Cryptography: Introduction

Markulf Kohlweiss & Myrto Arapinis
School of Informatics
University of Edinburgh

January 10, 2021

What is cryptography?

> *"The practice of creating and understanding codes that keep information secret."*
>
> Cambridge dictionary

What is cryptography?

> *"The practice of creating and understanding codes that keep information secret."*
>
> Cambridge dictionary

But nowadays cryptography encompasses many more things than just secret communications.

> *"Cryptography is the scientific study of techniques for securing [against internal or external attacks] digital information, transactions, and distributed computations."*
>
> Jonathan Katz and Yehuda Lindell
> in Introduction to Modern Cryptography

Cryptography is everywhere!

Cryptographic methods are powerful tools at the core of many security mechanisms used:

- to securely and confidentially access a website such as an online banking website;
- to attest the identity of the organization operating a web server;
- when talking over a mobile phone;
- to enforce access control in a multi-user operating system;
- to prevent thieves extracting trade secrets from stolen laptops;
- to prevent software copying;
- *etc*

Cryptography (and security more broadly) is becoming a more and more central topic within computer science

Important remark

Cryptography is not:

- The solution to all security problems (see other sections of the course)
- Secure if not implemented and/or deployed correctly
- Something you will be able to invent at the end of this course

Learning objectives for the Cryptography section

- Appreciate the variety of applications that use cryptography with different purposes
- Introduce the basic concepts of cryptography
- Understand the type of problems cryptography can address
- Understand the types of problems that need to be addressed when using cryptography

Topics in the Cryptography section

We will discuss constructions for:

- Symmetric Encryption
- Asymmetric (public-key) Encryption
- Hash functions and Message Authentication Codes (MACs)
- Digital Signatures
- Public Key Infrastructure (PKI)

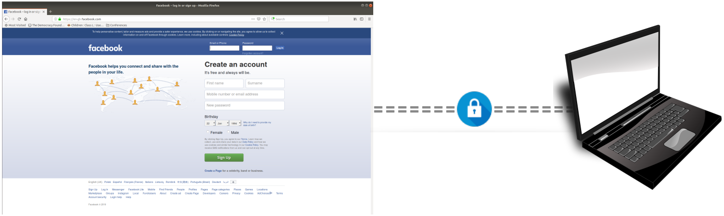
We present only the rudiments of the topic:

- What cryptography can achieve
- That cryptography can go wrong
- What is good practice when using cryptography

Cryptography: Symmetric encryption

Goal: confidentiality

- Secure communications



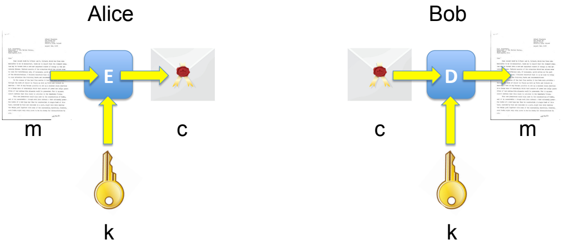
- File protection

Symmetric encryption schemes

A symmetric cipher consists of two algorithms

- encryption algorithm $E: \mathcal{K} \times \mathcal{M} \to \mathcal{C}$
- decryption algorithm $D: \mathcal{K} \times \mathcal{C} \to \mathcal{M}$

st. $\forall k \in \mathcal{K}$, and $\forall m \in \mathcal{M}$, $D(k, E(k, m)) = m$



- same key k to encrypt and decrypt
- the key k is secret: only known to Alice and Bob

What is a good encryption scheme?

An encryption scheme is secure against a given adversary, if this adversary cannot

- recover the secret key k
- recover the plaintext m underlying a ciphertext c
- recover any bits of the plaintext m underlying a ciphertext c
- ...

Kerckhoff's principle

The architecture and design of a security system/mechanism should be made public

No security through obscurity!

- The encryption (E) and decryption (D) algorithms are public
- The security relies entirely on the secrecy of the key

Open design allows for a system to be scrutinised by many users, white hat hackers, academics, *etc.*
$\rightsquigarrow$ early discovery and corrections of flaws/vulnerabilities

Adversary's capabilities

- A cryptographic scheme is secure under some assumptions, that is against a certain type of attacker
- A cryptographic scheme may be vulnerable to certain types of attacks but not others

Adversary's capabilities

- A cryptographic scheme is secure under some assumptions, that is against a certain type of attacker
- A cryptographic scheme may be vulnerable to certain types of attacks but not others

The attacker know the encryption/decryption algorithms but may have access to :

- Ciphertext only attack - some ciphertexts $c_1, \ldots, c_n$
- Known plaintext attack some plaintext/ciphertext pairs (m_1, c_1), ..., (m_n, c_n) st. $c_i = E(k, m_i))$
- Chosen plaintext attack - he has access to an encryption oracle - can maybe trick a user to encrypt messages $m_1, \ldots, m_n$ of his choice
- Chosen ciphertext attack - he has access to a decryption oracle - can maybe trick a user to decrypt ciphertexts $c_1, \ldots, c_n$ of his choice
- unlimited, or polynomial, or realistic ($\leq 2^{80}$) computational power

Brute-force attack - attack on all schemes

- Try all possible keys $k \in \mathcal{K}$ - requires some knowledge about the structure of plaintext

- Making exhaustive search unfeasible:
  - $\mathcal{K}$ should be sufficiently large, *i.e.* keys should be sufficiently long
  - Keys should be sampled uniformly at random from $\mathcal{K}$

A simple scheme: the substitution cipher

- shared secret: a permutation π of the set of characters
  $$\pi = \begin{array}{l} a \mapsto q \ b \mapsto w \ c \mapsto e \ d \mapsto r \ e \mapsto t \ f \mapsto y \ g \mapsto u \ h \mapsto i \ i \mapsto o \\ j \mapsto m \ k \mapsto a \ l \mapsto s \ m \mapsto d \ n \mapsto f \ o \mapsto g \ p \mapsto h \ q \mapsto j \ r \mapsto k \\ s \mapsto l \ t \mapsto z \ u \mapsto x \ v \mapsto c \ w \mapsto v \ x \mapsto b \ y \mapsto n \ z \mapsto p \end{array}$$

