3010 Submodule Applied Crypto

Dr Tay Kian Boon

NTU, SCSE, Week 10B

2022/2023 Semester 2

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Copyright Notice

All course materials (including and not limited to lecture slides, handouts, recordings, assessments and assignments), are solely for your own educational purposes at NTU only. All course materials are protected by copyright, trademarks or other rights.

All rights, title and interest in the course materials are owned by, licensed to or controlled by the University, unless otherwise expressly stated. The course materials shall not be uploaded, reproduced, distributed, republished or transmitted in any form or by any means, in whole or in part, without written approval from the University.

You are also not allowed to take any photograph, video recording, audio recording or other means of capturing images and/or voice of any of the course materials (including and not limited to lectures, tutorials, seminars and workshops) and reproduce, distribute and/or transmit in any form or by any means, in whole or in part, without written permission from the University.

Appropriate action(s) will be taken against you (including and not limited to disciplinary proceedings and/or legal action) if you are found to have committed any of the above or infringed copyright.

1 August 2022

Overview Week 10 Lectures-B

- Randomness
- WW2 machine ciphers
  - **PURPLE** (Japan) (wont cover)
  - **ENIGMA** (Germany)
- Stream ciphers
- Intro to Block Ciphers
- Intro to AES

Randomness –Birthday Paradox

- Assuming birthdays independent. How many people is needed in a room where you can find a chance of >50% of same birthday
- Ans: only 23 (tutorial 2)

Randomness –Birthday Paradox

- 23 seems to be a paradoxically small number since we have 365 possible dates for birthdays
- Note 23 approx 1.2*sqrt365

Randomness –Birthday Paradox

- In early IPOD days, some listeners complained hearing same song within 2 hours although they have 400 songs on their ipod. Assuming 4 min songs on average.
- Question: Is IPOD shuffling random? (tutorial)

Historical WW2 Ciphers

- After one-time pad (which is difficult to produce), during WW1 & WW2 period, military from many big countries begin to conceive of making machine encryptors for their use

- 2 most famous ones:
  - **PURPLE ciphers (Japan)** –wont cover
  - **ENIGMA ciphers (Germany)**

What is Cryptography - Informal

- As the German military grew in the late 1920s, it began looking for a better way to secure its communications.
- They built new cryptographic machine called **'ENIGMA.'**
- **It is polyalphabetic.**
- They believed the encryption generated by the machine to be unbreakable. With a theoretical number of ciphering possibilities of 3 × 10^114, their belief was not unjustified.

ENIGMA

- They say Necessity is the MOTHER of invention.
- Germany wanted to swallow up many, Poland one of the first.
- Germans did not count on tiny Poland to break ENIGMA!
- WW2 lasted 6 years with 80 million casualties:
- Experts believed breaking of Enigma cut short the war by at least 2 years, thus saving 27 million lives!
- How did the Poles do it?

ENIGMA

- Determining the exact wiring of each of the three rotors became the Polish cryptanalysts' first task.
- Poland's cipher bureau tested and hired three mathematicians in 1932.
- **Marian Rejewski**, Jerzy Rozycki, and Henryk Zygalski painstakingly analyzed the intercepted encrypted messages searching for clues.
- Rejewski eventually determined a mathematical equation that could find the wiring connections.
- However, the equation had too many unknown variables.
- He finally made the initial breaks into the wiring sequence only with the aid of a German traitor!

ENIGMA

- **Hans Schmidt** provided the French with documentation on the Enigma machine and some Enigma keys.
- Unfortunately, the information did not contain wiring diagrams for the rotors.
- France Capt Bertrand arranged a mtg the Polish cryptologic agency in Dec 1932.
- Rejewski studied the necessary complicated mathematical equations to determine the wiring of the Enigma rotors.
- Initially, there were too many unknown variables

ENIGMA

- In 1939 the Poles decided to inform the British of their successes.
- British team- Dilly Knox Tony Kendrick, Peter Twinn, **Alan Turing** and Gordon Welchman.
- They worked at Bletchley Park and that is where the first wartime Enigma messages were broken by the British in January 1940.
- Enigma traffic continued to be broken routinely at Bletchley Park for the remainder of the war.

