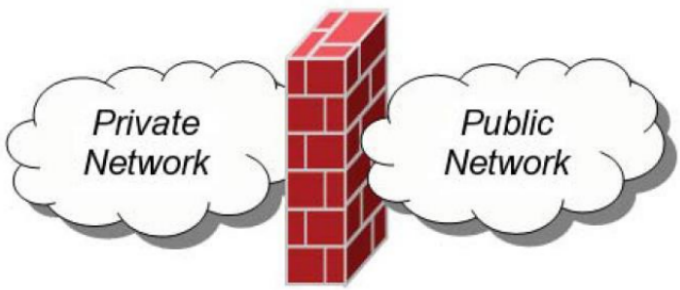
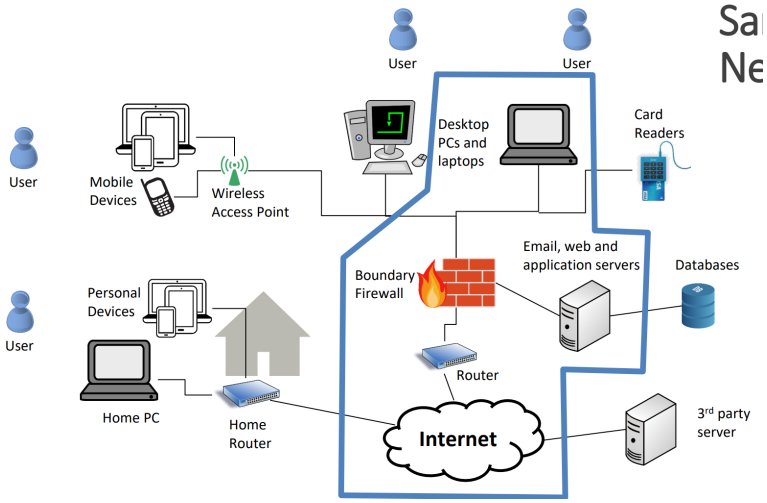
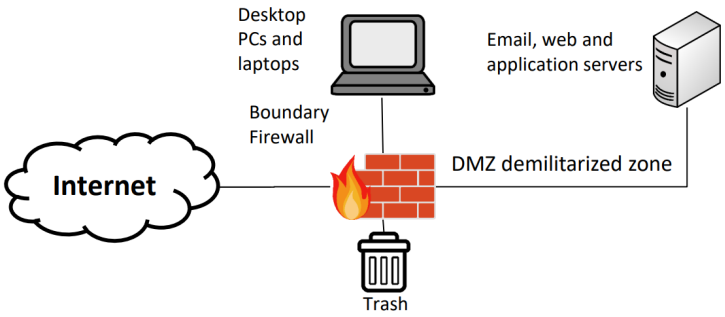
CS Revision Lecture 5, 6

- ## Lecture 5 - Firewalls, NAT and Intrusion Detection
  - ### Methods for observing, managing, and controlling network information flows
    - Firewalls
    - Network Address Translation (NAT)
    - Intrusion Detection System (IDS)
  - ### Firewalls
    - A firewall is a security measures designed to **prevent unauthorized electronic access** to a networked computer system
    - Intuition: Similar to firewalls in building construction. Intent is to isolate one “network” or “compartment “ from another



