also does. There are $\phi(26) = 12$ choices for a , and 26 choices for b , giving a total of $12 * 26 = 312$

transformations of this type (one of these is $C \equiv P \pmod{26}$ obtained when $a = 1$ and $b = 0$). If the relationship between plaintext and ciphertext is described above then the inverse relationship is given by

$$P \equiv a^*(C - b) \pmod{26}$$

where a^* is an inverse of $a \pmod{26}$, which can be found using the congruence

$$a^* \equiv a^{\phi(26)-1} = a^{11} \pmod{26}$$

.