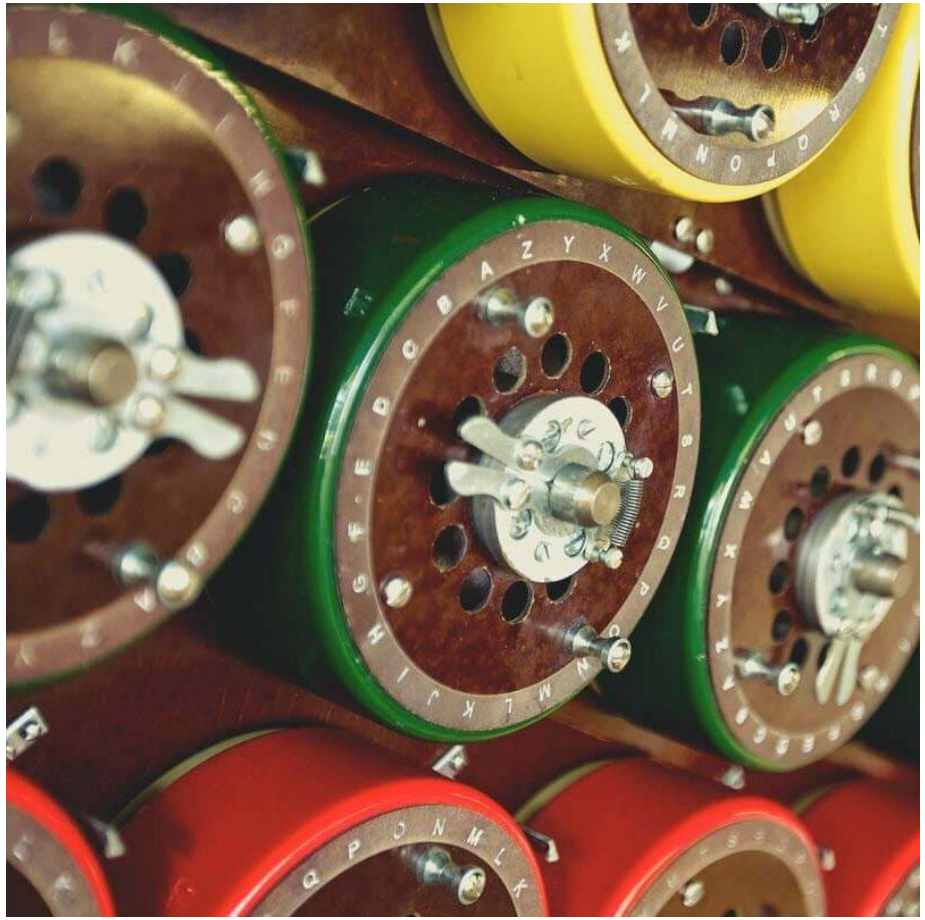
WW2 Fought Between them?

WW2 Fought Between them & Hitler!

Rejewski (1905-1980)

Turing (1912-1954)

Bletchley Park: Bombes Assembly



Facts abt Bombes

Why were the Bombes needed?

- Bletchley Park was set up to decode intercepted Nazi Engima messages. These devices typically changed settings every 24 hours and with 159 quintillion possible combinations every day, the staff at Bletchley Park worked around the clock to break the settings by hand. A mechanical method for identifying the keys was needed and Alan Turing designed the Bombe to speed up the process.

Facts abt Bombes

- **2. What was the impact of the Bombes?**
- By speeding up the process of breaking the day's Enigma settings, Turing's invention meant staff were able to decode quickly and pass on intelligence – often with enough time for it to be acted upon.
- **3. In which areas of the war did they have success?**
- The use of Bombes in intelligence gathering had a huge impact across many land, sea and air campaigns. The German battleship Bismarck was located with the assistance of Enigma decrypts and sunk by air and surface attack in 1941. Later, in 1944, Enigma decrypts provided details of German defensive preparations for, and reactions to the D-Day invasion.

Facts abt Bombes

- **5. What was the legacy of Turing's creation?**
- The Bombes represented the first mass production of a specially designed cryptanalytical machine. They heralded the industrialisation of codebreaking and the intelligence they provided was crucial to Allied success in WW2. They were a significant part of the Bletchley Park operation, which was so successful that the Germans remained unaware the information sent on their "unbreakable" Engima machines had actually been cracked by the Allies.

