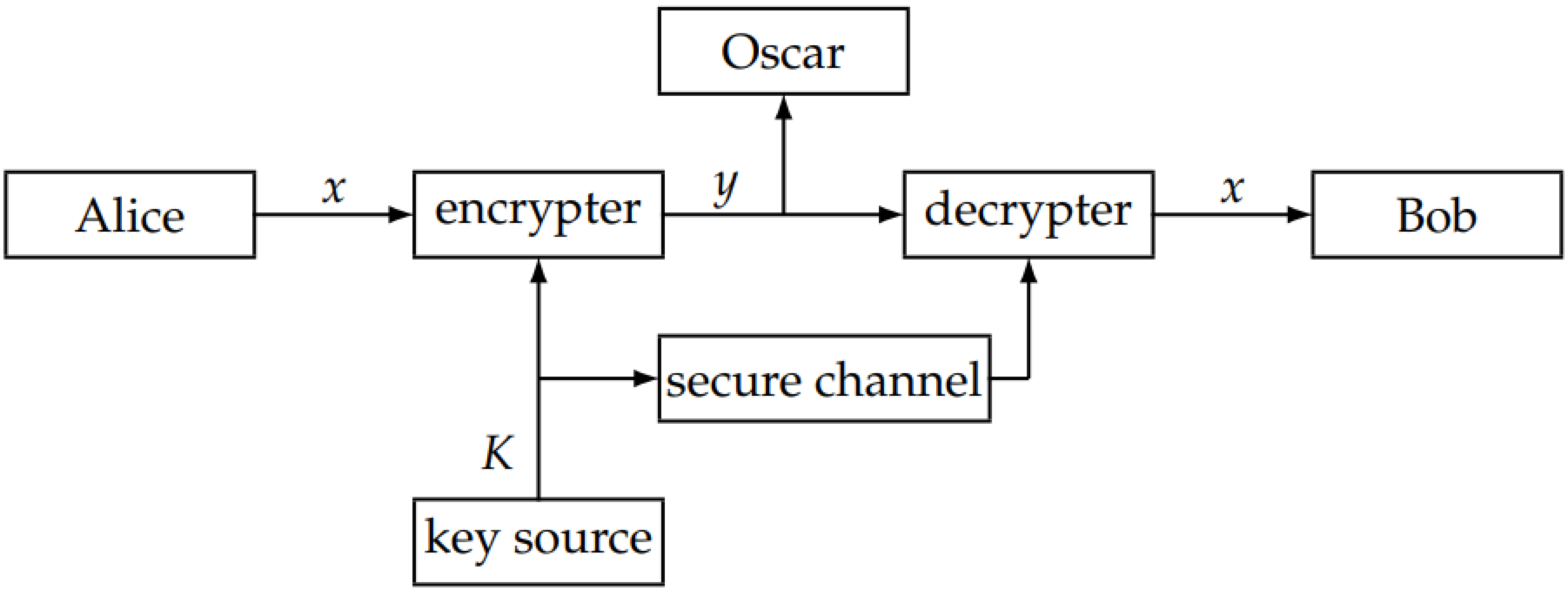
Cryptography

Lecture 4

Vahid Amin-Ghafari

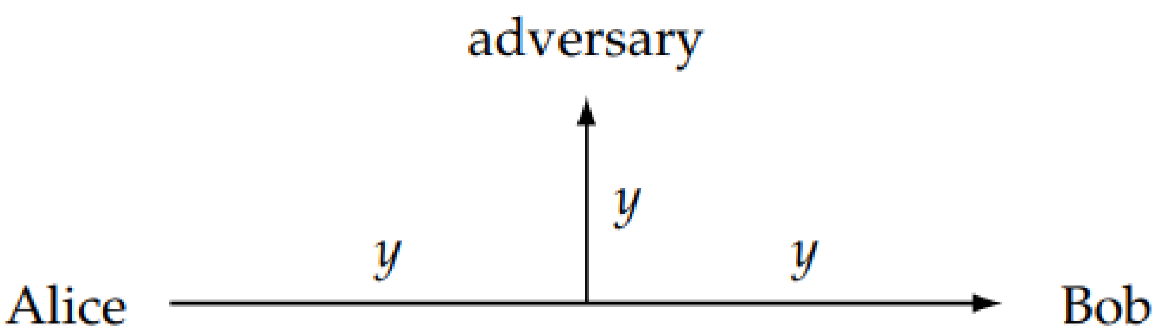
Vahidaming@cumt.edu.cn

- The fundamental objective of cryptography is to enable two people, usually referred to as Alice and Bob, to communicate over an insecure channel in such a way that an opponent, Oscar, cannot understand what is being said