- Encryption: apply π to each character of the plaintext
  $$E(\pi, p_1 \dots p_n) = \pi(p_1) \dots \pi(p_n)$$

- Decryption: apply π^{-1} to each character of the plaintext
  $$D(\pi, c_1 \dots c_n) = \pi^{-1}(c_1) \dots \pi^{-1}(c_n)$$

Substitution cipher: example

$\pi =$ $a \mapsto q$ $b \mapsto w$ $c \mapsto e$ $d \mapsto r$ $e \mapsto t$ $f \mapsto y$ $g \mapsto u$ $h \mapsto i$ $i \mapsto o$ $j \mapsto m$ $k \mapsto a$ $l \mapsto s$ $m \mapsto d$ $n \mapsto f$ $o \mapsto g$ $p \mapsto h$ $q \mapsto j$ $r \mapsto k$ $s \mapsto l$ $t \mapsto z$ $u \mapsto x$ $v \mapsto c$ $w \mapsto v$ $x \mapsto b$ $y \mapsto n$ $z \mapsto p$

m = THIS COURSE AIMS TO INTRODUCE YOU TO THE PRINCIPLES AND TECHNIQUES OF SECURING COMPUTERS AND COMPUTER NETWORKS WITH FOCUS ON INTERNET SECURITY. THE COURSE IS EFFECTIVELY SPLIT INTO TWO PARTS. FIRST INTRODUCING THE THEORY OF CRYPTOGRAPHY INCLUDING HOW MANY CLASSICAL AND POPULAR ALGORITHMS WORK E.G. DES, RSA, DIGITAL SIGNATURES, AND SECOND PROVIDING DETAILS OF REAL INTERNET SECURITY PROTOCOLS, ALGORITHMS, AND THREATS, E.G. IPSEC, VIRUSES, FIREWALLS. HENCE, YOU WILL LEARN BOTH THEORETICAL ASPECTS OF COMPUTER AND NETWORK SECURITY AS WELL AS HOW THAT THEORY IS APPLIED IN THE INTERNET. THIS KNOWLEDGE WILL HELP YOU IN DESIGNING AND DEVELOPING SECURE APPLICATIONS AND NETWORK PROTOCOLS AS WELL AS BUILDING SECURE NETWORKS.

c = ZIOL EGXKLT QODL ZG OFZKGRXET NGX ZG ZIT HKOFEOHSTL QFR ZTEIFOJXTL GY LTEXKOFU EGDHXZTKL QFR EGDHXZTK FTZVGKAL VOZI YGEXL GF OFZTKFTZ LTEXKOZN. ZIT EGXKLT OL TYYTEZOCTSN LHSOZ OFZG ZVG HQKZL. YOKLZ OFZKGRXEOFU ZIT ZITGKN GY EKNHZGUKQHIN OFESXROFU IGV DQFN ESQLLOEQS QFR HGHXSQK QSUGKOZIDL VGKA T.U. RTL, KLQ, ROUOZQS LOUFQZXKTL, QFR LTEGFR HKGCOROFU RTZQOSL GY KTQS OFZTKFTZ LTEXKOZN HKGZGEGSL, QSUGKOZIDL, QFR ZIKTQZL, T.U. OHLTE, COKXLTL, YOKTVQSSL. ITFET, NGX VOSS STQKF WGZI ZITGKTZOEQS QLHTEZL GY EGDHXZTK QFR FTZVGKA LTEXKOZN QL VTSS QL IGV ZIQZ ZITGKN OL QHHSOTR OF ZIT OFZTKFTZ. ZIOL AFGVSTRUT VOSS ITSH NGX OF RTLOUFOFU QFR RTCTSGHOFU LTEXKT QHHSOEQZOGFL QFR FTZVGKA HKGZGEGSL QL VTSS QL WXOSROFU LTEXKT FTZVGKAL.

Quiz

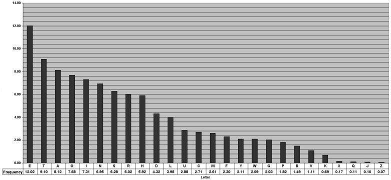
Breaking the substitution cipher

Breaking the substitution cipher

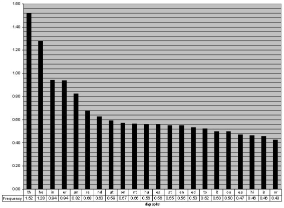
- Key space size: $|\mathcal{K}| = 26!$ $(\approx 2^{88})$ $\Rightarrow$ brute force infeasible!

Breaking the substitution cipher

- Key space size: $|\mathcal{K}| = 26!$ ($\approx 2^{88}$) $\Rightarrow$ brute force infeasible!
- Frequency analysis: exploit regularities of the language
  - Use frequency of letters in English text