The Enigma Machine: Overview

Attributed to German military during World War II

Invented by Arthur Scherbius at the end of WW-I

Key : Settings for the machine components

Wheel order, Ring settings, Plug connections, etc.

Initialization : Rotor positions chosen by operator.

Cracking Enigma

- Cannot just be brute-forced (huge key-space)
- Extremely complicated mathematical analysis
- Needs a huge amount of computing capability
- Almost impossible without “known plaintexts”
- Kudos to Marian Rejewski et al. (1932-1939)
- And of course, Alan Turing et al. (1939-1945)

Reading : Cracking Enigma in 2021 : https://youtu.be/RzWB5jL5RX0

Factors Leading to Break of ENIGMA

- Complacency
- Careless implementation by Germans
  - Reduced settings
- Espionage
- First rate mathematicians & computer scientists
- Rich Budget

Post-WWII History

- Claude Shannon — father of the science of information theory
- Computer revolution — lots of data to protect
- Data Encryption Standard (DES), 70's
- Public Key cryptography, 70's
- CRYPTO conferences, 80's
- Advanced Encryption Standard (AES), 90's
- The crypto genie is out of the bottle…

Claude Shannon

- The founder of Information Theory
- 1949 paper: *Comm. Thy. of Secrecy Systems*
- Fundamental concepts
  - **Confusion** — obscure relationship between plaintext and ciphertext
  - **Diffusion** — spread plaintext statistics through the ciphertext
- Proved one-time pad is secure
- One-time pad is confusion-only, while “double transposition” is diffusion-only

Real-World One-Time Pad-Vernam

- Project VENONA
  - Soviet spies encrypted messages from U.S. to Moscow in 30's, 40's, and 50's
  - Nuclear espionage, etc.
  - Thousands of messages
- Spy carried one-time pad into U.S.
- Spy used pad to encrypt secret messages
- Repeats within the "one-time" pads made cryptanalysis possible

SYMMETRIC CRYPTOGRAPHY

- **Symmetric Key**
  - Same key for encryption and decryption
  - 2 Modern types: **Stream ciphers, Block ciphers**
- **Stream ciphers** — 'generalize' one-time pad
  - Except that key is relatively short
  - Key is stretched into a INFINITE (periodic) **keystream**
  - Keystream is used just like a one-time pad, **XOR the keystream with plaintext bit by bit!**
- **Block Ciphers** – later

Stream Ciphers

Stream Ciphers

- Once upon a time, not so very long ago… stream ciphers were the king of crypto
- Today, not as popular as block ciphers
- We'll discuss some main examples stream ciphers:
- LFSRs - A5/1
  - Based on shift registers
  - Used in GSM mobile phone system (2G)
- RC4
  - Based on a changing lookup table
  - Used in many places (in the past, less often now)
- GRAIN – NFSR (secure non-linear feedback registers)

Stream Ciphers

- From Key K

1. Generate pseudorandom bits (by specific algorithm) –therefore deterministic, therefore not truly random

2. Then encrypt the plaintext by XORing it with the generated pseudorandom bits…

Stream Ciphers

- Stream ciphers resembles deterministic random bit generators (DRBGs) than they are to full-fledged pseudorandom number generators (PRNGs) because, like DRBGs, stream ciphers are deterministic.

- Stream ciphers' determinism allows you to decrypt by regenerating the same pseudorandom bits used to encrypt.

Stream Cipher Operations

- A stream cipher computes $KS = \mathbf{SC}(K, N)$, encrypts as $C = P \oplus KS$, and decrypts as $P = C \oplus KS$.

- The encryption and decryption functions are the same because both do the same thing—namely, XOR bits with the keystream.

Stream Ciphers –Most common class

- Feedback Shift Registers (FSR)
- Will now explain the basic mechanism behind hardware stream ciphers, called ***feedback shift registers (FSRs).***
- Almost all hardware stream ciphers rely on FSRs in some way, whether that's
  - the A5/1 cipher (encryption algo in 2G mobile phones) or
  - the more recent cipher Grain-128a.

eStream Project

- The eSTREAM project was a multi-year effort, running from 2004 to 2008, to promote the design of efficient and compact stream ciphers suitable for widespread adoption.
- As a result of the project, a portfolio of new stream ciphers was announced in April 2008. The eSTREAM portfolio was revised in September 2008, and currently contains seven stream ciphers.
- This website (below) is dedicated to ciphers in this final portfolio. For information on the eSTREAM *project* and selection process, including a timetable of the project and further technical background, please visit the original eSTREAM Project website.