    - Sample network

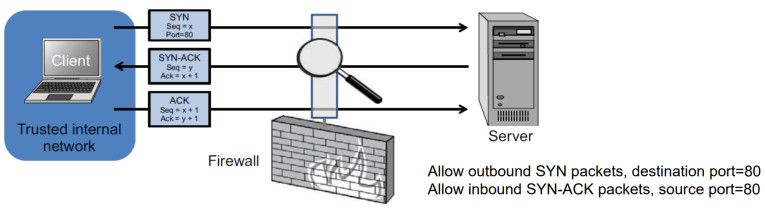




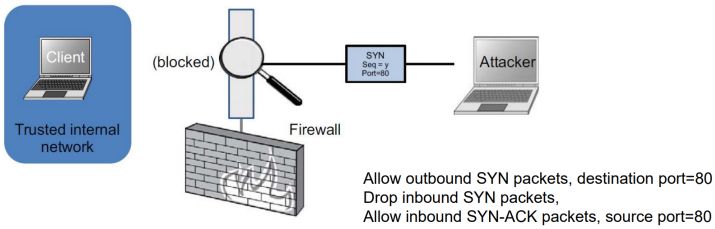
Rule	Type	Source Address	Destination Address	Destination Port	Action
1	TCP	*	192.168.1.*	22	Permit
2	UDP	*	192.168.1.*	69	Permit
3	TCP	192.168.1.*	*	80	Permit
4	TCP	*	192.168.1.18	80	Permit
5	UDP	*	192.168.1.*	*	Deny

- Malicious actions from the Internet AND local network
- Firewall applies a set of rules called firewall policies
- Based on rules, it allows or denies the traffic
- Blocklist: Allow-by-default
- Allowlist: Deny-by-default

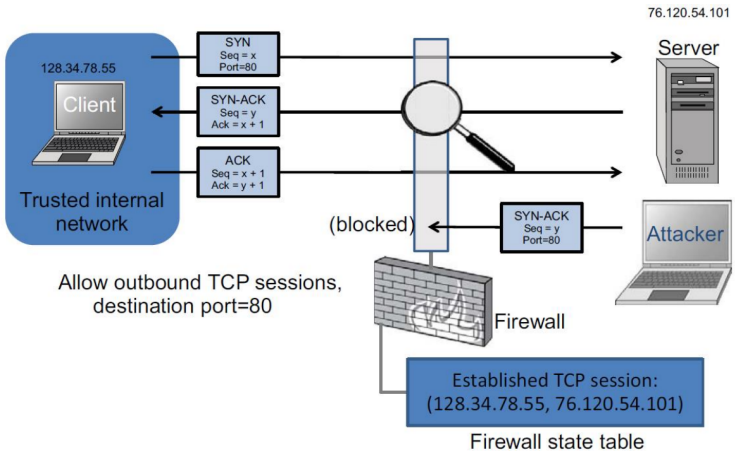
- Firewall Types
  - **Packet filters (stateless)**
    - If a packet matches the packet filter's set of rules, the packet filter will drop or accept it
    - A stateless firewall doesn't maintain any remembered context (or "state") with respect to the packets it is processing. Instead, it treats each packet attempting to travel through it in isolation without considering packets that it has processed previously.



    - Stateless firewalls may have to be fairly restrictive in order to prevent most attacks.



- **Stateful filters**
  - it maintains records of all connections passing through it and can determine if a packet is either the start of a new connection, a part of an existing connection, or is an invalid packet.
  - Stateful firewalls can tell when packets are part of legitimate sessions originating within a trusted network.
  - Stateful firewalls maintain tables containing information on each active connection, including the IP addresses, ports, and sequence numbers of packets.
  - Using these tables, stateful firewalls can allow only inbound TCP packets that are in response to a connection initiated from within the internal network
  - Example
    - Allow only requested TCP connections:



- **Application layer**
  - It works like a proxy it can “understand” certain applications and protocols
  - It may inspect the contents of the traffic, blocking what it views as inappropriate content (i.e. websites, viruses,

vulnerabilities, ...)

