MY PROFILE

NAME - KUSHAGRA MEHROTRA

CLASS - CS-MINOR-MARCH

TOPIC - Make A Report On Different Types Of Ciphers With Examples And Screenshots Of The Implementation (You can Use Online Tools To Do Ciphering)

ACKNOWLEDGEMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and I am extremely privileged to have got this all along with the completion of my project. All that I have done is only due to such supervision and assistance and I would not forget to thank them.

I respect and thank Teachnook, for providing me an opportunity to do the project work in "Different types of Ciphers" and giving us all support and guidance which made me complete the project duly. I am extremely thankful for providing such nice support and guidance.

I would not forget to remember our Mentor, Mr.Ahmed, for their encouragement and the spirit to convey knowledge helped me in understanding everything in a Self-Paced mode.

INTRODUCTION

For thousands of years, kings, queens, and generals have relied on efficient com-

munication to govern their countries and command their armies. At the

same time, they have all been aware of the consequences of their messages falling

into the wrong hands, revealing precious secrets to rival nations and betraying vital information to opposing forces. It was the threat of enemy interception that motivoted the development of codes and ciphers: techniques for disguising a message so that only the intended recipient can read it.

The desire for secrecy has meant that nations have operated code-making departments, responsible for ensuring the security of communications by inventing and implementing the best possible codes. At the same time, enemy codebreakers have attempted to break these codes, and steal secrets. Codebreakers are linguistic alchemists, a mystical tribe attempting to conjure sensible words out of meanfewer symbols. The history of codes and ciphers is the story of the centuries-old the battle between code makers and codebreakers, and an intellectual arms race that has had a dramatic impact on the course of history.

In writing The Code Book, I have had two main objectives. The first is to chart the evolution of codes. Evolution is a wholly appropriate term because the development of codes can be viewed as an evolutionary struggle. A code is a conconstantly under attack from codebreakers. When the codebreakers have developed a new weapon that reveals a code's weakness, then the code is no longer useful. It isther becomes extinct or it evolves into a new, stronger code. In turn, this new code thrives only until the codebreakers identify its weakness, and so on. This is anallogos to the situation facing, for example, a strain of infectious bacteria. The bacteria live, thrive and survive until doctors discover an antibiotic that exposes a weakness in the bacteria and kills them. The bacteria are forced to evolve and outwit the antibiotic, and, if successful, they will thrive once again and re-establish themselves. The bacteria are continually forced to evolve to survive the onslaught of new antibiotics.

The ongoing battle between code makers and codebreakers has inspired a whole

series of remarkable scientific breakthroughs. The code makers have continually striven to construct ever-stronger codes for defending communications, while codebreakers have continually invented more powerful methods for attacking them. In their efforts to destroy and preserve secrecy, both sides have drawn upon a diverse range of disciplines and technologies, from mathematics to linguistics, from information theory to quantum theory. In return, code makers and codebreakers have enriched these subjects, and their work has accelerated technological development, most notably in the case of the modern computer.

History is punctuated with codes. They have decided the outcomes of battles and led to the deaths of kings and queens. I have therefore been able to call upon stories of political intrigue and tales of life and death to illustrate the key turning points in the evolutionary development of codes. The history of codes is so inordinately rich that I have been forced to leave out many fascinating stories, which in turn means that my account is not definitive. If you would like to find out more about your favorite tale or your favorite codebreaker then I would refer you to the list of further reading, which should help those readers who would like to study the subject in more detail. Having discussed the evolution of codes and their impact on history, the book's second objective is to demonstrate how the subject is more relevant today than ever before. As information becomes an increasingly valuable commodity, and as the communications revolution changes society, so the process of encoding-ing messages, known as encryption, will play an increasing role in everyday life. Nowadays our phone calls bounce off satellites and our e-mails pass through various computers and both forms of communication can be intercepted with ease, so jeopardizing our privacy. Similarly, as more and more business is conducted over the Internet, safeguards must be put in place to protect companies and their clients. Encryption is the only way to protect our privacy and guarantee the success

of the digital marketplace. The art of secret communication, otherwise known as to cryptography, will provide the locks and keys of the Information Age.

However, the public's growing demand for cryptography conflicts with the needs of law enforcement and national security. For decades, the police and the intelligence services wiretaps used wiretaps to gather evidence against terrorists and organized crime syndicates, but the recent development of ultra-strong codes threatens to undermine wiretaps value of wire-taps. As we enter the twenty-first century, civil libertarians are pressing for the widespread use of cryptography to protect the privacy of the individual. Arguing alongside them are businesses, who require strong cryptography to guarantee the security of transactions within the fast-growing world of Internet commerce. At the same time, the forces of law and order are lobbying governments to restrict the use of cryptography. The question is, which do we value more – our privacy or an effective police force? Or is there a compromise?

Although cryptography is now having a major impact on civilian activities, it should be noted that military cryptography remains an important subject. It has been said that the First World War was the chemists' war because mustard gas and chlorine was employed for the first time, and the Second World War was the physicists' war, because the atom bomb was detonated. Similarly, it has been argued that the Third World War would be the mathematicians' war, because mathmagicians will have control over the next great weapon of war – information. Mathematicians have been responsible for developing the codes that are currently used to protect military information. Not surprisingly, mathematicians are also at the forefront of the battle to break these codes.

While describing the evolution of codes and their impact on history, I have allowed me a minor detour. Chapter 5 describes the decipherment of various an-

ancient scripts, including Linear B and Egyptian hieroglyphics. Technically, cryptography concerns communications that are deliberately designed to keep secrets from an enemy, whereas the writings of ancient civilizations were not intended to be indecipherable: it is merely that we have lost the ability to interpret them. However, the skills required to uncover the meaning of archaeological texts are closely related to the art of codebreaking. Ever since reading The Decipherment of Linear B, John Chadwick's description of how an ancient Mediterranean text was unraveled, I have been struck by the astounding intellectual achievements of those men and women who have been able to decipher the scripts of our ancestors, thereby allowing us to read about their civilizations, religions, and everyday lives.