  - Use frequency of digrams in English text



  - Use frequency of trigrams in English text

the > and > ing

  - Use expected words

Breaking the substitution cipher: example

$\pi =$

$c =$ ZIOL EGXKLT QODL ZG OFZKGRXET NGX ZG ZIT HKOFEOHSTL QFR ZTEIFOJXTL GY LTEXKOFU EGDHXZTKL QFR EGDHXZTK FTZVGKAL VOZI YGEXL GF OFZTKFTZ LTEXKOZN. ZIT EGXKLT OL TYYTEZOCTSN LHSOZ OFZG ZVG HQKZL. YOKLZ OFZKGRXEOFU ZIT ZITGKN GY EKNHZGUKQHIN OFESXROFU IGV DQFN ESQLLOEQS QFR HGHXSQK QSUGKOZIDL VGKA T.U. RTL, KLQ, ROUOZQS LOUFQZXKTL, QFR LTEGFR HKGCOROFU RTZQOSL GY KTQS OFZTKFTZ LTEXKOZN HKGZGEGSL, QSUGKOZIDL, QFR ZIKTQZL, T.U. OHLTE, COKXLTL, YOKTVQSSL. ITFET, NGX VOSS STQKF WGZI ZITGKTZOEQS QLHTEZL GY EGDHXZTK QFR FTZVGKA LTEXKOZN QL VTSS QL IGV ZIQZ ZITGKN OL QHHSOTR OF ZIT OFZTKFTZ. ZIOL AFGVSTRUT VOSS ITSH NGX OF RTLOUFOFU QFR RTCTSGHOFU LTEXKT QHHSOEQZOGFL QFR FTZVGKA HKGZGEGSL QL VTSS QL WXOSROFU LTEXKT FTZVGKAL.

Breaking the substitution cipher: example

$\pi =$

c = ZIOL EGXKLT QODL ZG OFZKGRXET NGX ZG ZIT HKOFEOHSTL QFR ZTEIFOJXTL GY
LTEXKOFU EGDHXZTKL QFR EGDHXZTK FTZVGKAL VOZI YGEXL GF OFZTKFTZ
LTEXKOZN. ZIT EGXKLT OL TYYTEZOCTSN LHSOZ OFZG ZVG HQKZL. YOKLZ
OFZKGRXEOFU ZIT ZITGKN GY EKNHZGUKQHIN OFESXROFU IGV DQFN ESQLLOEQS
QFR HGHXSQK QSUGKOZIDL VGKA T.U. RTL, KLQ, ROUOZQS LOUFQZXKTL, QFR
LTEGFR HKGCOROFU RTZQOSL GY KTQS OFZTKFTZ LTEXKOZN HKGZGEGSL,
QSUGKOZIDL, QFR ZIKTQZL, T.U. OHLTE, COKXLTL, YOKTVQSSL. ITFET, NGX
VOSS STQKF WGZI ZITGKTZOEQS QLHTEZL GY EGDHXZTK QFR FTZVGKA
LTEXKOZN QL VTSS QL IGV ZIQZ ZITGKN OL QHHSOTR OF ZIT OFZTKFTZ. ZIOL
AFGVSTRUT VOSS ITSH NGX OF RTLOUFOFU QFR RTCTSGHOFU LTEXKT
QHHSOEQZOGFL QFR FTZVGKA HKGZGEGSL QL VTSS QL WXOSROFU LTEXKT
FTZVGKAL.

Most common letters in c: t $>$ z $>$ o $>$ l

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t$

c = TIOL EGXKLE QODL TG OFTKGRXEE NGX TG TIE HKOFEOHSEL QFR TEEIFOJXEL GY LEEXKOFU EGDHXTEKL QFR EGDHXTEK FETVGKAL VOTI YGEXL GF OFTEKFET LEEXKOTN. TIE EGXKLE OL EYYEETOCESN LHSOT OFTG TVG HQKTL. YOKLT OFTKGRXEOFU TIE TIEGKN GY EKNHTGUKQHIN OFESXROFU IGV DQFN ESQLLOEQS QFR HGHXSQK QSUGKOTIDL VGKA E.U. REL, KLQ, ROUOTQS LOUFQTXKEL, QFR LEEGFR HKGCOROFU RETQOSL GY KEQS OFTEKFET LEEXKOTN HKGTGEGSL, QSUGKOTIDL, QFR TIKEQTL, E.U. OHLEE, COKXLEL, YOKEVQSSL. IEFEE, NGX VOSS SEQKF WGTI TIEGKETOEQS QLHEETL GY EGDHXTEK QFR FETVGKA LEEXKOTN QL VESS QL IGV TIQT TIEGKN OL QHHSOER OF TIE OFTEKFET. TIOL AFGVSERUE VOSS IESH NGX OF RELOUFOFU QFR RECESGHOFU LEEXKE QHHSOEQTOGFL QFR FETVGKA HKGTGEGSL QL VESS QL WXOSROFU LEEXKE FETVGKAL.

Most common letters in c: t $>$ z $>$ …

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t$

c = TIOL EGXKLE QODL TG OFTKGRXEE NGX TG TIE HKOFEOHSEL QFR TEEIFOJXEL GY
LEEXKOFU EGDHXTEKL QFR EGDHXTEK FETVGKAL VOTI YGEXL GF OFTEKFET
LEEXKOTN. TIE EGXKLE OL EYYEETOCESN LHSOT OFTG TVG HQKTL. YOKLT
OFTKGRXEOFU TIE TIEGKN GY EKNHTGUKQHIN OFESXROFU IGV DQFN ESQLLOEQS
QFR HGHXSQK QSUGKOTIDL VGKA E.U. REL, KLQ, ROUOTQS LOUFQTXKEL, QFR
LEEGFR HKGCOROFU RETQOSL GY KEQS OFTEKFET LEEXKOTN HKGTGEGSL,
QSUGKOTIDL, QFR TIKEQTL, E.U. OHLEE, COKXLEL, YOKEVQSSL. IEFEE, NGX
VOSS SEQKF WGTI TIEGKETOEQS QLHEETL GY EGDHXTEK QFR FETVGKA
LEEXKOTN QL VESS QL IGV TIQT TIEGKN OL QHHSOER OF TIE OFTEKFET. TIOL
AFGVSERUE VOSS IESH NGX OF RELOUFOFU QFR RECESGHOFU LEEXKE
QHHSOEQTOGFL QFR FETVGKA HKGTGEGSL QL VESS QL WXOSROFU LEEXKE
FETVGKAL.

Most common digrams in c: of > zi > ...