eStream Finalists

Profile 1 (SW)	Profile 2 (HW)
HC-128 (Wu Hong Jun,SPMS)	Grain v1
Rabbit	MICKEY 2.0
Salsa20/12	Trivium
SOSEMANUK	

Stream Ciphers

- Stream ciphers were popular in the past
  - Efficient in hardware
  - Speed was needed to keep up with voice, etc.
  - Today, processors are fast, so software-based crypto is usually more than fast enough
- Future of stream ciphers?
  - Expert Shamir declared “the death of stream ciphers” (esp if its linear)

Intro to Block Ciphers
Intro to AES

Block Ciphers

- Block ciphers, which operate on an n-bit block of plaintext are some of the most powerful, fastest, and most used cryptosystems in existence.
- Unlike public-key systems, which obtain their security from a well-known hard mathematical problem, block ciphers are secret key systems based on bit operations, and obtain their strengths from a mixture of non-linear operations, such as substitutions, and permutations.
- This Section introduces the block ciphers, and investigates the most important ciphers DES (historical, 1977) & AES (2000-present day).

Block Ciphers: Overview

- Rounds and key schedules.
- Modes of encryption, which are ways in which a block cipher can be used to encrypt a ciphertext longer than a single block.
- The Feistel construction, which underlies many modern block ciphers.
- The Data Encryption Standard (DES,1977) which was in use for about 30 years up to 2000, and has been intensively analyzed.
- The Advanced Encryption Standard (AES, 2000), winner of an international block cipher competition.

Block Ciphers

- A *block cipher* may be considered as two related functions, encryption and decryption, each of which takes two inputs.
- Input to encryption function are a plaintext block and a key K
- Input to decryption function are the ciphertext block, same key K
- Both plaintext and ciphertext blocks will have the same length.
- Denote the encryption function by E(P, K) and the decryption function by D(C, K).

Block Ciphers

1. The plaintext should be recoverable from the ciphertext; thus D(E(P, K), K) = P.

2. Given a ciphertext and a complete working knowledge algorithm, there should be no feasible way of recovering the plaintext. (If the plaintext were easy to recover, the cipher is WEAK!)

3. The functions should preferably be fast in both hardware and software.

Block Ciphers

- Block ciphers may be distinguished by the length of the blocks of plaintext and ciphertext (this is called the *block size*), and the length of the key.
- Basic aspects of block cipher security is that the key be at least 128 bits long; 256-bits if we want to thwart quantum computers
- Ciphers with smaller keys (esp < 64-bits) are vulnerable to brute force attacks.
- 56-bit cipher – only 7 days BF to crack via FPGA hardware crackers. Much shorter time via ASIC chip crackers.

Block Ciphers

- Almost all block ciphers work by applying a mixing function to the plaintext & key, and then applying that function to the output.
- Each application of the mixing function is called a *round*.
- The input to each round consists of the block obtained from the previous round, and a sequence of subkey-bits obtained in some way from the original key.
- These are the *round keys*.
- The method by which the round keys (subkeys) are determined from the original key is called the *key schedule*.

Block Ciphers

- A general schema is shown in Figure 8.1. It is necessary to choose the number of rounds to obtain a good level of security, but not so many as to slow down the encryption.
- It is also necessary for the round keys all to be different for maximum security.
- One shud not be able to derive future round keys purely from preceding round keys (kind of independence is hoped for)

Block Ciphers: Confusion & Diffusion

- These terms describe how well a cipher mixes the bits from the plaintext and the key.
- Informally, **diffusion** describes how a change in the plaintext affects the ciphertext.
- A small change in plaintext resulting in a large change in ciphertext shows a high level of *diffusion*.
- For example Vigenere ciphers provide LITTLE diffusion: a single change in the plaintext will result in only that particular character of the ciphertext changing.

Block Ciphers: Confusion & Diffusion

- *Confusion* is the property that key & ciphertext are not easily related; in particular that each character of the ciphertext should depend on many parts of the key.
- So if a single character of the key is changed, the ciphertext should change.
- In an ideal secure cipher, changing one character of the key will change *all* characters of the ciphertext.