Turning to the purists, I should apologize for the title of this book. The Code Book is about more than just codes. The word 'code' refers to a very particular type of secret communication, one that has declined in use over the centuries. In a code, a word or phrase is replaced with a word, number or symbol. For example, secret agents have codenames, words that are used instead of their real names in order to mask their identities. Similarly, the phrase Attack at dawn could be replaced by the codeword Jupiter, and this word could be sent to a commander in the battlefield as a way of baffling the enemy. If headquarters and the commander have previously agreed on the code, then the meaning of Jupiter will be clear to the intended recipient, but it will mean nothing to an enemy who intercepts it. The alternative to a code is a cipher, a technique that acts at a more fundamental level, by replacing letters rather than whole words. For example, each letter in a phrase could be replaced by the next letter in the alphabet, so that A is replaced by B, B by C, and so on. Attack at dawn thus becomes Buubdl bu box. Ciphers play an integral role in cryptography, and so this book should have been called The Code and Cipher Book. I have, however, forsaken accuracy for snappiness.

As the need arises, I have defined the various technical terms used within cryptography. Although I have generally adhered to these definitions, there will be occasions when I use a term that is perhaps not technically accurate, but which I feel is more familiar to the nonspecialist. For example, when describing a person's attempting to break a cipher, I have often a used codebreaker rather than the more accurate cipher breaker. I have done this only when the meaning of the word is obvious from the context. There is a glossary of terms at the end of the book. More often than not, though, crypto-jargon is quite transparent: for example, the plaintext is the message before encryption and ciphertext is the message after encryption.

A Codes & Ciphers Primer

* Codes and ciphers have been around for millennia, often proving of vital strategic importance. They are still widespread in the 21st century and are an important component of the operation of the internet. This document provides a brief introduction of codes and ciphers.

3-ROTOR ENIGMA (GVG / PD)

[1] BASIC CONCEPTS

* The oldest means of sending secret messages is to simply conceal them by one trick or another. The ancient Greek historian Herodotus wrote that when the Persian Emperor Xerxes moved to attack Greece in 480 BC, the Greeks were warned by a Greek named Demaratus who was living in exile in Persia. In those days, Demaratus wrote a message on the wooden tablet itself and then covered it with wax, allowing the vital information to be smuggled out of the country.

The science of sending concealed messages is known as "steganography", Greek for "concealed writing". Steganography has a long history, leading to inventions such as invisible ink and "microdots", or highly miniaturized microfilm images that could be hidden almost anywhere. Microdots are a common feature in old spy movies and TV shows. However, steganography is not secure by itself. If someone finds the hidden message, all its secrets are revealed. That led to the idea of obscuring the message that it could not be read even if it were intercepted, and the result was "cryptography", Greek for "hidden writing".

Cryptography takes two forms: "codes" and "ciphers". The distinction between codes and ciphers is commonly misunderstood "code" is essentially a secret language invented to conceal the meaning of a message. The simplest form of a code is the "jargon code", in which a particular arbitrary phrase, for an arbitrary example:

The nightingale sings at dawn.

-- corresponds to a particular predefined message that may not, in fact shouldn't have, anything to do with the jargon code particularly. The actual meaning of this might be:

The supply drop will take place at 0100 hours tomorrow.

Jargon codes have been used for a long time, most significantly in World War II, when they were used to send commands on broadcast radio to resistance fighters. However, from a cryptographic point of view, they're not very interesting. A proper connection would run something like this:

BOXER SEVEN SEEKS TIGER5 AT RED CORAL

This uses "codewords" to report that a friendly military force codenamed BOXER SEVEN is now hunting an enemy force codenamed TIGER5 at a location codenamed RED CORAL. This particular code is weak in that the "SEEK" and "AT" work provide information to a codebreaker on the structure of the message. In practice, military codes are often defined using "code numbers" instead of codewords, listed in a codebook that provides a dictionary of code numbers and their equivalent words. For example, this message might be coded as:

85772 24799 10090 59980 12487

-- where "85772" maps to BOXER SEVEN, "12487" maps to "RED CORAL", and so on. Codewords and code numbers are referred to collectively as "code groups". The words they represent are referred to as "plaintext" or, more infrequently, "clear", "plain code", "placode", or "plaindata".

Codes are unsurprisingly defined by "codebooks", which are dictionaries of code groups listed with their corresponding plaintext. Codes originally had the codegroups in the same order as their plaintext. For example, in a code based on code numbers, a word starting with "a" would have a low-value code number, while one starting with "z" would have a high-code number. This meant that the same codebook could be used to "encode" a plaintext message into a coded message or"context", and "decode" a code text back into a plaintext message.

However, such "one-part" codes had certain predictability that made it easier for outsiders to figure out the pattern and "creates" or "break" the message, revealing its secrets. To make life more difficult for

codebreakers, code makers then designed codes where there was no predictable relationship between the order of the code groups and the order of the matching plain te. This meant that two codebooks were required, one to look up plaintext to find code groups for encoding, the other to look up code groups to find plaintext for decoding. This was in much the same way that a student of a foreign language, say, French, an English-French, and a French-English dictionary to translate back and forth between the two languages. Such "two-part" required more effort to implement and use, but they were harder to crack.

* In contrast to a code, a "cipher" conceals a plaintext message by replacing or scrambling its letters. This process is known as "enciphering" and results in a "ciphertext" message. Converting a ciphertext message back to a plaintext message is known as "deciphering". Coded messages are often enciphered to improve their security, a process known as "superencipherment".

There are two classes of ciphers. A "substitution cipher" changes the letters in a message to another set of letters, or "cipher alphabet", while a "transposition cipher" shuffles the letters around. In some usages, the term "cipher" always means "substitution cipher", while "transpositions" are not referred to as ciphers at all. The term "cipher" will mean both substitution ciphers and transposition ciphers in this document. It is useful to refer to them together since the two approaches are often combined in the simplicity cipher scheme. However, transposition ciphers will be referred to in specific as "transpositions" for simplicity.

"Encryption" covers both encoding and enciphering, while "decryption" covers both decoding and deciphering. This is also in the term "crypto text" to cover both code text and ciphertext, though the term "encode" is sometimes seen instead. The science of creating codes and ciphers is known, as mentioned, as "cryptography", while the science of breaking them is known as "cryptanalysis". Together, the two fields make up the science of "cryptology".