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t$

c = TIOL EGXKLE QODL TG OFTKGRXEE NGX TG TIE HKOFEOHSEL QFR TEEIFOJXEL GY
LEEXKOFU EGDHXTEKL QFR EGDHXTEK FETVGKAL VOTI YGEXL GF OFTEKFET
LEEXKOTN. TIE EGXKLE OL EYYEETOCESN LHSOT OFTG TVG HQKTL. YOKLT
OFTKGRXEOFU TIE TIEGKN GY EKNHTGUKQHIN OFESXROFU IGV DQFN ESQLLOEQS
QFR HGHXSQK QSUGKOTIDL VGKA E.U. REL, KLQ, ROUOTQS LOUFQTXKEL, QFR
LEEGFR HKGCOROFU RETQOSL GY KEQS OFTEKFET LEEXKOTN HKGTGEGSL,
QSUGKOTIDL, QFR TIKEQTL, E.U. OHLEE, COKXLEL, YOKEVQSSL. IEFEE, NGX
VOSS SEQKF WGTI TIEGKETOEQS QLHEETL GY EGDHXTEK QFR FETVGKA
LEEXKOTN QL VESS QL IGV TIQT TIEGKN OL QHHSOER OF TIE OFTEKFET. TIOL
AFGVSERUE VOSS IESH NGX OF RELOUFOFU QFR RECESGHOFU LEEXKE
QHHSOEQTOGFL QFR FETVGKA HKGTGEGSL QL VESS QL WXOSROFU LEEXKE
FETVGKAL.

Most common digrams in c: of > zi > ...

t↦z suggests h↦i

Breaking the substitution cipher: example

$\pi =$ $t \mapsto z$ $e \mapsto t$ $h \mapsto i$

c = THOL EGXKLE QODL TG OFTKGRXEE NGX TG THE HKOFEOHSEL QFR TEEHFOJXEL GY LEEXKOFU EGDHXTEKL QFR EGDHXTEK FETVGKAL VOTH YGEXL GF OFTEKFET LEEXKOTN. THE EGXKLE OL EYYEETOCESN LHSOT OFTG TVG HQKTL. YOKLT OFTKGRXEOFU THE THEGKN GY EKNHTGUKQHHN OFESXROFU HGV DQFN ESQLLOEQS QFR HGHXSQK QSUGKOTHDL VGKA E.U. REL, KLQ, ROUOTQS LOUFQTXKEL, QFR LEEGFR HKGCOROFU RETQOSL GY KEQS OFTEKFET LEEXKOTN HKGTGEGSL, QSUGKOTHDL, QFR THKEQTL, E.U. OHLEE, COKXLEL, YOKEVQSSL. HEFEE, NGX VOSS SEQKF WGTH THEGKETOEQS QLHEETL GY EGDHXTEK QFR FETVGKA LEEXKOTN QL VESS QL HGV THQT THEGKN OL QHHSOER OF THE OFTEKFET. THOL AFGVSERUE VOSS HESH NGX OF RELOUFOFU QFR RECESGHOFU LEEXKE QHHSOEQTOGFL QFR FETVGKA HKGTGEGSL QL VESS QL WXOSROFU LEEXKE FETVGKAL.

Most common digrams in c: of $>$ zi $>$ ...

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i$

c = THOL EGXKLE QODL TG OFTKGRXEE NGX TG THE HKOFEOHSEL QFR TEEHFOJXEL GY LEEXKOFU EGDHXTEKL QFR EGDHXTEK FETVGKAL VOTH YGEXL GF OFTEKFET LEEXKOTN. THE EGXKLE OL EYYEETOCESN LHSOT OFTG TVG HQKTL. YOKLT OFTKGRXEOFU THE THEGKN GY EKNHTGUKQHHN OFESXROFU HGV DQFN ESQLLOEQS QFR HGHXSQK QSUGKOTHDL VGKA E.U. REL, KLQ, ROUOTQS LOUFQTXKEL, QFR LEEGFR HKGCOROFU RETQOSL GY KEQS OFTEKFET LEEXKOTN HKGTGEGSL, QSUGKOTHDL, QFR THKEQTL, E.U. OHLEE, COKXLEL, YOKEVQSSL. HEFEE, NGX VOSS SEQKF WGTH THEGKETOEQS QLHEETL GY EGDHXTEK QFR FETVGKA LEEXKOTN QL VESS QL HGV THQT THEGKN OL QHHSOER OF THE OFTEKFET. THOL AFGVSERUE VOSS HESH NGX OF RELOUFOFU QFR RECESGHOFU LEEXKE QHHSOEQTOGFL QFR FETVGKA HKGTGEGSL QL VESS QL WXOSROFU LEEXKE FETVGKAL.

Most common digrams in c: of > zi > ...
we guess in$\mapsto$of

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f$

c = THIL EGXKLE QIDL TG INTKGRXEE NGX TG THE HKINEIHSEL QNR TEEHNIJXEL GY
LEEXKINU EGDHXTEKL QNR EGDHXTEK NETVGKAL VITH YGEXL GN INTEKNET
LEEXKITN. THE EGXKLE IL EYYEETICESN LHSIT INTG TVG HQKTL. YIKLT
INTKGRXEINU THE THEGKN GY EKNHTGUKQHHN INESXRINU HGV DQNN ESQLLIEQS
QNR HGHXSQK QSUGKITHDL VGKA E.U. REL, KLQ, RIUITQS LIUNQTXKEL, QNR
LEEGNR HKGCIRINU RETQISL GY KEQS INTEKNET LEEXKITN HKGTGEGSL,
QSUGKITHDL, QNR THKEQTL, E.U. IHLEE, CIKXLEL, YIKEVQSSL. HENEE, NGX
VISS SEQKN WGTH THEGKETIEQS QLHEETL GY EGDHXTEK QNR NETVGKA
LEEXKITN QL VESS QL HGV THQT THEGKN IL QHHSIER IN THE INTEKNET. THIL
ANGVSERUE VISS HESH NGX IN RELIUNINU QNR RECESGHINU LEEXKE
QHHSIEQTIGNL QNR NETVGKA HKGTGEGSL QL VESS QL WXISRINU LEEXKE
NETVGKAL.

