Data and Network Security

Jordan Pyott

2021-05-06 17:02

Contents

Course Information

Lecturers Details

- Lecturer: Dr. Clementine Gritti
  - Office: Erskine 304
  - Email: clementine.gritti@canterbury.ac.nz
- Tutor: Ryan Beaumont
  - Email: rbe72@uclive.ac.nz

Weekly Lab Times

- Tuesday, 16:00–18:00, Jack Erskine 136 Lab 4.
- Wednesday, 14:00–16:00, Jack Erskine 136 Lab 1.
- Friday, 16:00–18:00, Jack Erskine 136 Lab 4.

Other Information

- Labs and Quiz's will be available on learn
- Textbooks
  - Cryptography and network security : principles and practice, William Stallings, 5th edition
    * This course is inspired from this book as the bulk of the course is founded in cryptography.
    * The exam will only be on content in the slides not from the book
  - Computer security : principles and practice, William Stallings and Lawrie Brown, 3rd edition

2.1 Term 1 Plan

Week starting date	Course week	Monday lecture	Thursday lecture	Lab
19/07/2021	1	L1: Course introduction	L2: Course overview	no lab
26/07/2021	2	L3: Discrete mathematics	L4: CrypTool (at home)	Lab 1: Introduction
02/08/2021	3	L5: Classical encryption part 1	L6: Classical encryption part 2	Lab 2: Discrete maths exercises
09/08/2021	4	L7: Block ciphers	L8: Block cipher modes	Lab 3: CrypTool part 1
16/08/2021	5	L9: Stream ciphers	L10: Number theory	Lab 4: CrypTool part 2
23/08/2021	6	L11: Hash functions and MACs	no lecture	Lab 5: Number theory exercises

2.2 Term 2 Plan

Week starting date	Course week	Monday lecture	Thursday lecture	Lab
13/09/2021	7	L12: Public key crypto part 1	L13: Public key crypto part 2	Lab 6: Hash functions and MACs exercises
20/09/2021	8	L14: Digital signatures	L15: PKI and certificates	Lab 7: CrypTool part 3
27/09/2021	9	L16: Key establishment	L17: TLS part 1	Lab 8: PKI and certificates
04/10/2021	10	L18: TLS part 2	L19: IPSec and VPN	Lab 9: Digital signatures and key establishment exercises
11/10/2021	11	L20: Email security	L21: Malware and attacks	Lab 10: TLS
18/10/2021	12	L22: Recap lecture	no lecture	Lab 11: IPSec and email security exercises

Figure 1: Timetable

Assessment

1. Labs (10%) - attendance and participation:

- Labs are done individually but you are encouraged to discuss and share with your peers (you are allowed to see each other during labs).
- Attending one lab each week over the semester automatically gives you full mark: – The tutor will assess your attendance.
- If you cannot attend one lab session, then a report (along with a justification of student absence) will be required and assessed:
  - The report needs to be submitted by one week after the missed session.
  - Example: if you miss Tuesday lab on Week X then you are asked to submit a report by Tuesday of Week X+1.
  - The report needs to be sent to *both* the lecturer and the tutor.

2. Weekly quizzes (20%):

- They can be found and done on LEARN.
- 9 quizzes in total.
- Each quiz contains 10 questions. Each question contains 4 choices such that only one choice is correct.

- 2 attempts per quiz, such that the highest grade is taken into account.
- A quiz is given on Friday of Week X, and should be done before Friday of Week X+1 (except for the one released just before the break):

3. Assignment (20%):

- *Deadline*: 17 September 2021.
- Small exercises on what has been covered so far.
- The assignment will be released on LEARN on 20 August 2021.
- Your report should be uploaded to LEARN.