[2] SIMPLE SUBSTITUTION CIPHERS

* A simple substitution cipher in which the same cipher letter is always exchanged for the same plaintext letter is known as a "monoalphabetic substitution cipher". For example, we could define a cipher alphabet as follows:

plaintext alphabet: abcdefghijklmnopqrstuvwxyz

ciphertext alphabet: TDNUCBZROHLGYVFPWIXSEKAMQJ

Given the plaintext:

erase the tapes

-- and the cipher alphabet above, we get:

CITXC SRC STPCX

Note that in this example plaintext is printed in lowercase, while ciphertext is printed in uppercase. This convention will be followed in the rest of this document.

Monoalphabetic substitution ciphers go back to at least the fourth century BC. One of the simplest monoalphabetic substitution ciphers, known as a "Ceasar shift", associated with Julius Ceasar, involves shifting letters by several positions, say through:

plaintext alphabet: abcdefghijklmnopqrstuvwxyz

cipher alphabet: XYZABCDEFGHIJKLMNOPQRSTUVW

Using this cipher alphabet, Alice can convert the plaintext:

beware the ideas of march

-- into the ciphertext:

YBTXOB QEB FABP LC JXOZE

This can be made even more cryptic by removing the spaces:

YBTXOBQEBFABPLCJXOZE

-- and it still remains more or less readable when translated back to plaintext:

bewaretheidesofmarch

With 26 letters in the English alphabet, there are of course 25 different possible Ceasar shift cipher alphabets. All Bob needed, to read the cipher is a number from 1 to 25 to define the shift. This number can be thought of as a "key" associated with the Cesar shift enciphering "algorithm".

A Ceasar shift cipher is ridiculously easy to crack since all one has to do is try all 25 Ceasar shift cipher alphabets until one works. Interestingly, however, it is still in use, in the form of the "rot13" scheme used on internet forums, which is a 13-place Ceasar-shift used to hide the punchlines of jokes and the like.

* A more secure way to build a substitution cipher is to completely mix up the mappings between the plaintext and ciphertext alphabets. There are a vast number, trillions of trillions, of possible ways to scramble the alphabet, and such a scrambling might seem on the face of it very secure.

One way to come up with a mixed cipher alphabet is for Alice to take a "keyphrase" consisting of, say, a name, such as RICHARD MILHAUS NIXON, write it down while eliminating any redundant letters, and then complete the cipher by writing down the remaining letters of the alphabet in alphabetical order:

plaintext alphabet: abcdefghijklmnopqrstuvwxyz

cipher alphabet: RICHADMLUSNXOBEFGJKPQTVWYZ

This is a simple cipher algorithm, but even if a codebreaker knows that this general scheme was used, the message still cannot be read without the keyphrase, and a brute-force approach to cracking it is very difficult. This is a fundamental principle of cryptography, stated by a 19th-century Dutch linguist & cryptographer, Auguste Kerckhoffs von Niewenhof), and known as"Kerckhoffs' Principle": The security of a cipher should not depend on an enemy's ignorance of the enciphering algorithm, on enemy's ignorance of the key. Codebreaking is often focused on discovering keys.

[3] SIMPLE TRANSPOSITIONS

* For an example of a transposition, suppose Alice wants to send Bob the message:

meet me near the clock tower at twelve midnight tonite

One way to transpose this message is for Alice to "write-in" the words vertically in five rows without any spaces as follows

m e e t m

e n e a r

t h e c l

o c k t o

w e r a t

t w e l v

e m i d n

i g h t t

o n i t e

-- and then "read off" each column, top to bottom, as follows:

metowteio enhcewmgn eeekreihi tactaldtt mrlotvnte

METOWTEIOENHCEWMGNEEEKREIHITACTALDTTMRLOTVNTE

Bob then "writes" the message in five parts:

M E T O W T E I O

E N H C E W M G N

E E E K R E I H I

T A C T A L D T T

M R L O T V N T E

-- and then "reads off" the message from the columns:

MEET ENEAR THECL OCKTO WERAT TWELV EMIDN IGHTT ONITE

MEETMENEARTHECLOCKTOWERATTWELVEMIDNIGHTTONITE

meet me near the clock tower at twelve midnight tonite

Transposition of the form shown above is extremely easy to crack. Holmes just writes down the letters of the transposition rows, increasing the length of the rows until he sees something that makes sense. However, Alice could make things more doable for a codebreaker a bigger headache by reading off columns in an alternating "down" and "up" fashion, or by reading off the transposition in a "spiral" pattern -- "down" on the left side, "right" across the bottom, "up" on the right side, "left" across that the second-to-left column, "down" again, and so on until all letters were transposed. Even more sophisticated transpositions "checkerboard" pattern -- such as a "knight's tour", a grid of numbers that specify the sequence of movements of a chess knight across the grid.

[4] FREQUENCY ANALYSIS AGAINST CIPHERS

* Given a large number of possible monoalphabetic substitution cipher alphabets, it might seem like a substitution cipher be very hard to break. In reality, it's very easy if given a reasonably large ciphertext message to analyze, but it took over a

thousand years to figure out how.

The basic approach for cracking a monoalphabetic substitution cipher was invented by a multi-talented medieval Arabic scholar named al-Kindi, and is now known as "frequency analysis". His work was an outgrowth of efforts by Arabs to perform text analyses of religious texts to see if they were written by the Prophet.

Frequency analysis is a statistical method. In every language, some letters are used on average more than others, and the percentages of letters in different languages tend to be constant. For example, in English text, the three most common letters in the average are "e", "t", and "a", while the three least common are "x", "q", and "z". (Average letter frequencies will differ in different languages.) This means that if a codebreaker sees that "O", "G", and "B" are the most common letters in ciphertext they are likely to represent "e", "t", and "a".

Frequencies of characters in any text may deviate from the average, so this mapping may not be perfectly accurate. However frequency analysis can also be performed on pairs of letters, or "digraphs", in the ciphertext -- "ee" is common in English text not "aa". A codebreaker can also spot common triplets, or "trigraphs", or entire words -- "the" is the most common word in English text. Text may have predictable elements -- for example, Nazi correspondence might start with "Heil Hitler!" Such predictable phrases are known as "cribs". Military correspondence tends to follow standard formats and is often loaded with cribs.

Once given these clues, a codebreaker can then use educated guesswork to "fill in the blanks". For example, an incomplete of the form "-u-m-RI-e" in-text sent from a naval base is likely to be "submarine". This is the same skill that is used to solve crossword puzzles, and is known to cryptologists

as "anagramming". The usage of the word in cryptology is somewhat diff from the popular usage, which refers to a scrambling of the letters of one word into a different word.