Most common digrams in c: of $>$ zi $> \ldots$

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f$

c = THIL EGXKLE QIDL TG INTKGRXEE NGX TG THE HKINEIHSEL QNR TEEHNIJXEL GY
LEEXKINU EGDHXTEKL QNR EGDHXTEK NETVGKAL VITH YGEXL GN **INTEKNET**
LEEXKITN. THE EGXKLE IL EYYEETICESN LHSIT INTG TVG HQKTL. YIKLT
INTKGRXEINU THE THEGKN GY EKNHTGUKQHHN INESXRINU HGV DQNN ESQLLIEQS
QNR HGHXSQK QSUGKITHDL VGKA E.U. REL, KLQ, RIUITQS LIUNQTXKEL, QNR
LEEGNR HKGCIRINU RETQISL GY KEQS INTEKNET LEEXKITN HKGTGEGSL,
QSUGKITHDL, QNR THKEQTL, E.U. IHLEE, CIKXLEL, YIKEVQSSL. HENEE, NGX
VISS SEQKN WGTH THEGKETIEQS QLHEETL GY EGDHXTEK QNR NETVGKA
LEEXKITN QL VESS QL HGV THQT THEGKN IL QHHSIER IN THE INTEKNET. THIL
ANGVSERUE VISS HESH NGX IN RELIUNINU QNR RECESGHINU LEEXKE
QHHSIEQTIGNL QNR NETVGKA HKGTGEGSL QL VESS QL WXISRINU LEEXKE
NETVGKAL.

We identify in c the word **INTEKNET**

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f$

c = THIL EGXKLE QIDL TG INTKGRXEE NGX TG THE HKINEIHSEL QNR TEEHNIJXEL GY
LEEXKINU EGDHXTEKL QNR EGDHXTEK NETVGKAL VITH YGEXL GN **INTEKNET**
LEEXKITN. THE EGXKLE IL EYYEETICESN LHSIT INTG TVG HQKTL. YIKLT
INTKGRXEINU THE THEGKN GY EKNHTGUKQHHN INESXRINU HGV DQNN ESQLLIEQS
QNR HGHXSQK QSUGKITHDL VGKA E.U. REL, KLQ, RIUITQS LIUNQTXKEL, QNR
LEEGNR HKGCIRINU RETQISL GY KEQS INTEKNET LEEXKITN HKGTGEGSL,
QSUGKITHDL, QNR THKEQTL, E.U. IHLEE, CIKXLEL, YIKEVQSSL. HENEE, NGX
VISS SEQKN WGTH THEGKETIEQS QLHEETL GY EGDHXTEK QNR NETVGKA
LEEXKITN QL VESS QL HGV THQT THEGKN IL QHHSIER IN THE INTEKNET. THIL
ANGVSERUE VISS HESH NGX IN RELIUNINU QNR RECESGHINU LEEXKE
QHHSIEQTIGNL QNR NETVGKA HKGTGEGSL QL VESS QL WXISRINU LEEXKE
NETVGKAL.

We identify in c the word **INTEKNET**
suggests r$\mapsto$k

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f \; r \mapsto k$

c = THIL EGXRLE QIDL TG INTRGRXEE NGX TG THE HRINEIHSEL QNR TEEHNIJXEL GY
LEEXRINU EGDHXTERL QNR EGDHXTER NETVGRAL VITH YGEXL GN **INTERNET**
LEEXRITN. THE EGXRLE IL EYYEETICESN LHSIT INTG TVG HQRTL. YIRLT
INTRGRXEINU THE THEGRN GY ERNHTGURQHHN INESXRINU HGV DQNN ESQLLIEQS
QNR HGHXSQR QSUGRITHDL VGRA E.U. REL, RLQ, RIUITQS LIUNQTXREL, QNR
LEEGNR HRGCIRINU RETQISL GY REQS INTERNET LEEXRITN HRGTGEGSL,
QSUGRITHDL, QNR THREQTL, E.U. IHLEE, CIRXLEL, YIREVQSSL. HENEE, NGX
VISS SEQRN WGTH THEGRETIEQS QLHEETL GY EGDHXTER QNR NETVGRA
LEEXRITN QL VESS QL HGV THQT THEGRN IL QHHSIER IN THE INTERNET. THIL
ANGVSERUE VISS HESH NGX IN RELIUNINU QNR RECESGHINU LEEXYE
QHHSIEQTIGNL QNR NETVGRA HRGTGEGSL QL VESS QL WXISRINU LEEXRE
NETVGRAL.

We identify in c the word **INTEKNET**

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f \; r \mapsto k$

c = THIL EGXRLE QIDL TG INTRGRXEE NGX TG THE HRINEIHSEL QNR TEEHNIJXEL GY
LEEXRINU EGDHXTERL QNR EGDHXTER NETVGRAL VITH YGEXL GN INTERNET
LEEXRITN. THE EGXRLE IL EYYEETICESN LHSIT INTG TVG HQRTL. YIRLT
INTRGRXEINU THE THEGRN GY ERNHTGURQHHN INESXRINU HGV DQNN ESQLLIEQS
QNR HGHXSQR QSUGRITHDL VGRA E.U. REL, RLQ, RIUITQS LIUNQTXREL, QNR
LEEGNR HRGCIRINU RETQISL GY REQS INTERNET LEEXRITN HRGTGEGSL,
QSUGRITHDL, QNR THREQTL, E.U. IHLEE, CIRXLEL, YIREVQSSL. HENEE, NGX
VISS SEQRN WGTH THEGRETIEQS QLHEETL GY EGDHXTER QNR NETVGRA
LEEXRITN QL VESS QL HGV THQT THEGRN IL QHHSIER IN THE INTERNET. THIL
ANGVSERUE VISS HESH NGX IN RELIUNINU QNR RECESGHINU LEEXYE
QHHSIEQTIGNL QNR NETVGRA HRGTGEGSL QL VESS QL WXISRINU LEEXRE
NETVGRAL.