Oscar
Alice
x
encrypter
y
decrypter
x
Bob
secure channel
K
key source

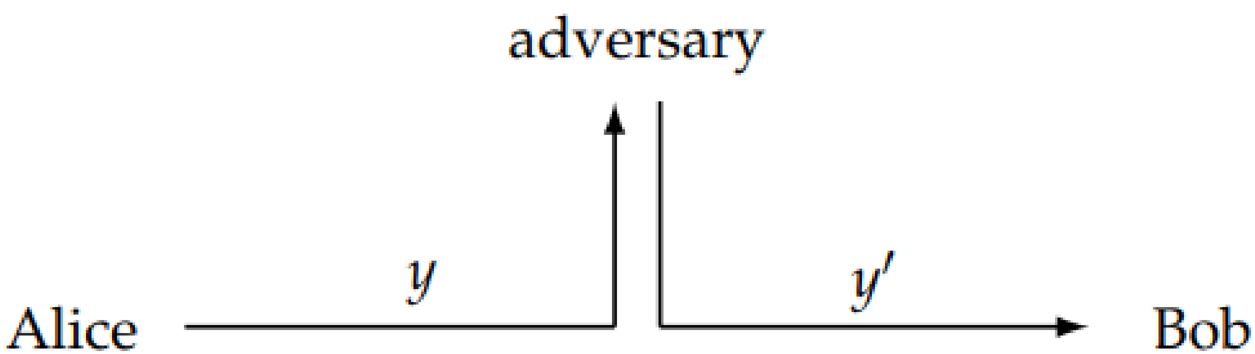
A passive adversary

- Cryptosystems provide secrecy (equivalently, confidentiality) against an eavesdropping adversary, which is often called a passive adversary.
- A passive adversary is assumed to be able to access whatever information is being sent from Alice to Bob;