If the frequency analysis of a ciphertext shows a seeming match to normal text, then the cipher is likely to be a transposition frequency analysis seems to show a different mapping of characters from normal text, but trying to plug the proper character back into the ciphertext gives absolutely no useful results, then the cipher is likely a combination substitution and transpositive course, frequency analysis requires enough text to permit the construction of a useful table of letter frequencies. This is a general truth of cryptanalysis: the more crypto text available, the easier the crypto text is to crack, and on the other side of the short or fragmentary crypto texts can be difficult to crack even if they used an insecure cryptographic scheme.

[5] CRACKING CODES / CODES VERSUS CIPHERS

* Solving a monoalphabetic substitution cipher is easy. Solving even a simple code is difficult. Decrypting a coded message little like trying to translate a document written in an alien language, with the task amounting to building up a "disco of the code groups and the plaintext words they represent.

One finger hold on a simple code is the fact, mentioned in the previous section, that some words are more common than other such as "the" or "a" in English. In telegraphic messages, the code group for "STOP" (end of the sentence) is usually very commohelps define the structure of the message in terms of sentences, if not their meaning.

Further progress can be made against a code by collecting many messages encrypted with the same code and then obtaining intelligence background on the messages, such as the location from where a message was sent, and where it was being sent time the message was sent; events occurring before and after the message was sent, and the normal habits of the people send the coded messages. Cribs can be helpful as well.

Various tricks can be used to "plant" or "sow" information into a code, for example by executing a raid at a particular time allocation against an enemy, and then examining code messages in response to the raid. Coding errors are a particularly usefufingerhold in a code, and naturally, people are bound to make errors, sometimes disastrous ones, sooner or later. Of course, planting information and exploiting errors works against ciphers as well.

* The most obvious and, in principle at least, the simplest way of cracking a code is to steal the codebook through bribery, and burglar raiding parties. Constructing a new code is like building a new language and writing a dictionary for it, which is a big job code that is compromised, the whole task has to be done all over again, and that means a lot of work for both cryptographers and code users. In practice, when codes were in widespread use, they were usually changed periodically.

Once codes have been created, their distribution is logistically clumsy, and the probability that the code will be compromised is high. In contrast, the security of ciphers is, as mentioned earlier, generally dependent on protecting the cipher keys. Ciphers can be stolen and people can betray them, but they are much easier to change and communicate.

[6] THE VIGENERE CIPHER

* The West learned about frequency analysis in the 15th century, forcing the development of better encryption schemes:

❑ One was the "nomenclator", which was a substitution cipher combined with a set of codewords.

❑ Another trick was to use "nulls", or unused symbols, in ciphers. Suppose that a cipher uses the numbers 00 through 9represent text. Even if numbers and punctuation are enciphered, that leaves an unused subset of numbers, and these ununumbers or nulls could be littered through the ciphertext. They were simply ignored when the text was deciphered.

❑ Yet another trick was the "homophonic substitution cipher". Given a cipher based on substituting the numbers 00 thro99 for text, there was no reason that multiple numbers couldn't be used to match a single common character, such as "help defeat frequency analysis.

One of the most significant developments was the "polyalphabetic substitution cipher", which was described in its definitive paper published in 1586 by a French diplomat named Blaise de Vigenere. A "Vigenere cipher" uses 26 substitutions cipherorganized using a "Vigenere square" as shown below, with some spacing added here to make it legible:

a bcd efgh ijk lmno pqr stuv wyxz

01: A BCD EFGH IJK LMNO PQR STUV WXYZ

02: B CDE FGHI JKL MNOP QRS TUVW XYZA

03: C DEF GHIJ KLM NOPQ RST UVWX YZAB

04: D EFG HIJK LMN OPQR STU VWXY ZABC

05: E FGH IJKL MNO PQRS TUV WXYZ ABCD

06: F GHI JKLM NOP QRST UVW XYZA BCDE

07: G HIJ KLMN OPQ RSTU VWX YZAB CDEF

08: H IJK LMNO PQR STUV WXY ZABC DEFG

09: I JKL MNOP QRS TUVW XYZ ABCD EFGH

10: J KLM NOPQ RST UVWX YZA BCDE FGHI

11: K LMN OPQR STU VWXY ZAB CDEF GHIJ

12: L MNO PQRS TUV WXYZ ABC DEFG HIJK

13: M NOP QRST UVW XYZA BCD EFGH IJKL

14: N OPQ RSTU VWX YZAB CDE FGHI JKLM

15: O PQR STUV WXY ZABC DEF GHIJ KLMN

16: P QRS TUVW XYZ ABCD EFG HIJK LMNO

17: Q RST UVWX YZA BCDE FGH IJKL MNOP

18: R STU VERY ZAB CDEF GHI JKLM NOPQ

19: S TUV WXYZ ABC DEFG HIJ KLMN OPQR

20: T UVW XYZA BCD EFGH IJK LMNO PQRS

21: U VWX YZAB CDE FGHI JKL MNOP QRST

22: V WXY ZABC DEF GHIJ KLM NOPQ RSTU

23: W XYZ ABCD EFG HIJK LMN OPQR STUV

24: X YZA BCDE FGH IJKL MNO PQRS TUVW

25: Y ZAB CDEF GHI JKLM NOP QRST UVWX

26: Z ABC DEFG HIJ KLMN OPQ RSTU VWXY

a bcd efgh ijk lmno pqr stuv wyxz

This defines 26 different Ceasar shift ciphers, each of which is weak in itself but which in combination result in a much morcipher. The idea in the Vigenere cipher is to use a cipher key to select different cipher alphabets in succession as letters are

enciphered. Suppose Alice wants to encipher the phrase:

use the force of luke

-- with a Vigenere cipher, using the cipher keyword "WARTHOG". All she has to do is scan down the square defined abovmatch the cipher alphabet letter to a particular row, then select the cipher character matching the plaintext letter for that row W: row 23 gives u -> Q

A: row 01 gives s -> S

R: row 18 gives e -> V

T: row 20 gives t -> M

H: row 08 gives h -> O

O: row 15 gives e -> S

G: row 07 gives f -> L

W: row 23 gives o -> L

A: row 01 gives r -> R

R: row 18 gives c -> T

T: row 20 gives e -> X

H: row 08 gives l -> S

O: row 15 gives u -> I

G: row 07 gives k -> Q

W: row 23 gives e -> A

This gives:

QSV MOS LLRTX SIQA

Simple frequency analysis cannot crack a Vigenere cipher, and the number of possible keys is so great that finding the right trial-and-error is effectively impossible. Despite the simplicity and elegance of the Vigenere cipher, it was mostly ignored next several centuries, since it was regarded as too cumbersome for general use.