The first word is THIL

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f \; r \mapsto k$

c = THIL EGXRLE QIDL TG INTRGRXEE NGX TG THE HRINEIHSEL QNR TEEHNIJXEL GY
LEEXRINU EGDHXTERL QNR EGDHXTER NETVGRAL VITH YGEXL GN INTERNET
LEEXRITN. THE EGXRLE IL EYYEETICESN LHSIT INTG TVG HQRTL. YIRLT
INTRGRXEINU THE THEGRN GY ERNHTGURQHHN INESXRINU HGV DQNN ESQLLIEQS
QNR HGHXSQR QSUGRITHDL VGRA E.U. REL, RLQ, RIUITQS LIUNQTXREL, QNR
LEEGNR HRGCIRINU RETQISL GY REQS INTERNET LEEXRITN HRGTGEGSL,
QSUGRITHDL, QNR THREQTL, E.U. IHLEE, CIRXLEL, YIREVQSSL. HENEE, NGX
VISS SEQRN WGTH THEGRETIEQS QLHEETL GY EGDHXTER QNR NETVGRA
LEEXRITN QL VESS QL HGV THQT THEGRN IL QHHSIER IN THE INTERNET. THIL
ANGVSERUE VISS HESH NGX IN RELIUNINU QNR RECESGHINU LEEXYE
QHHSIEQTIGNL QNR NETVGRA HRGTGEGSL QL VESS QL WXISRINU LEEXRE
NETVGRAL.

The first word is THIL
suggests s$\mapsto$l

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f \; r \mapsto k \; s \mapsto l$

c = THIS EGXRSE QIDS TG INTRGRXEE NGX TG THE HRINEIHSES QNR TEEHNIJXES GY SEEXRINU EGDHXTERS QNR EGDHXTER NETVGRAS VITH YGEXS GN INTERNET SEEXRITN. THE EGXRSE IS EYYEETICESN SHSIT INTG TVG HQRTS. YIRST INTRGRXEINU THE THEGRN GY ERNHTGURQHHN INESXRINU HGV DQNN ESQSSIEQS QNR HGHXSQR QSUGRITHDS VGRA E.U. RES, RSQ, RIUITQS SIUNQTXRES, QNR SEEGNR HRGCIRINU RETQISS GY REQS INTERNET SEEXRITN HRGTGEGSS, QSUGRITHDS, QNR THREQTS, E.U. IHSEE, CIRXSES, YIREVQSSS. HENEE, NGX VISS SEQRN WGTH THEGRETIEQS QSHEETS GY EGDHXTER QNR NETVGRA SEEXRITN QS VESS QS HGV THQT THEGRN IS QHHSIER IN THE INTERNET. THIS ANGVSERUE VISS HESH NGX IN RESIUNINU QNR RECESGHINU SEEXRE QHHSIEQTIGNS QNR NETVGRA HRGTGEGSS QS VESS QS WXISRINU SEEXRE NETVGRAS.

The first word is THIL

Breaking the substitution cipher: example

$\pi = \quad t \mapsto z \; e \mapsto t \; h \mapsto i \; i \mapsto o \; n \mapsto f \; r \mapsto k \; s \mapsto l$

c = THIS EGXRSE QIDS TG INTRGRXEE NGX TG THE HRINEIHSES QNR TEEHNIJXES GY SEEXRINU EGDHXTERS QNR EGDHXTER NETVGRAS VITH YGEXS GN INTERNET SEEXRITN. THE EGXRSE IS EYYEETICESN SHSIT INTG TVG HQRTS. YIRST INTRGRXEINU THE THEGRN GY ERNHTGURQHHN INESXRINU HGV DQNN ESQSSIEQS QNR HGHXSQR QSUGRITHDS VGRA E.U. RES, RSQ, RIUITQS SIUNQTXRES, QNR SEEGNR HRGCIRINU RETQISS GY REQS INTERNET SEEXRITN HRGTGEGSS, QSUGRITHDS, QNR THREQTS, E.U. IHSEE, CIRXSES, YIREVQSSS. HENEE, NGX VISS SEQRN WGTH THEGRETIEQS QSHEETS GY EGDHXTER QNR NETVGRA SEEXRITN QS VESS QS HGV THQT THEGRN IS QHHSIER IN THE INTERNET. THIS ANGVSERUE VISS HESH NGX IN RESIUNINU QNR RECESGHINU SEEXRE QHHSIEQTIGNS QNR NETVGRA HRGTGEGSS QS VESS QS WXISRINU SEEXRE NETVGRAS.

Going back to letter frequency and a few more guesses!!

Breaking the substitution cipher: example

$\pi =$ $a \mapsto q\ b \mapsto w\ c \mapsto e\ d \mapsto r\ e \mapsto t\ f \mapsto y\ g \mapsto u\ h \mapsto i\ i \mapsto o\ j \mapsto m\ k \mapsto a\ l \mapsto s$
$m \mapsto d\ n \mapsto f\ o \mapsto g\ p \mapsto h\ q \mapsto j\ r \mapsto k\ s \mapsto l\ t \mapsto z\ u \mapsto x\ v \mapsto c\ w \mapsto v\ x \mapsto b$
$y \mapsto n\ z \mapsto p$

m = THIS COURSE AIMS TO INTRODUCE YOU TO THE PRINCIPLES AND TECHNIQUES OF SECURING COMPUTERS AND COMPUTER NETWORKS WITH FOCUS ON INTERNET SECURITY. THE COURSE IS EFFECTIVELY SPLIT INTO TWO PARTS. FIRST INTRODUCING THE THEORY OF CRYPTOGRAPHY INCLUDING HOW MANY CLASSICAL AND POPULAR ALGORITHMS WORK E.G. DES, RSA, DIGITAL SIGNATURES, AND SECOND PROVIDING DETAILS OF REAL INTERNET SECURITY PROTOCOLS, ALGORITHMS, AND THREATS, E.G. IPSEC, VIRUSES, FIREWALLS. HENCE, YOU WILL LEARN BOTH THEORETICAL ASPECTS OF COMPUTER AND NETWORK SECURITY AS WELL AS HOW THAT THEORY IS APPLIED IN THE INTERNET. THIS KNOWLEDGE WILL HELP YOU IN DESIGNING AND DEVELOPING SECURE APPLICATIONS AND NETWORK PROTOCOLS AS WELL AS BUILDING SECURE NETWORKS.