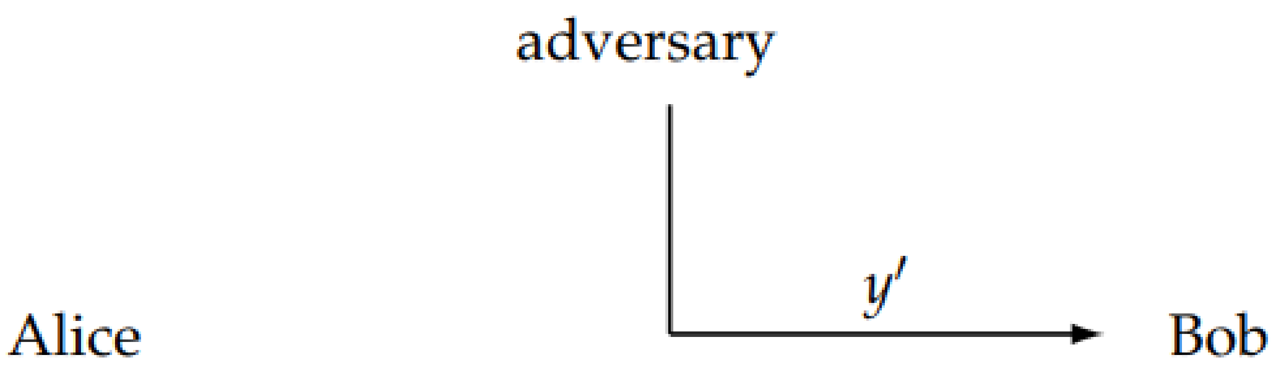


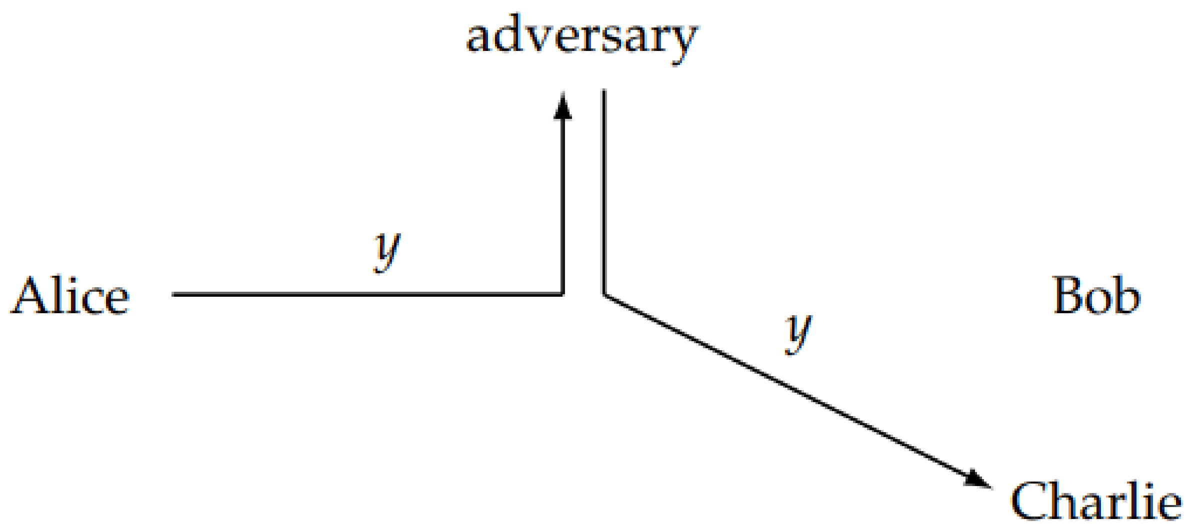
An active adversaries

- An active adversary is one who can alter information that is transmitted from Alice to Bob.



An active adversaries





Review

- Active and passive attacks
- Brute force (exhaustive search) attack
- Cryptanalysis (different types of attacks based on available data for attack)
- Private key (symmetric-key) cryptography
- Public key (asymmetric) cryptography

Cryptanalysis

- The general assumption that is usually made is that the opponent, Oscar, knows the cryptosystem being used. This is usually referred to as Kerckhoffs’ Principle.

- The attack model specifies the information available to the adversary when he mounts his attack.

Cryptanalysis

- **Ciphertext-only attack:**

The opponent observes a string of ciphertext.

- **Known plaintext attack:**

The opponent observes a string of plaintext, and the corresponding ciphertext.

- **Chosen plaintext attack:**

The opponent has obtained temporary access to the encryption machinery. Hence he can choose a plaintext string, and produce the corresponding ciphertext string.

- **Chosen ciphertext attack**

The opponent has obtained temporary access to the decryption machinery. Hence he can choose a ciphertext string, and produce the corresponding plaintext string, x.

A chosen-plaintext attack

- In May 1942, US Navy cryptanalysts intercepted an encrypted message from the Japanese that they were able to partially decode. The result indicated that the Japanese were planning an attack on AF, where AF was a ciphertext fragment that the US was unable to decode.
- The US believed that Midway Island was the target. Unfortunately, their attempts to convince planners in Washington that this was the case were fruitless .
- The general belief was that Midway could not possibly be the target. The Navy cryptanalysts devised the following plan: They instructed US forces at Midway to send a fake message that their freshwater supplies were low. The Japanese intercepted this message and immediately sent an encrypted message to their superiors that “AF is low on water”

A chosen-plaintext attack

- The Navy cryptanalysts now had their proof that AF corresponded to Midway, and the US dispatched three aircraft carriers to that location. The result was that Midway was saved, and the Japanese incurred significant losses.
- This battle was a turning point in the war between the US and Japan in the Pacific.
- The Navy cryptanalysts here carried out a chosen-plaintext attack, as they were able to influence the Japanese to encrypt the word "Midway"
- If the Japanese encryption scheme had been secure against chosen-plaintext attacks, this strategy by the US cryptanalysts would not have worked