[7] TELEGRAPHY & CRYPTOLOGY / FRACTIONATING CIPHERS

* The invention of telegraphy in the mid-19th century revolutionized communications, and also had a significant impact on cryptology. Since telegrams were paid for on a word-by-word basis, there was an incentive to reduce the word count, and so"commercial codes" were developed from the outset to provide what might be called in modern terms "data compression". Talso provided some security against casual reading, though they were one-part codes and not generally intended to provide

security. Commercial codes did generally offer superencipherment schemes for users who wanted more secrecy.

Commercial codes remained in use until well after the First World War. Some of the entries in a commercial code published1920s seemed almost tailored for encrypting contemporary pulp fiction:

BUKSI Avoid arrest if possible.

OBNYX Escape at once.

PYTUO collided with an iceberg.

YBDIG Plundered by natives.

CULKE Bad as possibly can be.

The telegraph had a particular impact on military operations, contributing along with the invention of the locomotive and firearms to a revolution in warfare. However, the telegraph was vulnerable to interception, meaning that good cryptographic schemes were required to ensure security. Codes really couldn't do the job for an army on the move in the field, because of the logistical difficulty of handling the codebooks. The Vigenere cipher came into common use, assisted by a simple invention "cipher disk", described below. However, although the Vigenere cipher had been regarded as indecipherable for a long timecrytanalysts were finally able to get a handle on it. For example, if multiple messages were encrypted using a Vigenere cipher and the same cipher key was intercepted, frequency analysis could be used across the first letter of all the messages in parallel, the second letter, and so on.

* Other cipher schemes were developed for military telegraphic use. During the American Civil War, Union forces used a "cipher", basically a transposition cipher based on entire words, not characters. A small set of codewords were implemented to conceal words that might provide too much of a clue to snoopers, with cipher books distributed describing various transposipatterns and the codewords to use.

A general technique for cracking transpositions called "multiple anagramming" was developed in the 1870s. It requires two messages that have been transposed in the same way -- for a simple example, consider two messages consisting only of one letter word each:

TASPR

PRSCO

A little examination shows that both of these two transpositions could be unscrambled into two different words:

TASPR -> PARTS, TRAPS

PRSCO -> CROPS, CORPS

Given each message on its own, there's no way to figure out which of the two unscrambling would be correct for each mesHowever, performing the same unscrambling on both messages in parallel shows that there's only one unscrambling that yieintelligent answer for both messages:

TASPR -> TRAPS PRATS PARTS

PRSCO -> PORCS CORPS CROPS

12345 15243 45213 42513

* Several new cipher techniques were also developed based on the "Polybius square" or "checkerboard". This was a converting of the letters into pairs of numbers, devised in classical Greek times by a scholar named Polybius. Updated into EnglishPolybius square consists of the letters of the alphabet arranged as a 5-by-5 checkerboard grid with rows and columns numbered "I" and "J" assigned to the same grid location:

1 2 3 4 5

1 A B C D E

2 F G H I/J K

3 L M N O P

4 Q R S T U

5 V W X Y Z

The row-and-column index numbers are then substituted for letters. For example, "help" becomes "23 15 31 35". For added security, the arrangement of the letters in the grid can be scrambled, preferably by using a cipher keyword to determine the

scrambling. For example, if we use the key "RICHARD MILHAUS NIXON", we get a checkerboard that looks like this:

1 2 3 4 5

1 R I/J C H A

2 D M L U S

3 N X O B E

4 F G K P Q

5 T V W Y Z

No matter how the grid is arranged, however, as stated it's just a monoalphabetic substitution cipher, with pairs of ciphertext in place of plaintext letters. It offers little security, though it can be used as a basis for "semaphore" codes, in which a signal holds two flags in five different positions each to send the full alphabet. However, the checkerboard has a much more potential plaintext letter is represented by two ciphertext digits, and that gives a new option for encrypting the message.

Suppose Alice has a checkerboard based on the key "RICHARD MILHAUS NIXON" as above, and the plaintext:

down with big brother

She can obtain row-and-column indexes for the letters in this plaintext from the grid as follows:

d: row = 2 / column = 1

o: row = 3 / column = 3

w: row = 5 / column = 3

n: row = 3 / column = 1

...

Now she divides the plaintext into blocks of, say, five letters, and writes the indexes vertically underneath the letters:

downw ithbi gbrot her

23535 15131 43135 131

13313 21442 24131 451

To produce the final encryption, she concatenates the two rows on a block-by-block basis:

2353513313 1513121442 4313524131 131451

The message is divided into blocks to make encryption and decryption more manageable; there's no particular reason to use a block size of five, any other reasonable value will work as well. This particular scheme breaks down each letter into two cip

letters, a concept known as "fractionation"; and then it breaks the message down into blocks and concatenates the two have a concept known as "seriation". The resulting message cannot be cracked by simple frequency analysis. This scheme is known "bifid" cipher and was introduced by a Frenchman named Felix Marie Delastelle in a book published in 1901.

It is also possible to "map" plaintext letters to three or more digits that can then be transposed, and such a three-digit mapping is called a "trifid" cipher. Some variations on fractionating ciphers are extremely devious and hard to crack.

[8] ONE-TIME PAD CIPHER

* The First World War introduced wireless telegraphy into military operations, greatly extending the communications revolt begun by the telegraph. However, radio communications placed new demands on security, since it was so very easy to intermessages sent over the airwaves.

Since warships at sea could now communicate with shore stations over wireless, navies adopted code systems. Naval codes were often bound with metal plates in the covers so they could be thrown overboard in an emergency and sink. Since the world on land quickly bogged down on land into immobile trench warfare, armies also adopted codes for frontline operations. This relatively little problem in distributing the codebooks, the armies not being on the move, and "trench codes" came into common use.

Sophisticated cipher systems were also developed. One of the most significant achievements in cryptography in the First WWar was a cipher that was, and remains, uncrackable even in principle. As is

typical of black magic, it had a major catch.