Going back to letter frequency and a few more guesses!!

A better substitution cipher: The One-Time Pad (OTP)

A better substitution cipher: The One-Time Pad (OTP)

- $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}^n$

A better substitution cipher: The One-Time Pad (OTP)

- $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}^n$
- Encryption: $\forall k \in \mathcal{K}.\ \forall m \in \mathcal{M}.\ E(k,m) = k \oplus m$

A better substitution cipher: The One-Time Pad (OTP)

- $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}^n$
- Encryption: $\forall k \in \mathcal{K}.\ \forall m \in \mathcal{M}.\ E(k,m) = k \oplus m$

$$
\begin{array}{cccccccccc}
k & = & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
m & = & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\
\hline
c & = & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0
\end{array}
$$

A better substitution cipher: The One-Time Pad (OTP)

- $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}^n$
- Encryption: $\forall k \in \mathcal{K}.\ \forall m \in \mathcal{M}.\ E(k,m) = k \oplus m$

$$
\begin{array}{ccccccccccc}
k & = & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\
m & = & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\
\hline
c & = & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0
\end{array}
$$

- Decryption: $\forall k \in \mathcal{K}.\ \forall c \in \mathcal{C}.\ D(k,c) = k \oplus c$

A better substitution cipher: The One-Time Pad (OTP)

- $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}^n$
- Encryption: $\forall k \in \mathcal{K}.\ \forall m \in \mathcal{M}.\ E(k,m) = k \oplus m$

$$\begin{array}{cccccccccc} k & = & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ m & = & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ \hline c & = & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{array}$$

- Decryption: $\forall k \in \mathcal{K}.\ \forall c \in \mathcal{C}.\ D(k,c) = k \oplus c$

$$\begin{array}{cccccccccc} k & = & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ c & = & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline m & = & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}$$

A better substitution cipher: The One-Time Pad (OTP)

- $\mathcal{M} = \mathcal{C} = \mathcal{K} = \{0,1\}^n$
- Encryption: $\forall k \in \mathcal{K}.\ \forall m \in \mathcal{M}.\ E(k,m) = k \oplus m$

$$\begin{array}{ccccccccccc} k & = & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ m & = & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ \hline c & = & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \end{array}$$

- Decryption: $\forall k \in \mathcal{K}.\ \forall c \in \mathcal{C}.\ D(k,c) = k \oplus c$

$$\begin{array}{ccccccccccc} k & = & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ c & = & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ \hline m & = & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}$$

- Consistency: $D(k, E(k,m)) = k \oplus (k \oplus m) = m$

Perfect secrecy

Definition

A cipher (E, D) over $(\mathcal{M}, \mathcal{C}, \mathcal{K})$ satisfies perfect secrecy if for all messages $m_1, m_2 \in \mathcal{M}$ of same length ($|m_1| = |m_2|$), and for all ciphertexts $c \in \mathcal{C}$

$$|Pr(E(k, m_1) = c) - Pr(E(k, m_2) = c)| \leq \epsilon$$

where $k \xleftarrow{r} \mathcal{K}$ and ϵ is some "negligible quantity".

OTP satisfies perfect secrecy

Theorem (Shannon 1949)

The One-Time Pad satisfies perfect secrecy

OTP satisfies perfect secrecy

Theorem (Shannon 1949)

The One-Time Pad satisfies perfect secrecy

Proof: We first note that for all messages $m \in \mathcal{M}$ and all ciphertexts $c \in \mathcal{C}$

$$Pr(E(k, m) = c)$$

where $k \xleftarrow{r} \mathcal{K}$.

OTP satisfies perfect secrecy

Theorem (Shannon 1949)

The One-Time Pad satisfies perfect secrecy

Proof: We first note that for all messages $m \in \mathcal{M}$ and all ciphertexts $c \in \mathcal{C}$

$$Pr(E(k,m) = c) \quad = \quad \frac{\#\{k \in \mathcal{K} : \ k \oplus m = c\}}{\#\mathcal{K}}$$

where $k \xleftarrow{r} \mathcal{K}$.

OTP satisfies perfect secrecy

Theorem (Shannon 1949)

The One-Time Pad satisfies perfect secrecy

<u>Proof:</u> We first note that for all messages $m \in \mathcal{M}$ and all ciphertexts $c \in \mathcal{C}$

$$\begin{aligned} Pr(E(k,m)=c) &= \tfrac{\#\{k\in\mathcal{K}:\ k\oplus m=c\}}{\#\mathcal{K}} \\ &= \tfrac{\#\{k\in\mathcal{K}:\ k=m\oplus c\}}{\#\mathcal{K}} \end{aligned}$$

where $k \xleftarrow{r} \mathcal{K}$.

