MATH 4161 Mathematics of Cryptography **Lecture 1**
Lecture Joe Tran

Contents

1. What this class is about

This class will cover topics including

- Probability theory
- Information theory
- Number theory

2. Course Outline

2.1. **Course Description.** Cryptography deals with the study of making and breaking secret codes.

In this course we will be studying situations that are often framed as a game between three parties: a sender (e.g., an embassy), a receiver (the government office) and an opponent (a spy). We assume that the sender needs to get an urgent message to the receiver through communication channels which are vulnerable to the opponent. To do this communication, the sender and receiver agree in advance to use some sort of code which is unlocked by a keyword or phrase. The opponent will be able to intercept the message. Is he/she able to unlock the message without knowing the key?

In this course we will learn some probability theory, information theory and number theory to answer questions about how vulnerable the methods of sending secrets are. This has a great number of applications to internet credit card transactions, wireless communication and electronic voting. We will start by learning some classical codes (used up through WWI) and analyzing those. The last third of the course we will start to learn the methods that are used in modern cryptography.

"Few persons can be made to believe that it is not quite an easy thing to invent a method of secret writing which shall bate investigation. Yet it may roundly be asserted that human ingenuity cannot concoct a cipher which human ingenuity cannot resolve" (Edgar Allan Poe, 1809-1849)

2.2. **Course components.**

- 50% Assignments.
  - Email then submit on eClass.

  - Roughly 20, best 15 of 20 for this grade remaining are extra credit towards final.
- 30% Tests (Open book)
  - 3 tests 10% each
  - Test 1: Probability theory
  - Test 2: Information theory
  - Test 3: Number Theory
- 20% Final Exam (not comprehensive, open book)

Tests and exam: open books, open notes, closed friends and enemies. Open computers.

3. Vocabulary

- **Cryptography:** The art of secret writing.
- **Plaintext:** Text to be encoded for secrecy.
- **Ciphertext:** Encoded text. Short ciphertext is sometimes called a "cryptogram"
- **Cipher:** A method of secret writing.
- **n-gram:** A string of n-letters.
- **Encipherment, Encryption:** The process of encoding plaintext into ciphertext.
- **Decipherment, Decryption:** The process of decoding ciphertext back into plaintext.
- **Encrypt, Decrypt:** These are corresponding verbs.
- **Sender:** The person or organization that is to send the encrypted message.
- **Receiver:** The person or organization which is to receive and decrypt the message.
- **Opponent:** The person or organization which intercepts the message and attempts the unauthorized decipherment.
- **Key:** The information, usually a sequence of digits or symbols, used to determine the method by which plaintext is to be transformed into ciphertext.
- **Cryptographic System, Encipherment Scheme:** A family of ciphers (transformations of plaintext into cipherment to be used for encryption and decryption). Each member of the family is determined by a particular key.
- **Message Space:** The collection of all messages that may occur in a particular cryptographic transaction.
- **Key Space:** The collection of all keys that may occur in a given cryptographic system.
- **Cryptanalysis:** The process by which the opponent attempts to recover the original plaintext from the intercepted ciphertext.
- **Code Breaking:** The process by which a cryptographic system is made vulnerable to cryptanalysis.
- **One-Time Pad:** A key to be used only once.

In a typical cryptographic transaction the sender and receiver choose a cryptographic system and, at some time before the message is to be sent, the sender chooses the key. This determines which transformation of the system will be used to encrypt the message. The key is then sent to the receiver by some safe path (inaccessible to the opponent). Upon obtaining the key the receiver determines which transformation of the system is to be used to decrypt the message.

A number of assumptions are usually made without explicit mention about cryptographic transactions. It is assumed that safe paths between sender and receiver do exist though generally they may be impractical to use for the message itself (for instance the path may require hand carrying by an especially trusted messenger). While, for practical reasons (such as speed of delivery for instance) the path taken by the message itself may have to be "unsafe". Furthermore, security of the message is not usually expected to be achieved through the opponent's ignorance of the encryption system but rather from lack of knowledge as to which particular transformation of the family has been used in the encipherment. That is, the opponent's task usually consists of reconstructing the

key from an analysis of the ciphertext. Security is achieved by assuring that the key space is too large for an exhaustive trial and error attack to be practical.

Of course, the basic goal of the opponent is to recover the original plaintext. This may not necessarily involve reconstructing the key.

The two main methods of encryption are Substitution and Transposition and most known modern methods are a mixture of both. These two methods may be described as follows:

4. Introductory Cyphers: Caesar and Vigenere

Definition 1 (Substitution)**.** When individual letters or n-grams of plaintext are replaced by letters or n-grams of ciphertext.

Example 1. Consider the following text:

THE END OF THE WORLD AS WE KNOW IT

and substitute the following letters to transform the text to

WKH HQG RI WKH ZRUOG DV ZH NQRZ LW

Definition 2 (Transposition)**.** When the characters of the original message are rearranged according to some particular pattern.

Example 2. Consider the following text:

FRANKLY MY DEAR

and applying a transposition

MADLA RKYEN FRY

Definition 3. The *cypherspace* is the set of all messages or keys in a space. No messages in this space will overlap. (Visual representation of an encipherment scheme that is vulnerable to attack)

Let $C(m, k)$ be the encrypt m using encipherment scheme C using key k. If $C(m, k) = C(m', k')$, then it is impossible to tell from the cyphertext alone whether a cyphertext corresponds to m or m'. In terms of math, can be thought of as an one-to-one function.

Example 3. Given the plaintext:

TOBEO RNOTT OBE

And suppose we're performing the following shift:

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
R S T U V W X Y Z A B C D E F G H I J K L M N O P Q
```

Then the above plaintext becomes the cyphertext given by

KFSVF IEFKK FSV

Then observe that (look vertically)

```
K   L M N O P Q R S T U V W X Y Z A B C D E F G H I J
F   G H I J K L M N O P Q R S T U V W X Y Z A B C D E
S   T U V W X Y Z A B C D E F G H I J K L M N O P Q R
V   W X Y Z A B C D E F G H I J K L M N O P Q R S T U
F   G H I J K L M N O P Q R S T U V W X Y Z A B C D E
```

```
I   J K L M N O P Q R S T U V W X Y Z A B C D E F G H
E   F G H I J K L M N O P Q R S T U V W X Y Z A B C D
F   G H I J K L M N O P Q R S T U V W X Y Z A B C D E
K   L M N O P Q R S T U V W X Y Z A B C D E F G H I J
K   L M N O P Q R S T U V W X Y Z A B C D E F G H I J

F   G H I J K L M N O P Q R S T U V W X Y Z A B C D E
S   T U V W X Y Z A B C D E F G H I J K L M N O P Q R
V   W X Y Z A B C D E F G H I J K L M N O P Q R S T U
```

The `TOBEO RNOTT OBE` is located at the 9th column, which we basically observe the English word that is being produced each time.

5. Further Notes: Caesar and Vigenere Substitutions

Suppose we are given the following plaintext given by

A penny saved is a penny earned

and we consider the following substitution given by the following

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

to encrypt it. We first ignore the spaces and write the text in uppercase

APENNYSAVEDISAPENNYEARNED

and we substitute each letter with the letter below it in the above substitution. So then our cyphertext will then become

DSHQQBVBYHGLVDSHQQBHDUPHG

This substitution is an example of one of the earliest known cyphers, known as the Caesar cypher or Caesar substitution. There are 26 different forms of Caesar substitutions. For example, we could have made the following substitution instead

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓	⇓
P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O

Our message would then appear as

PETCCNHPKTSXHPETTCCNTPGCTS