4. Final exam (50%)

- 3-hours duration
- 25 multiple-choice questions
- 5 open questions, such that if additional information is needed to solve the problem then it will be provided.
- Covers all content from all lectures study *definitions, mechanisms, processes*
  - Not expected to remember the code of each standard (e.g. RFC1234)

Lectures

Lecture One - Course Introduction

- All materials will be found on learn, including lectures, labs, quizzes and assignments.
- Course outline available
- Labs must be done in person or a report *will not get full marks if do not attend*
- Labs start next week
- Weekly quizzes go over two lectures each *multi-choice*
- Midterm and final will all be entirely open questions

Why do we need cyber security

- Privacy
- Security
- Risk management

Famous recent attacks

- Dark hotel attack
  - Targeted phishing attacks using spy-ware

  - Infiltrating guests computers through WIFI networks at hotels
  - Loss of confidentiality

- POODLE attack

  - man in the middle exploit
  - Communications can be decryped and exploited

- EncroChat

  - A communications network and service provider
  - Infiltrated by police in 2020

- WannaCry

  - Loss of availability
  - stolen government hacking tools
  - Worm encrypting files on computers hard drive

    * Was a form of ransom-ware

- Botnet

  - Botnet attacking IoT devices with default admin credentials
  - DDos
  - Loss of availability

Because of these attacks, some users have lost confidence in the service provided not storing/selling data.

Course Focus

- Cryptography as a foundation for information security
- Applications of cryptography
- History of cryptography
- Modern cryptography

  - Block ciphers, stream ciphers
  - public key crypto
  - Hashing and MAC

- Some mathematics

  - Modular arithmetic
  - Number theory
  - Elliptic curves

- Using all of the cryptography
  - Public key infrastructure
  - Secure email
  - TLS (HTTPS)

Lecture Two - Course Overview

What is cyber security?

Definition from the NIST computer security handbook:

- The protection afforded to an automated information system in order to attain the applicable objectives of preserving the integrity, availability and confidentiality of information system resources.
  - Some literature might differentiate between *computer security and cyber security*

Definitions

- A threat
  - Represents a potential security harm to an asset
- An attack
  - is a threat that is carried out
- The threat agent
  - carrying out the attack is referred to as an attacker
- A countermeasure
  - Any means taken to deal with an attack
- A residual level of risk to the assets
  - represented by vulnerabilities possibly exploited by threat agents
- **Assets**
  - Computer systems and other data processing, storage and data communication devices
  - OS, system utilities and applications
  - Files and databases and further data
  - Local and wide area network links
- **Vulnerabilities**

  - A computer system can be:
    * Leaky
      · meaning gives access to information through the network, violates confidentiality
    * Corrupted
      · meaning that it does the wrong thing, violates integrity principle
    * Unavailable
      · meaning that it becomes impossible or impractical to use, violates availability

- Passive attacks
  - Interception
  - Traffic analysis
    * Spoofing, finding information and observing traffic
- Active attacks
  - Altering information and system resources
  - May be hard to prevent but easy to detect
  - Masquerade
    * the attacker claims to be someone else *authorized*
  - Falsification (Man in middle)
    * the attacker changes messages during transmission
  - Misappropriation (DDOS)
    * the attacker prevents legitimate users from accessing resources
- Inside attacks
  - initiated by an entity INSIDE the security perimeter
  - authorization to access system resources but use of them in a malicious way
  - Exposure
    * the attacker intentionally releases sensitive information to an outsider.
  - Falsification
    * the attacker alters or replaces valid data or introduces false data into a file or database.
- Outside attacks
  - initiated from OUTSIDE the perimeter, by an unauthorised or illegitimate user of the system
  - Obstruction
    * the attacker disables communication links or alters communication control information.

  - Intrusion
    * the attacker gains unauthorised access to sensitive data by overcoming the access control protections.

Security functional requirements

- Information security management requires to:

  1. Identify threats
  2. Classify all threats according to liklihood and severity
  3. Apply security controls based on cost benefit analysis

- Countermeasures to vulnerabilities and threats comprise:

  1. Computer security technical measures

  - access control, authentication and system protection

  2. Management measures

  - awareness and training

  3. Both

  - configuration management

Defining Information Security

Definition from the NIST computer security handbook:

- The term security is used in the sense of minimizing the vulnerabilities of assets and resources. An asset is anything of value. A vulnerability is any weakness that could be exploited to violate a system or the information it contains. A threat is a potential violation of security.