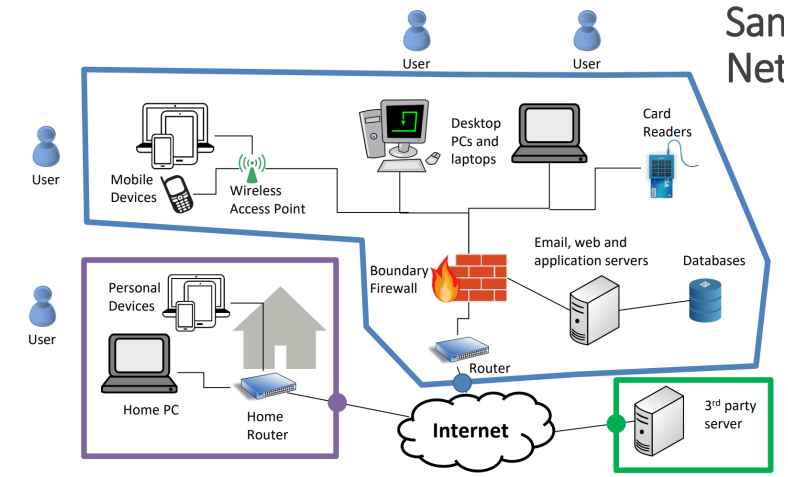
- Simulates the (proper) effects of an application
- Effectively a **protective mal actor-in-the-middle** that screens information at an application layer
- Allows an administrator to block certain application requests.
- For example:
  - Block all web traffic containing certain words (aka censorship)
  - Remove all macros from Microsoft Word files in email
  - Prevent anything that looks like a credit card number from leaving a database

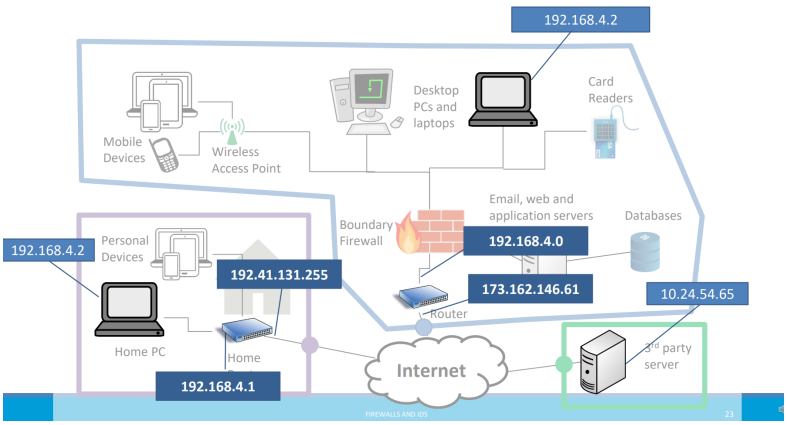
- Personal firewalls
  - Runs on the workstation that it protects (software)
  - Provides basic protection, especially for home or mobile devices
  - Any rootkit type software can disable the firewall
- Firewalls Pros and Cons
  - **They do** prevent straightforward attacks and information leakages
  - **They can be surpassed though**, and may have unintended consequences
  - Increasing their effectiveness increases their operational cost substantially (overhead/configuration).
  - May give false sense of security.
  - **Bottom-line**: you have to have one but do not count on it for much.

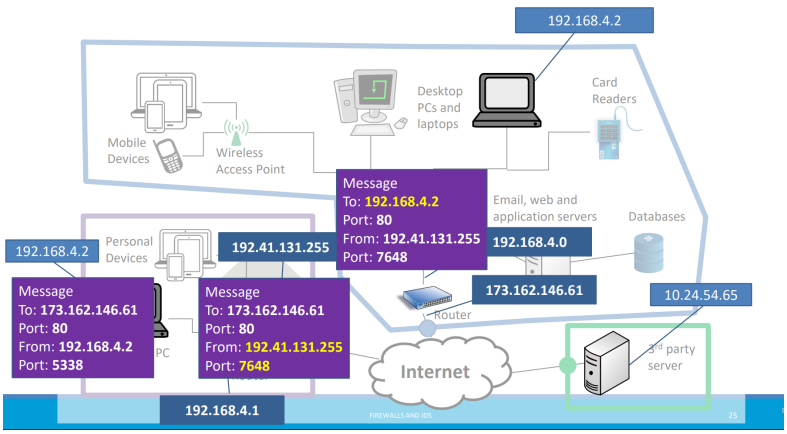
- Network Address Translation (NAT)
  - IPv4 and address space exhaustion
    - There are less than 4.3 billion IPv4 addresses available