Shannon’s Theory

- In 1949, Claude Shannon published a paper entitled Communication Theory of Secrecy Systems in the Bell Systems Technical Journal
- A great influence on the scientific study of cryptography

computational security

- This measure concerns the computational effort required to break a cryptosystem. We might define a cryptosystem to be computationally secure if the best algorithm for breaking it requires at least N operations, where N is some specified, very large number.
- The problem is that no known practical cryptosystem can be proved to be secure under this definition. In practice, people often study the computational security of a cryptosystem with respect to certain specific types of attacks (e.g., an exhaustive key search).
- Of course, security against one specific type of attack does not guarantee security against some other type of attack

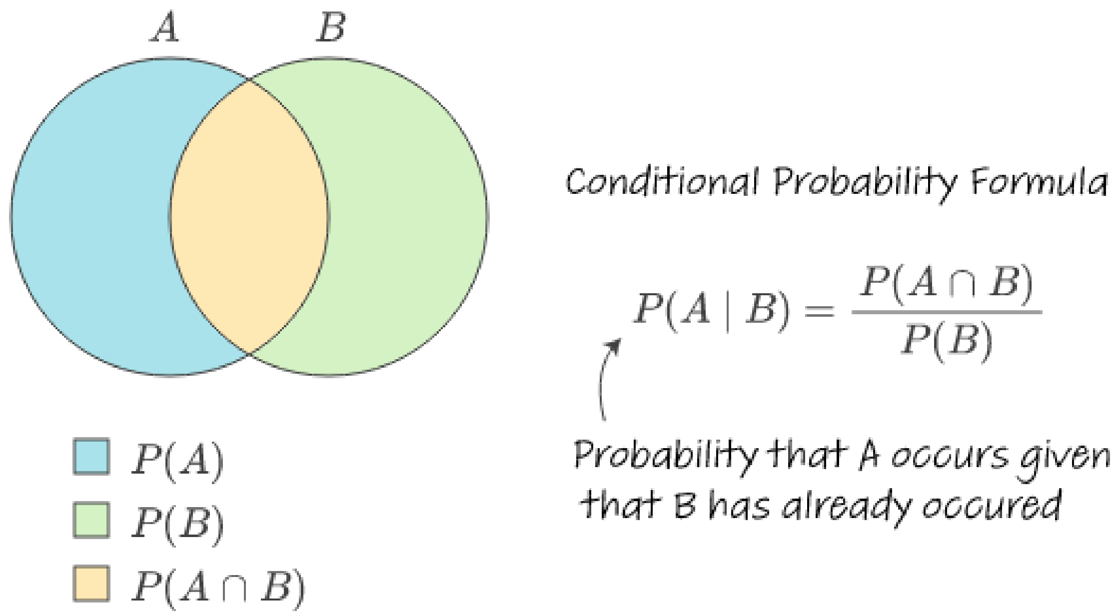
Provable security

- Another approach is to provide evidence of security by means of a reduction.
- In other words, we show that if the cryptosystem can be “broken” in some specific way, then it would be possible to efficiently solve some well studied problem that is thought to be difficult. For example, it may be possible to prove a statement of the type “a given cryptosystem is secure if a given integer n cannot be factored.”
- Cryptosystems of this type are sometimes termed provably secure, but it must be understood that this approach only provides a proof of security relative to some other problem, not an absolute proof of security

Unconditional security

- This measure concerns the security of cryptosystems when there is no bound placed on the amount of computation that Oscar is allowed to do.
- A cryptosystem is defined to be unconditionally secure if it cannot be broken, even with infinite computational resources.

Perfect Secrecy



- we assu) is specifie for only one enc
- A crypto] = Pr[x] for all x robability that the ext y is observed, is identical to the a priori probability that the plaintext is x.
- Perfect secrecy means that Oscar can obtain no information about the plaintext by observing the ciphertext. (intersection (∩) symbol, Joint events refer to events that occur simultaneously or together)

Defining secure encryption

Crypto definitions (generally)