The CIA Triad

- Confidentiality
  - Preventing unauthorised disclosure of information
- Integrity
  - Preventing modification or destruction of information
- Availability
  - ensuring resources are accessible when required

Information Security Definitions

- Security Service
  - a processing or communication service to give a specific kind of protection to system resources.
  - Types of security services [Lecture Two - Slide 20/27]:
    * Peer entity authentication
    * Data origin authentication
    * Access control
    * Data confidentiality
    * Traffic flow confidentiality
    * Data integrity
    * Non-repudiation
    * Availability
- Security Mechanism
  - a method of implementing one or more security services.
  - Types of security mechanisms
    * Encipherment
    * Digital signature
    * Access control
      · access control lists, password or tokens which may be used to indicate access rights
    * Data integrity
    * Authentication exchange
    * Traffic padding
    * Routing control
    * Notarization

Risk Management

A key tool in information security management:

1. Identify threats
2. Classify all treats according to likelihood and severity
3. Apply security controls based on cost benefit analysis

Lecture Three - Number Theory and Finite Fields

Factorisation

The set of all integers is denoted by $\mathbb{Z} = ..., -3, -2, -1, 0, 1, 2, 3, ...$, given $a, b \in \mathbb{Z}$, a divides b if there exists $k \in \mathbb{Z} \quad s.t \quad ak = b$.

- This means that a is a factor of b
- $a|b$

We use p to denote a prime, an integer $p \geq 1$ is a *prime* if its divisors are $(1, p)$.

- Testing a prime number p by trial numbers up to the square root of p = There are more efficient ways to check for primality *later in the course*

Properties of factorisation and useful formulae:

- If $a|b$ and $a|c$, then $a|bc$
- If p is prime and $p|ab$ then either $p|a$ or $p|b$
- **Division algorithm**
    - given $a, b \in \mathbb{Z}, \quad s.t \quad a > b$, then there exists $q, r \in mathbbZ \quad s.t \quad a = bq + r$
    - $a = bq + r$ and $0 \leq r \ leqb$, we can use this to show $r < \frac{a}{2}$.
- **Greatest common divisor (GCD)**
    - $gcd(a, b) = d$ if $d|a$ and $d|b$
    - if $c|a$ and $c|b$ then $c|d$
    - a and b are *relatively prime / co-prime* when $gcd(a, b) = 1$
- **Euclidean Algorithm**
    - Find $d = gcd(a, b)$

$$a = bq * 1 + r_1 \quad for \quad 0 < r_1 < b$$

$$b = r_1 q_1 + r_2 \quad for \quad 0 < r_2 < r_1$$

$$r_1 = r_2 q_1 + r_3 \quad for \quad 0 < r_3 < r_2$$

$$...$$

$$f * k - 3 = r * k - 2q * k - 1 + r * k - 1 \quad for \quad 0 < f * k - 1 < f * k - 2$$

$$f * k - 2 = r * k - 1q * k + r * k \quad for \quad 0 < f * k < f * k - 1$$

$$f * k - 1 = r * kq * k + 1 + r * k + 1 \quad with \quad r * k + 1 = 0$$

- Hence $d = r_k = gcd(a, b)$
- **Back Substitution - Extending Euclidean Algorithm**

  - Finding x, y in $ax + ay = d = r_k$
  - This is essentially reversing the Euclidean algorithm

- **Modular Arithmetic:**
- Given $a \equiv b$ (mod n) and $c \equiv d$ (mod n), then the following conditions hold:

  - $a + b \equiv c + d$ (mod n)
  - $ac \equiv bd$ (mod n)
  - $ka \equiv kb$ (mod n)

- **Groups**

  - A group $\mathbb{G}$ is a set with *binary operation* and:
    * Closure: $a \cdot b \in \mathbb{G}$ for $a, b \in \mathbb{G}$
    * Identity: there is an element 1, s.t. $a \cdot 1 = 1 \cdot a = a$ for $a \in \mathbb{G}$
    * Inverse: there is an element b s.t. $a \cdot b = 1$ for $a \in \mathbb{G}$
    * Associativity: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ for $a, b, c \in \mathbb{G}$
    * Commutativity: $a \cdot b = b \cdot a$ for $a, b \in \mathbb{G}$
      · If this condition holds, the group is said to be *abelian*.