- We do not have enough addresses for every device on the planet
- Answer: Network Address Translation
  - **Internal IP different than external IP**
  - **Border router maps between its own IP and the internal ones**
- Alternative answer: IPv6

- Sample Network

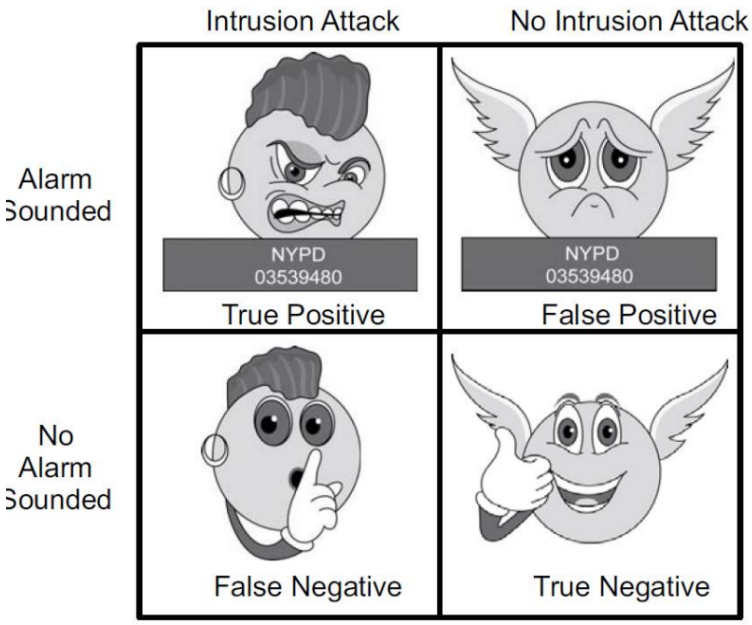






- Intrusion Detection Systems (IDS)

- Firewalls are preventative, IDS detects a potential incident in progress
  - At some point you have to let some traffic into and out of your network (otherwise users get upset)
  - Most security incidents are caused by a user letting something into the network that is malicious, or by being an insider threat themselves
  - These cannot be prevented or anticipated in advance
  - The next step is to identify that something bad is happening quickly so you can address it
- Possible Alarm Outcomes
  - • Alarms can be sounded (positive) or not (negative)



- Rule-Based Intrusion Detection
  - **Rules identify the types of actions that match certain known intrusion attack. Rule encode a signature for such an attack.**
  - Requires that admin anticipate attack patterns in advance
  - Attacker may test attack on common signatures
  - Impossible to detect a new type of attack
  - **High accuracy, low false positives**
- Statistical Intrusion Detection

- **Dynamically build a statistical model of acceptable or "normal" behavior and flag anything that does not match**
- Admin does not need to anticipate potential attacks
- System needs time to warm up to new behavior
- Can detect new types of attacks
- **Higher false positives, lower accuracy**

- Base-Rate Fallacy
  - Suppose an IDS is 99% accurate, having a 1% chance of false positives or false negatives.
  - Suppose further…
    - An intrusion detection system generates 1,000,100 log entries.
    - Only 100 of the 1,000,100 entries correspond to actual malicious events.
    - Because of the success rate of the IDS, of the 100 malicious events, 99 will be detected as malicious, which means we have 1 false negative.
    - Nevertheless, of the 1,000,000 benign events, 10,000 will be mistakenly identified as malicious. That is, we have 10,000 false positives!
    - Thus, there will be 10,099 alarms sounded, 10,000 of which are false alarms. That is, roughly 99% of our alarms are false alarms.
  - Number of alarms is a big problem
    - In the 2013 Target breach the IDS did correctly identify that there was an attack on the Target network
    - There were too many alarms going off to investigate all of them in great depth
    - Some cyberattack insurance policies state that if you know about an attack and do nothing they will not cover the attack.