A Vigenere cipher based on a keyword becomes more secure with a longer keyword. It also becomes more secure if the

keyword is made more unpredictable, to prevent a codebreaker from using guesswork to determine the keyword. That means that the most secure possible Vigenere cipher is one where the keyword is as long as the message, with the keyword made completely unpredictable random characters.

To implement such a cipher, sets of random keys could be printed on a pad of paper, with each page on the pad used once then thrown away. As a result, this scheme is called a "one-time pad" cipher. The interesting thing about the one-time pad is logically impossible to crack by analytic means. Imagine being given a page of completely random letters: it's impossible "crack" because there's no message there, it's just noise. Taking a message and changing all the characters at random also rein a page of noise, and it's just as uncrackable. The only way the message can be extracted is by telling the recipient of the

a message with a one-time pad what random changes were made to the letters in the text.

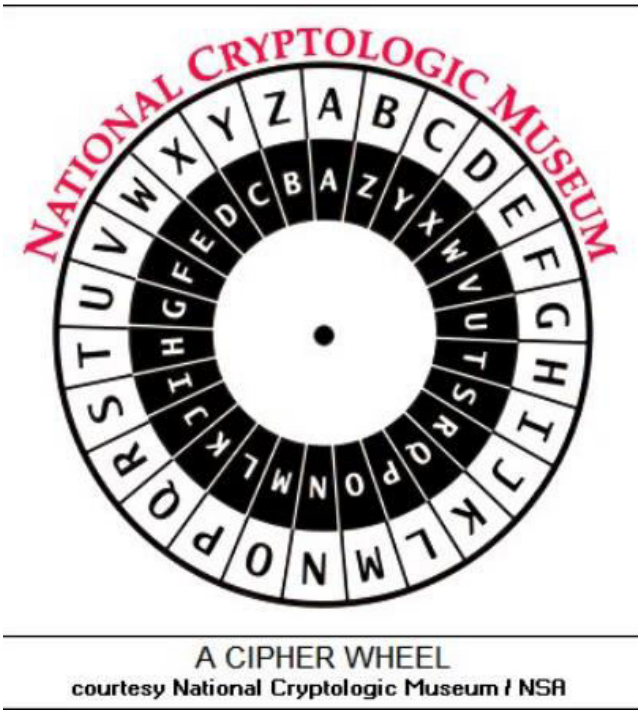
A one-time pad cipher has an interesting property: it is not so much true that no signal can be extracted from pure noise, but any signal can be extracted from noise. Since there's no fixed pattern in the ciphertext or the key, a key can be easily synthetic to produce every possible message that will fit into the number of letters, such as instructions to attack the enemy, a shoppinor dirty limericks.

* The catch to the one-time pad cipher is that a specific random key can only be used once to encipher one message. If two messages are enciphered using the same key, then it is no longer impossible to crack the cipher, since a (painful) exercise in multiple anagramming could be used to decipher the texts, trying every possible key on both messages until both messages are revealed. This makes the one-time pad cipher logistically clumsy to deal with, worse than a code, requiring distribution of updates to every user -- and so it is only used in very high-security communications.

[9] CIPHER MACHINES

* The "cipher disk", mentioned above, was one of the first cipher machines, invented in the 15th century. The cipher disk coof two nested disks, including an outer disk labeled with all the letters of the alphabet, and an inner disk also labeled with

letters of the alphabet -- but not necessarily in the same order. The outer disk defines the plaintext alphabet, while the inner defines a monoalphabetic substitution cipher alphabet.



A CIPHER WHEEL
courtesy National Cryptologic Museum / NSA

Suppose Alice takes her cipher disk and rotates the inner disk so that the letter "A" lines on the inner disk line up with, the said letter "q" on the outer disk. She can then encipher a message by taking each letter, looking up the plaintext letter on the route, and then writing down the matching ciphertext letter next to it on the inner disk. The cipher disk can be also made in the for"slide", with two alphabets written on strips of cardboard, or whatever, that can be slid next to each other. Of course, each should have two repeating alphabets to permit them to be read conveniently when offset.

The cipher as described remains a feeble monoalphabetic substitution cipher, which is not only easy to crack but also so eause that the cipher disk hardly seems very handy. Where the cipher disk comes in useful is in dealing with a Vigenere cipher example, suppose Alice wants to encrypt a message using a Vigenere cipher with the cipher keyword "WARTHOG". Tencipher the first letter, she moves the "W" on the inner disk to match up with the "a" on the outer disk and then finds the

ciphertext letter on the inner disk matching the desired plaintext letter on the outer disk. Next, Alice enciphers the second lethe inner wheel by moving the "A" on the inner disc to match the "a" on the outer disk, then looking up the ciphertext letter inner disk that matches the desired plaintext letter on the outer disk. She repeats this process for the rest of the keyword and starts all over again at the beginning of the keyword.

* The cipher disk led to a more the sophisticated cipher machine, invented late in the 19th century by French military officer Etienne Bazeries, and known as the "Bazeries cylinder". (The basic idea was invented by American President ThomJefferson a century earlier, but then forgotten.) A Bazeries cylinder consists of a set of roughly 20 to 30 numbered disks, a different cipher alphabet on the edge of each disk, and a hole in the center of the disks to allow them to be stacked on an axle disk is removable and

can be mounted on the axle in any order desired. The order of the disks can be considered the cipher key for the Bazeries cylinder, with both Alice and Bob arranging the disks in the same predefined order.

To encrypt a message, Alice rotates the disks to produce the plaintext message along one "row" of the stack of disks and selects another row as the ciphertext. To decrypt the message, Bob rotates the disks on his cylinder to produce the ciphertexalong a row. It is handy if both Alice and Bob know the offset of the row, but not necessary since Bob can simply look around the cylinder to find a row that makes sense. It was also possible to implement the same scheme with slides containincipher alphabets on a frame. Both schemes were used in World War I and well into World War II; they provided strong encryption, but they were hardly indecipherable.

SIMPLE CIPHER DEVICES
(GVG / PD)

* After World War I, several much more sophisticated cipher machines were built, the most famous being the "Enigmapatented in 1918 by a German engineer named Arthur Scherbius.