- **Cyclic Groups**

  - the order $|\mathbb{G}|$ of a group $\mathbb{G}$ is the number of elements in $\mathbb{G}$
  - g^k denote the repeated application of $g \in \mathbb{G}$, using the group operation
  - The order $|g|$ of $g \in \mathbb{G}$ is the smallest integer k s.t. $g^k = 1$
  - g is a generator of $\mathbb{G}$
  - A group is said to be *cyclic* if it has a generator

- **Finding inverse**

  - Use extended Euclidean algorithm if GCD is 1.

- **Fields**

  - A field $\mathbb{F}$ is a set with binary operations:
  - $\mathbb{F}$ is an *abelian group* under the operation $+$ with identity element of 0.

- **Finite Fields**

  - Setting up secure communications requires fields with a finite number of elements.
  - Notation is $GF(p) = \mathbb{Z}\_p$
  - Addition modulo 2: XOR LOGIC

Lecture Four - CrypTool

What is cryptool?

- Open source program that focuses on the free e-learning software, illustrating cryptographic concepts.
- We will be using CrypTool 2

Video to go back and watch

Lecture Five - Classic Encryption (Part One)

Goals:

- Study historical ciphers
- Establish basic notation and terminology
- Introduce basic cryptographic operations
- Explore typical attacks and adversary capabilities

Terminology:

- Cryptography: The study of designing crypto-systems, key components are:
  - Confidentiality: A key is needed to *read* the message
  - Authentication: A key is needed to *write* the message
- Cryptanalysis: The study of breaking crypto-systems
- Stenography: The study of concealing information
- Crypto-system
  - A set of plain-texts
  - A set of cipher-texts
  - A set of keys
  - A function called *encryption* which transforms plain text to cipher text
  - An inverse function called *decryption*, which reverses the encrypted message to plain-text
- Symmetric and Asymmetric Cryptography
  - Symmetric key cipher:
    * Encryption and decryption keys are known to the sender and receiver
    * Secure channel for transmission of the keys
  - Asymmetric key cipher
    * Each participant has a public key and a private key
    * Possibly working for both encryption of messages and the creation of digital signatures

Symmetric Encryption **Notation for Symmetric Encryption Algorithms:**

- Encryption function: E
- Decryption function: D
- Message or plain-text: M
- Cipher-text: C
- Shared secret key: K

Encryption is denoted as $C = E(M, K)$
Decryption is denoted as $M = D(C, K)$

Methods to break Symmetric encryption:

An adversary has access to many methods to break a cryptosystem, however such methods depend on some conditions and known knowledge, such as:

- What are the resources available to the adversary?
  - Examples: *computational capability, inputs/outputs, knowledge of the crypto-system*
- What does the adversary want?
  - Do we want to know the secret key, distinguish two messages, **what are we trying to achieve?**.

Exhaustive Key Search:

This is a brute force attack, that checks all combinations of possible keys, note there is no way to prevent such an attack, our best option is just to raise complexity to where it is to combinationally difficult to compute, however this does not stop people from getting lucky when breaking your system.

- We may be able to break the system without exhaustive search.
- We may be able to break the system without finding the key.

Attack classifications:

These are ordered from the least powerful to the most dangerous information an attacker could have.

1. Cipher-text Only Attack - the attacker only has access to intercepted encrypted messages
2. Known Plain-text Attack - The attacker knows some of the plain text
3. Chosen Plain-text Attack - The attacker can obtain the cypher-text from some plain-text that has been selected
4. Chosen Cipher-text Attack - The attacker can obtain the plain-text from some cipher-text that it has selected

Which of these should be prevented?