OTP satisfies perfect secrecy

Theorem (Shannon 1949)

The One-Time Pad satisfies perfect secrecy

<u>Proof:</u> We first note that for all messages $m \in \mathcal{M}$ and all ciphertexts $c \in \mathcal{C}$

$$\begin{aligned} Pr(E(k,m) = c) &= \frac{\#\{k \in \mathcal{K} : \ k \oplus m = c\}}{\#\mathcal{K}} \\ &= \frac{\#\{k \in \mathcal{K} : \ k = m \oplus c\}}{\#\mathcal{K}} \\ &= \frac{1}{\#\mathcal{K}} \end{aligned}$$

where $k \xleftarrow{r} \mathcal{K}$.

OTP satisfies perfect secrecy

Theorem (Shannon 1949)

The One-Time Pad satisfies perfect secrecy

Proof: We first note that for all messages $m \in \mathcal{M}$ and all ciphertexts $c \in \mathcal{C}$

$$
\begin{aligned}
Pr(E(k, m) = c) &= \frac{\#\{k \in \mathcal{K} \colon\ k \oplus m = c\}}{\#\mathcal{K}} \\
&= \frac{\#\{k \in \mathcal{K} \colon\ k = m \oplus c\}}{\#\mathcal{K}} \\
&= \frac{1}{\#\mathcal{K}}
\end{aligned}
$$

where $k \xleftarrow{r} \mathcal{K}$.
Thus, for all messages $m_1, m_2 \in \mathcal{M}$, and for all ciphertexts $c \in \mathcal{C}$

$$|Pr(E(k, m_1) = c) - Pr(E(k, m_2) = c)| \leq$$

OTP satisfies perfect secrecy

Theorem (Shannon 1949)

The One-Time Pad satisfies perfect secrecy

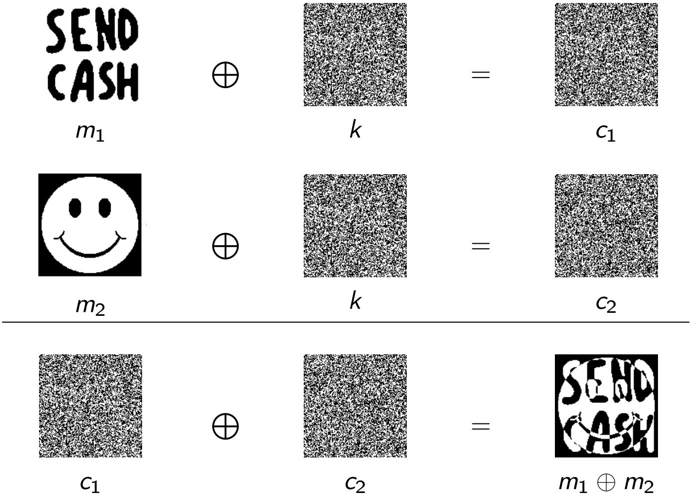
Proof: We first note that for all messages $m \in \mathcal{M}$ and all ciphertexts $c \in \mathcal{C}$

$$\begin{aligned} Pr(E(k,m) = c) &= \frac{\#\{k \in \mathcal{K}:\ k \oplus m = c\}}{\#\mathcal{K}} \\ &= \frac{\#\{k \in \mathcal{K}:\ k = m \oplus c\}}{\#\mathcal{K}} \\ &= \frac{1}{\#\mathcal{K}} \end{aligned}$$

where $k \stackrel{r}{\leftarrow} \mathcal{K}$.
Thus, for all messages $m_1, m_2 \in \mathcal{M}$, and for all ciphertexts $c \in \mathcal{C}$

$$|Pr(E(k,m_1) = c) - Pr(E(k,m_2) = c)| \leq \left| \frac{1}{\#\mathcal{K}} - \frac{1}{\#\mathcal{K}} \right| = 0$$

Two-time pad attacks



Source: Cryptosmith and David Lowry-Duda, `crypto.stackexchange.com/`

Limitations of OTP

Limitations of OTP

- Key-length!
  - The key should be as long as the plaintext

Limitations of OTP

- Key-length!
  - The key should be as long as the plaintext

- Getting true randomness!
  - The key should not be guessable from an attacker
  - If the key is not truly random, frequency analysis might again be possible

Limitations of OTP

- Key-length!
  - The key should be as long as the plaintext

- Getting true randomness!
  - The key should not be guessable from an attacker
  - If the key is not truly random, frequency analysis might again be possible

- Perfect secrecy does not capture all possible attacks

Limitations of OTP

- Key-length!
  - The key should be as long as the plaintext
- Getting true randomness!
  - The key should not be guessable from an attacker
  - If the key is not truly random, frequency analysis might again be possible
- Perfect secrecy does not capture all possible attacks
  - OTP is subject to two-time pad attacks
    given $m_1 \oplus k$ and $m_2 \oplus k$, we can compute
    $m_1 \oplus m_2 = (m_1 \oplus k) \oplus (m_2 \oplus k)$
    English has enough redundancy s.t. $m_1 \oplus m_2 \to m_1, m_2$

Limitations of OTP

- Key-length!
  - The key should be as long as the plaintext

- Getting true randomness!
  - The key should not be guessable from an attacker
  - If the key is not truly random, frequency analysis might again be possible

- Perfect secrecy does not capture all possible attacks
  - OTP is subject to two-time pad attacks
    given $m_1 \oplus k$ and $m_2 \oplus k$, we can compute
    $m_1 \oplus m_2 = (m_1 \oplus k) \oplus (m_2 \oplus k)$
    English has enough redundancy s.t. $m_1 \oplus m_2 \rightarrow m_1, m_2$

  - OTP is malleable
    given the ciphertext $c = E(k, m)$ with $m = "to\ bob : secret\ msg"$, it is possible to compute the ciphertext $c' = E(k, m')$ with
    $m' = "to\ eve : secret\ msg"$
    $c' := c \oplus "to\ bob : 00 \ldots 00" \oplus "to\ eve : 00 \ldots 00"$

Concluding remark

The confidentiality problem is now reduced to a key management problem:

- Where are keys generated?
- How are keys generated?
- Where are keys stored?
- Where are the keys actually used?
- How are key revoked and replaced?