As it emerged, the Enigma was a wooden box with a keyboard and a bank of lettered lights corresponding to the keys. To

encrypt a message, a plaintext character was typed in, and after scrambling the appropriate light was turned on to give the ciphertext character. The ciphertext was then sent using Morse code over radiotelegraph. The operation of the machine was "reciprocal", in that if the "A" key was pressed and lit up the "Z" lamp, then pressing the "Z" key would light up the "A" is meant that at the receiving end, as

long as the operator had an Enigma machine set up in the same way, he could just type in the ciphertext character to get the plaintext character: the procedure for encryption and decryption was the same.

The scramblings between the input and output were performed by what was called a "rotor" system. A rotor was a wheel with electrical contacts on each side and scrambled wiring linking the contacts on the two sides. The scrambled wiring represented a mixed substitution cipher alphabet. There were three rotors in series fed into one another, with this assembly referred to as a"basket".

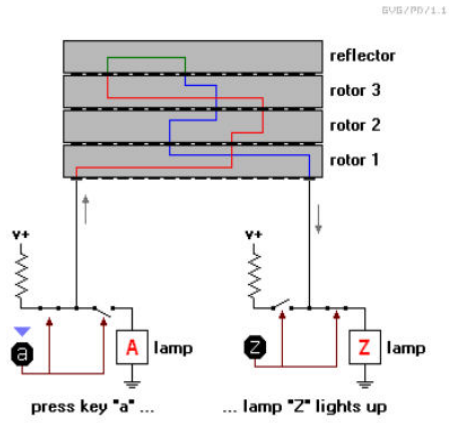
Every time a key was pressed, the first rotor would turn one position. Every time the first rotor went full circle (after 26

keypresses), the second rotor would turn one position. Every time the second rotor went full circle (after 26 * 26 = 676

keypresses), the third rotor would turn one position. When the third rotor went full circle (after 26 * 26 * 26 = 17,576

keypresses), the sequence would start all over again.

The third rotor was wired into a disk called the "reflector" that simply rerouted the connections on the output of the third disinto themselves. In operation, the electrical signal from the input key ran up through the basket of rotors to the reflector, and back down through the basket to the output lamp. This is how the Enigma achieved reciprocal operation, allowing both entrants to decryption using the same configuration.



RECIPROCAL OPERATION OF THE ENIGMA

Each of the rotors had a different wiring scheme. The three rotors were removable and could be rearranged in six different It was also possible in principle to have a stockpile of more than three rotors and select three from the set for use, further multiplying the possibilities. The Enigma also had a plugboard that allowed patch cords to be swapped around to mix six sets of letters. Setup of the Enigma involved selecting the order of the three rotors, selecting their starting positions, and seethe the patch cord arrangement. This setup amounted to the "key" or "day key" for the Enigma cipher, and there were a vast number of impossible day keys.

* The German military standardized the Enigma for encryption before World War II. Despite the complexity of the Enigcipher, Polish cryptanalysts were able to figure out patterns in the operation of the Enigma that allowed them to determine the thdaykey, using a set of Enigma simulators known for some obscure reason as a "bombe".

However, in 1938 the Germans improved their Enigma security by providing five rotors, increasing the number of possible configurations of three rotors selected from the set, and increasing the number of pairs of letters swapped by the plugboard greatly multiplying the possible numbers of day key configurations, demanding a much more complicated and expensive to figure out the settings. By this time, Nazi dictator Adolf Hitler was making threatening noises against Poland, so they got in touch with British intelligence to tell them how Enigma had been cracked.

The Nazis invaded Poland in September 1939, setting off World War II. They quickly overran the country. The British were hard-pressed by defeats early in the war, and to help fight back set up a sophisticated codebreaking operation at an estate in London named Bletchley Park. Although the Germans had tightened up their procedures for the Enigma operation, eliminatin' loopholes that the Poles had exploited to ferret out day keys, Bletchley Park cryptanalysts -- most significantly a brilliant Oxmathematician named Alan Turing -- were able to develop a new, more sophisticated bombe system to once again crack Enigma.

[10] MODERN CIPHER SYSTEMS

* A range of cryptographic schemes were used in World War II, including traditional military codes, much like the naval and trench codes used in World War I, as well as the Bazeries cylinder and slide schemes. The war also saw considerable us"telecipher" systems for high-level secure communications.

A telecipher system was an encrypted teletypewriter. Teletypewriters used a sequence of on-off electrical signals to transmit-receive letters, using a scheme known as the "Baudot code". If the "on" signal is represented as a "1" and the "off" signal

represented as a "0", then the Baudot code had the form of a set of "binary" numbers as follows:

00011 A -

11001 B ?

01110 C :

...

This scheme uses five "binary digits" or "bits" per letter; there were "shift" characters to permit the use of a secondary set of punctuation characters.

A telecipher machine scrambled the "stream" of Baudot plain data using what was known in the old days as "modulo-two

arithmetic" and is now known as an "exclusive OR" or "XOR" operation. The basic rules of the XOR operation are as follow 0 XOR 0 = 0

0 XOR 1 = 1

1 XOR 0 = 1

1 XOR 1 = 0

In simpler terms, if two bits are XORed that have the same value, the result is a "0". If two bits are XORed that have a diffevalue, the result is a "1". XOR can be performed with multibit values on a bit-by-bit basis -- for example, two five-bit values and Y can be XORed to give a 5-bit value Z:

X: 10011

Y: 00110

_____

Z: 10101

The XOR operation has the interesting property of being "reversible", in that if two binary values X and Y are XORed together the result Z can be XORed with Y to give X again:

Z: 10101

Y: 00110

_____

X: 10011

A telecipher machine used an arrangement of electromagnetic relays to XOR the plaindata with a stream of seemingly randoas a "mask":

data: 10100 00001 00110 10010 00100 10100 00110 10000 10010 00001 01010

mask: 01001 00101 11011 10110 00111 11001 10101 01010 10000 11001 10001

_____

cipher: 11101 00100 11101 00100 00011 01101 10011 11010 00010 11000 11011

The enciphered binary stream could be deciphered by simply XORing it again with the same mask.

The mask was generated by an arrangement of rotating "wheels", with sets of pins on them that could be set in or out. They would generate very long streams of bits unpredictably, and telecipher machine messages could be very hard to crack. The British developed one of the first electronic computers, named "Colossus" and built with vacuum tubes, german teleciphers messages.

* In the postwar period, electronic computers became much more widespread, and modern cryptography is effectively based on computer-based ciphers. Two classes of "digital" ciphers were developed: "stream" ciphers and "block" ciphers.