- A crypto-system is seen as very insecure if it can be practically attacked using only intercepted cypher texts
- A crypto-system should be secure against chosen plain-text and chosen cipher-text attacks (*this is the modern standard*)
- History shows that chosen cipher-text attacks are practical to set up for an attacker

Kerchhoffs' Principle: We assume an attacker has complete knowledge of the cipher, the only item that is unknown about the crypto-system is the private key.

- Using a secret, non-standard algorithm can cause severe problems
  - This would be an example of *security through obscurity* (note this is a bad habit)

Statistical Attacks:

- Depends on using the redundancy of the alphabet
- Information from distribution of letters
- Recognising patterns to guess information about the crypto-system

Basic cipher operations:

Historical ciphers combine two basic operations:

- Transposition: Characters on the plain-text are mixed up with each other
  - Permutating characters in a fixed period d and a permutation f
  - Plain-text is seen as a matrix of rows of length d
  - Key is (d, f)
  - Each block of d characters is re-ordered using permutation f
  - Complexity is $d!$
  - We will go through how to do this by hand and a formulation of how to automate this process using Crypt Tool
- Substitution: Characters are replaced by a different character (or set of characters)
  - Each character in the plain-text alphabet replaced by a character in the cipher-text alphabet, following a cypher-text table.
  - *Caesar cipher* is a simple example of a substitution cipher
    * Functions for encryption $C_i = (M_i + j) mod n$
    * Functions for decryption $M_i = (C_i + j) mod n$
    * To break this we can do a frequency analysis in order to find common characters (*example, use ' ' as most common letter*)

Lecture Six - Classic Encryption (Part Two)

Defining Polyalphabetic Substitution

- Using multiple mappings from plaintext to ciphertext
- The effect with multiple alphabets is to smooth frequency distrabution
  - Direct frequency analysis should no longer be effective
- Typical polyalphabetic ciphers are periodic substitution ciphers based on period d
- Given d ciphertext alphabets $C_0, C_1, ..., C_{d-1}$ let $f_i : A \to C_i$ be a mapping from the plaintext alphabet A to the i_{th} ciphertext alphabet $C_i \forall 0 \leq i \leq d-1$
- Encryption Process:
  - Plaintext: $M = M_0...M_{d-1}M_d...M_{2d-1}M_{2d}...$
  - is encrypted to:
    * $E(K, M) = f_0((M_0))...f_{d-1}(M_{d-1})f_0(M_d)...f_{d-1}(M_{2d-1})f_0(M_{2d})...$
  - Special case with $d = 1$: the cipher is monoalphabetic (simple substitution cipher)
- Key generation:
  - Select a block d
  - Generate d random simple substitution table
- Encryption:
  - Encrypting the character by using the substitution table number j such that $i \equiv j(modd)$
- Decryption
  - Using the same substitution table as in encryption in order to reverse the simple substitution

Let $d = 3$, thus there are 3 ciphertext alphabets.

Pltxt char.	ABC	DEF	GHI	JKL	MNO
C_1	UWY	SX∇	TVZ	CEI	AFG
C_2	QLM	PJO	RKN	∇XS	YUW
C_3	MLQ	RNQ	GFA	ZVT	YWU

Pltxt char.	PQR	STU	VWX	YZ∇	
C_1	BDH	KNR	JOP	LMQ	
C_2	ZVT	FGA	HDB	EIC	
C_3	POJ	HDB	IEC	∇XS	

If the plaintext is IT∇IS∇A∇BEAUTIFUL∇DAY then the ciphertext is ZGSZFSUCLXQBNNKRSSSQ∇.