        - Having a noisy IDS can potentially be a liability
    - Key take-aways
        - Well configured Firewalls are helpful tools to **defend against known attacks**
        - Network Address Translation **allows traffic to flow from routable Internet addresses and private local area networks**, but we have to be careful
        - Intrusion detection systems may be able to detect malicious activity that the Firewall allows (due to usability reasons)
        - A layered approach (Firewalls+IDS) is more resilient (but not perfect!)

Lecture 6 - Cryptography: Introduction

- Introduction
    - What is cryptography?
        - “The practice of creating and understanding codes that keep information secret.”
        - 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.
    - 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;

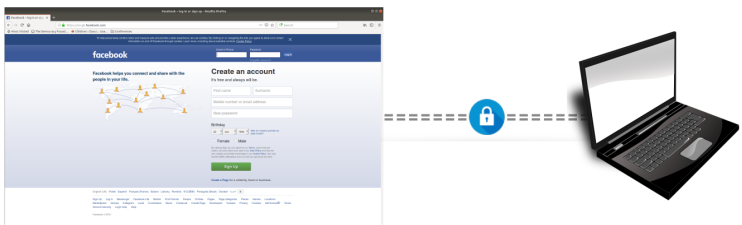
        - to enforce access control in a multi-user operating system;
        - to prevent thieves extracting trade secrets from stolen laptops;
        - to prevent software copying;
        - etc
    - Cryptography (and security more broadly) is becoming a more and more central topic within computer science
- Important remark
    - Cryptography is not:
        - The solution to all security problems
        - 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

Symmetric encryption

Goal: confidentiality

- Secure communications



- File Protection

Symmetric encryption schemes

- A symmetric cipher consists of two algorithms
  - Encryption algorithm $E : K \times M \rightarrow C$

        - decryption algorithm $D : K \times C \rightarrow M$
        - st. $\forall k \in K$, and $\forall m \in M$, $D(k, E(k, m)) = m$
    - **same key** k to encrypt and decrypt
    - the **key k is a secret**: only know 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.
        - 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
    - The attacker know the encryption/decryption algorithms but may have access to :
        - **Ciphertext only attack** - some ciphertexts $c1, \ldots, cn$

- **Known plaintext attack** some plaintext/ciphertext pairs (*m1, c1*), . . . , (m_n, c_n) st. c_i = *E*(*k*, m_i))
- **Chosen plaintext attack** - he has access to an encryption oracle - canmaybe trick a user to encrypt messages m_1, . . . , 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, . . . , c_n of his choice
- unlimited, or polynomial, or realistic (≤ 2^{80}) **computational power**

- Brute-force attack - attack on all schemes
  - Try all possible keys *k* ∈ *K* - requires some knowledge about the structure of plaintext
  - Making exhaustive search unfeasible:
    - *K* should be sufficiently large, i.e. keys should be sufficiently long
    - Keys should be sampled uniformly at random from *K*
- A simple scheme: the substitution cipher
  - shared secret: a permutation π of the set of characters