The old telecipher systems essentially implemented stream ciphers. A stream cipher operates by taking a stream of plain data and XORing them on a bit-by-bit basis with a mask of more or less random bits. The cipher stream can then be deciphered by boring it again with the same mask. The computer generates the mask using some variation of the "pseudo-random" bitstream

generation algorithm. Block ciphers, in contrast, break up the stream of data bits and perform a complicated set of shuffling or operations on each block. In either case, a "key" consisting of several bits is used to specify encryption, and decrying the 1970s, the US National Bureau of Standards (NBS, now the National Institute of Standards & Technology / NIST)

created a specification for a "Data Encryption Standard (DES)", using a 56-bit block cipher scheme. DES became the standard most of the rest of the century. Modern digital ciphers will use a key at least 100 bits long.

By the 1970s, the use of computing and computer networking was beginning to take off, and some cryptologists were looking

forward to the future. One of the issues that were considered the "key exchange" problem. All ciphers

up to that time we"symmetric", meaning the same key was used to encrypt and decrypt a message. However, symmetric ciphers posed some

problems in the emerging world of computer networking. In modern terms, suppose customers are visiting a business web want to establish secure communications, say to make a credit-card payment. In that case, the business would have to each different keys with each customer to guarantee security, which would be very cumbersome.

In 1975 a cryptologist named Whitfield Diffie at Stanford University in California came up with an idea that seems almost in hindsight but must have sounded preposterous at the time: "public key" cryptography. The idea was that a public key clip would have two keys: a "public" key, available to anyone, that can be used to encipher a message; and a "private" key, known only to one person, that can be used to decipher a message. The bizarre thing about a public key cipher is that the public key is used to encipher a message, but once that message is enciphered, only the private key can decipher the message.

In public-key cryptography, if Bob wants to send an encrypted message to Alice, he obtains her public key. Alice could put a public key on her website or business card if she liked for anyone to use. Bob uses Alice's public key to encipher a message and then sends it off. He doesn't need to worry about anyone reading the message since the public key can't be used to decrypt only Alice can decrypt it, using her private key.

Diffie had no idea of how to do this, but in 1975 he and his colleagues published the concept anyway to inspire other

cryptologists. About two years later, three researchers at the Massachusetts Institute of Technology (MIT) computer science department named Ron Rivest, Adi Shamir, and Ron Adleman announced that they had developed a public key cipher, named"RSA" after their initials. RSA, and other public-key ciphers that followed, had limitations in that they required a lot of composers and were relatively easy to crack. As a result, they weren't used to encipher long messages, instead of being used as a key for a block cipher that was used to send the actual message.

* One of the interesting things that could be done with public-key cryptography was "message authentification" -- that is, pra guarantee that a message was sent by the person who was supposed to have sent it. Normally, a public-key clip message is encrypted with the public key and decrypted with the private key. It also works the other way around: a message is encrypted with the private key and decrypted with the public key. Such an "inside-out" encryption, to give it a name, is so absurd, since anyone could decrypt the message, but the point is that only the person with the private key could have encrypted validate a message, a "hash" can be made of the message. A hash is a relatively short string of bits that is created from the message, with the interesting property that the slightest change in the message will result in a very different hash. Alice can encrypt a message to Bob using normal cryptographic methods, and then can send a hash to Bob using inside-out encryption extracts the hash with Alice's public key, and then performs a hash on the decrypted message to see if it matches the has sent by Alice. If Alice didn't send the message, Bob won't be able to decipher it; if the message has been tampered wiBob's hash won't match.

There's another aspect of message authentification. Suppose Bob hasn't heard from Alice in years, has lost track of her, and starts getting messages from her again. How does he know the messages are really from Alice and not somebody, say Zeldapretending to be Alice? Zelda can hand him her public key and Bob would be no wiser.

If Bob wants to make sure Alice is Alice, he can obtain a "digital certificate" or "cert" from a "certification server". Secure network servers operating on a "trusted third party" basis that store the public keys associated with specific users, and provide them on demand. Bob can get the cert associated with Alice from a certification server and then check to see if the key he has matches that in Alice's cert.

* Modern digital encryption schemes are commonly used on the internet. For one particularly important example, most peopwho surf the internet to make online purchases have used cryptography to provide their credit card numbers to vendors, using the "Secure Sockets Layer (SSL)" protocol. Pages protected by SSL are designated with an "HTTPS" prefix instead of the conven"HTTP" prefix.

SSL works more or less transparently to the user and uses both public-key and symmetric ciphers. Suppose Alice wants to purchase from an online vendor and needs to give the vendor her credit-card number. Her web browser will downvote the vendor's public key, validate it against the vendor's cert -- a web browser usually maintains a list of certs for prominent businesses -- and use it to encrypt a secret symmetrical key produced automatically at (more or less) random by the Web

browser. The Web browser will pass the encrypted secret key to the vendor, and this secret key will be used to encrypt the

session. Several secret keys are exchanged, but a detailed discussion of the complexities is a bit beyond the scope document.

[11] FOOTNOTE: BIBLE CODES & THEIR KIN

* One of the more dubious aspects of cryptology is the hunt for hidden ciphers in great works of literature, particularly the BFor want of a better term it might be called "pseudocryptology". The usual approach of pseudocryptologists is to scan tho text of interest and take out letters at intervals, for example, every 13th letter, every 100th letter, or every 1,617th letter. They arrange the letters in a block and see if patterns pop out. The patterns are typically identified as scrambled words buried in a stream of gibberish. It is not too surprising that one can extract seemingly sensible remarks using this scheme. It's a logical level with, say, playing the music of the Beatles backward and listening for secret messages left by John Lennon. A modern advocate of such hidden codes, Michael Drosnin, has published a series of BIBLE CODE books that made

predictions of a wide number of events hidden in the Bible. The books made the best-seller lists, even though most predictions identified by Drosnin in his books were, conveniently, of events that had already happened. His predictions of future events, such as the end of the world in the year 2000, were somewhat inaccurate.

Drosnin incautiously challenged critics to find similar predictions about, say, the assassination of a prime minister in MOBYand so an Australian mathematician named Brendan McKay went out and did precisely that, identifying "predictions" of assassinations. Similar exercises were performed on WAR AND PEACE and other classics. However, pseudocryptology has been around for a long time, and though it is easily debunked, it still refuses to die off.