Figure 2: Polyalphabetic example

Vigenere Cipher

- Popular form of periodic substitution ciphers based on *shifted* alphabets
- The key K is a sequence of characters
  - $K = K_0K_1...K_{d-1}$

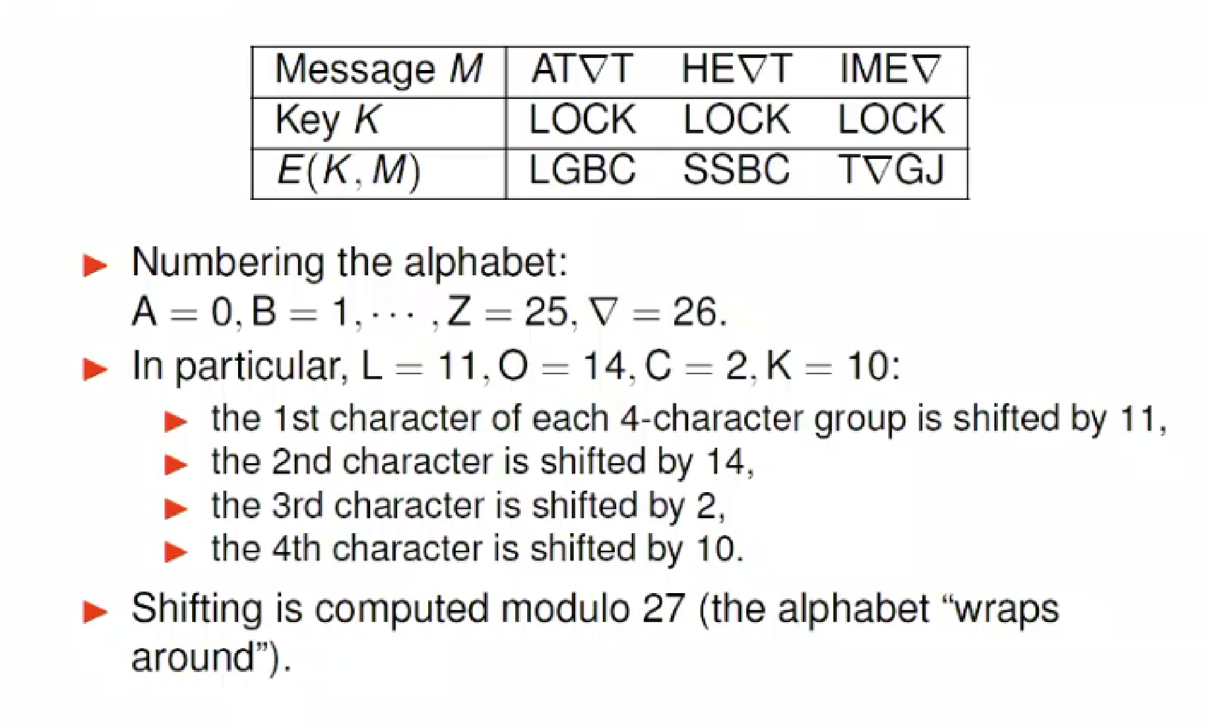

Figure 3: Vigenere example

Crypto-analysis:

- Identify the period length
  - Kasiski method
  - Cryptool uses autocorrelation to estimate the period
- Attack separately d substitution tables
  - Each substitution is just a shift
    - If there is sufficient ciphertext then it is trivial

Autocorrelation method

- Method used to find the period length d of any periodic polyalphabetic cipher
- Given ciphertext C, computing the correlation between C and its shift C_i for all values of i, of the period
- Seeing peaks in the value will help to find i

Hill Cipher

- THe americal mathematition Lester S. Hill published his cipher in 1929

- *Polygram cipher*
  - Simple substitution cipher on an extended alphabet consisting of multiple characters
  - Example: Diagram substatution in which the alphabet consists of all pairs of characters
- Major weakness: its linearity, hence known plaintext attacks are easy

Performing a linear transformation on d plaintext characters to get d ciphertext characters:

- Encryption involves multiplying a $d \times d$ matrix K by the block of plaintext M.
  - $C = KM$
- Decryption involves multiplying the matrix K^{-1} by the block of ciphertext C
  - $M = K^{-1}C$