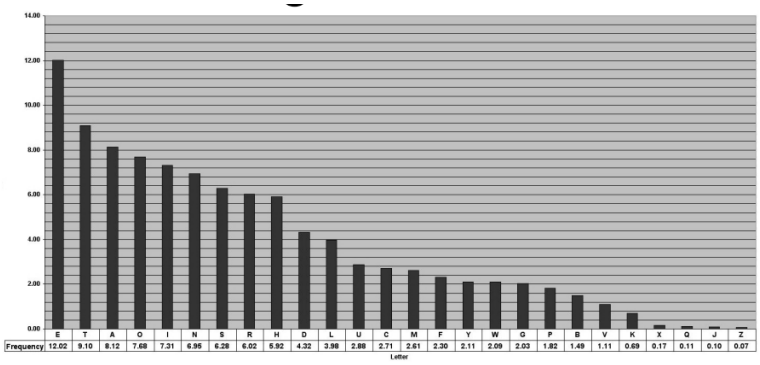
$$\pi = \begin{array}{l} a \mapsto q\ b \mapsto w\ c \mapsto e\ d \mapsto r\ e \mapsto t\ f \mapsto y\ g \mapsto u\ h \mapsto i\ i \mapsto o \\ j \mapsto m\ k \mapsto a\ l \mapsto s\ m \mapsto d\ n \mapsto f\ o \mapsto g\ p \mapsto h\ q \mapsto j\ r \mapsto k \\ s \mapsto l\ t \mapsto z\ u \mapsto x\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

  - **Encryption**: apply π to each character of the plaintext
    - *E*(π, $p_1 \ldots p_n$) = π(p_1). . . π(p_n)
  - **Decryption**: apply π^{-1} to each character of the plaintext
    - D(π, $c_1 \ldots c_n$) = $\pi^{-1}(c_1) \ldots \pi^{-1}(c_n)$
  - Example

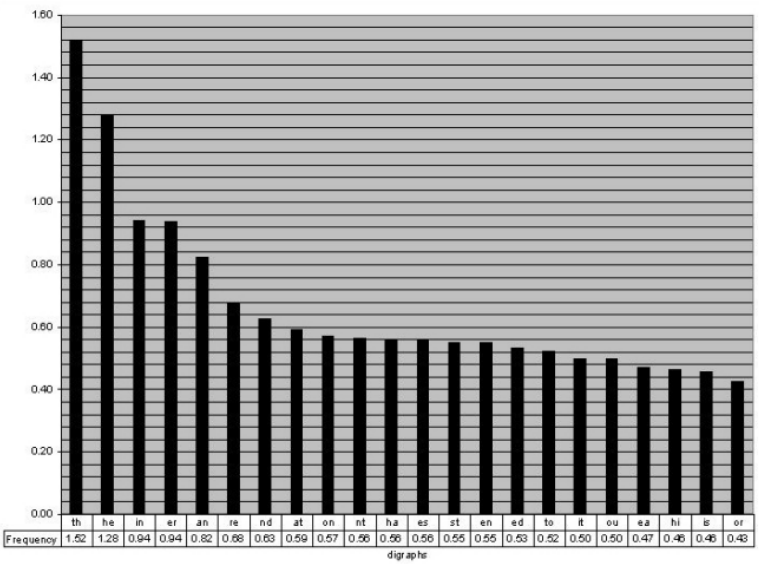
m = THIS COURSE AIMS TO INTRODUCE YOU TO THE PRINCIPLES AND TECHNIQUES OF SECURING COMPUTERS AND COMPUTER NETWORKS WITH FOCUS ON INTERNET SECURITY. THE COURSE IS EFFECTIVELY SPLIT INTO TWO PARTS. FIRST INTRODUCING THE THEORY OF CRYPTOGRAPHY INCLUDING HOW MANY CLASSICAL AND POPULAR ALGORITHMS WORK E.G. 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.

- Breaking the substitution cipher
  - Key space size: $|K| = 26!$ ($\approx 2^{88}$) ⇒ **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

- Example

π =

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

- π = t → z e → t
- Most common digrams in c: of > zi > . . .
- t → z suggests h → i (th → zi)
- guess in → of
- We identify in c the word **INTEKNET**
- guess r → 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 → l
- Going back to letter frequency and a few more guesses!!

π = a ↦ q b ↦ w c ↦ e d ↦ r e ↦ t f ↦ y g ↦ u h ↦ i i ↦ o j ↦ m k ↦ a l ↦ s m ↦ d n ↦ f o ↦ g p ↦ h q ↦ j r ↦ k s ↦ l t ↦ z u ↦ x v ↦ c w ↦ v x ↦ b y ↦ n z ↦ 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.

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

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

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

- **Decryption**: $\forall k \in^* K^*.\ \forall c \in C.\ D(k,c) = k \oplus c$

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

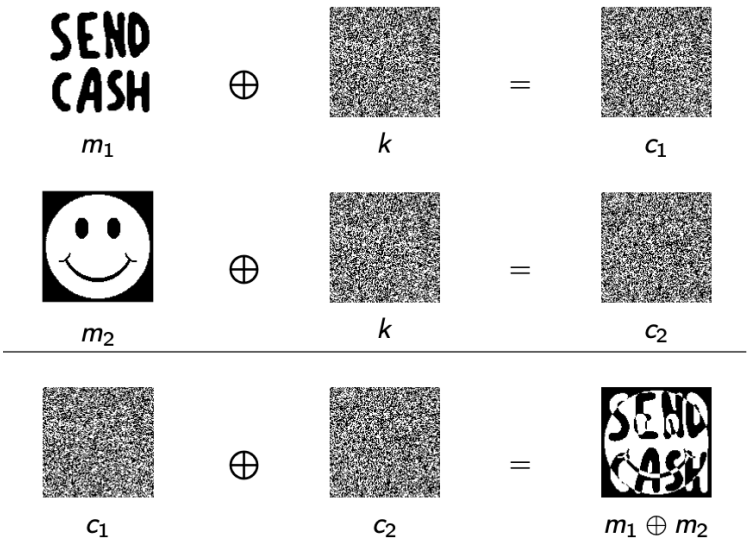
- **Consistency**: $D(k, E(k,m)) = k \oplus (k \oplus m) = m$

- Perfect secrecy
  - A cipher (E, D) over (M, C, K) satisfies perfect secrecy if for all messages $m_1, m_2 \in M$ of same length ($|m_1|=|m_2|$), and for all ciphertexts $c \in C$
    - $|Pr(E(k, m_1) = c) - Pr(E(k, m_2) = c)| \leq \varepsilon$
  - where $k \xleftarrow{r} K$ and ε is some "negligible quantity".
  - $P(M = m \mid C = c) = P(M = m)$ i.e. **seeing a ciphertext doesn't give you any extra information about the plaintext.** The probability of seeing a message m after the ciphertext has been observed is the same as the probability of the message without the ciphertext.
  - $P(C = c \mid M = m_0) = P(C = c \mid M = m_1)$ i.e. **the probability of ciphertext c is equally likely for 2 different messages.**
  - The key is as long as the message and a key should be used uniquely with a probability $1/|K|$ where $|K|$ is the key space.
- OTP satisfies perfect secrecy
  - Proof:
    - We first note that for all messages $m \in M$ and all ciphertexts $c \in C$

- $Pr(E(k, m) = c) = \frac{\#\{k \in K : k \oplus m = c\}}{\#K}$
- $= \frac{\#\{k \in K : k = m \oplus c\}}{\#K}$
- $= \frac{1}{\#K}$

- where $k \xleftarrow{r} K$
- Thus, for all messages $m_1, m_2 \in M$, and for all ciphertexts $c \in C$
  - $|Pr(E(k, m_1) = c) - Pr(E(k, m_2) = c)| \leqslant |\frac{1}{\#K} - \frac{1}{\#K}| = 0$

- Two-time pad 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**

      Can't use the key for two times
      - 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**

  can be modified

  - given the ciphertext $c = E(k, m)$ with m = "to bob : secret msg", it is poosible to compute the ciphertext $c' = E(k, m')$ with m' = "to eve: secret msg"
  - c′ := c ⊕ ”to bob : 00...00” ⊕ ”to eve : 00...00”
    - Cancel out the original ciphertext
    - Replace with new text
    - the rest of 00000 will keep the secret msg the same

- 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?

以上内容整理于 幕布文档