- Security guarantee/goal
  - What we want to achieve (or what we want to prevent the attacker from achieving)
- Threat model
  - What (real-world) capabilities the attacker is assumed to have

encryption scheme

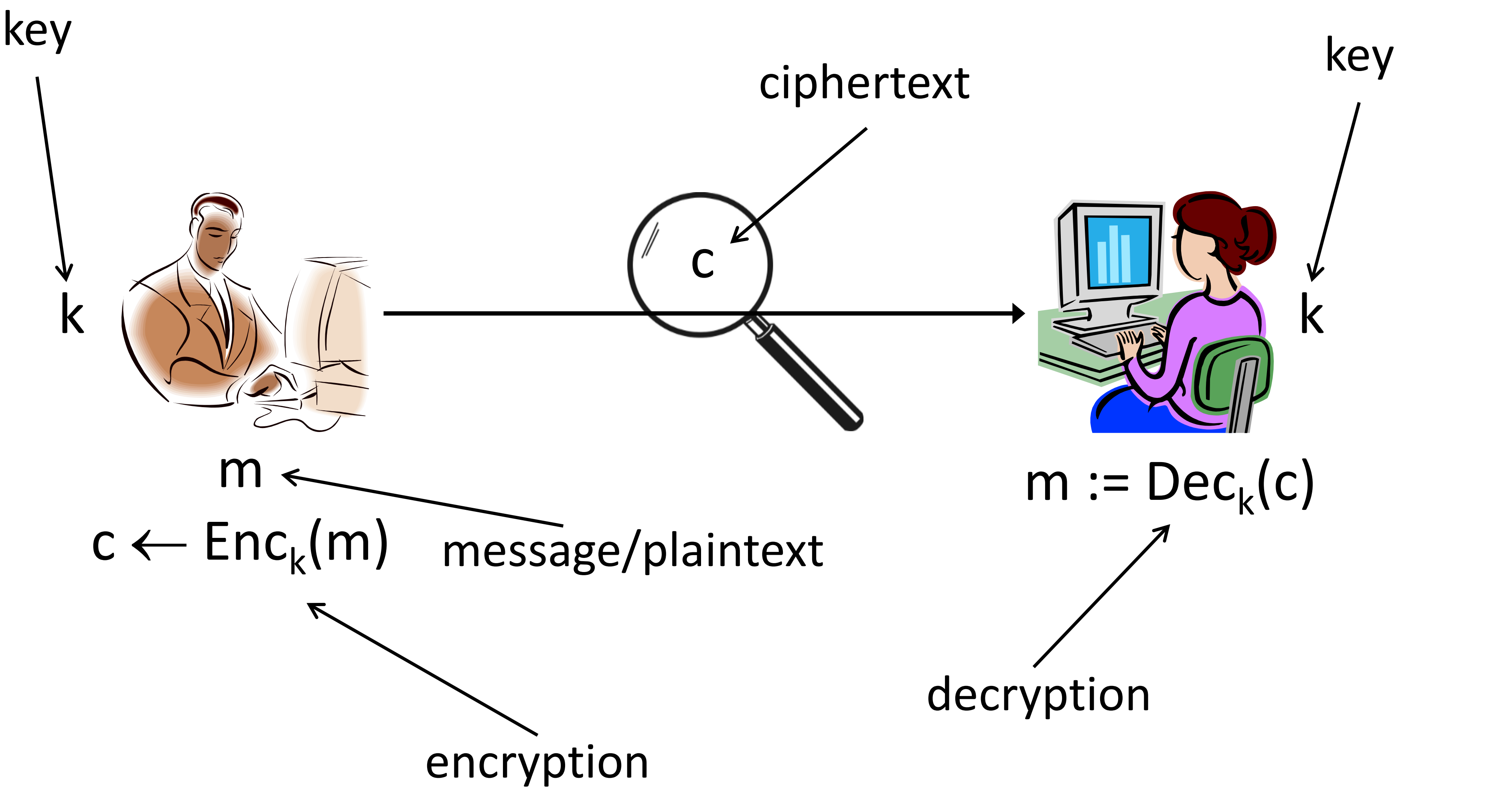
- A *private-key encryption scheme* is defined by a message space $\mathcal{M}$ and algorithms (Gen, Enc, Dec):
  - Gen (key-generation algorithm): generates k
  - Enc (encryption algorithm): takes key k and message $m \in \mathcal{M}$ as input; outputs ciphertext c

    $$c \leftarrow \mathrm{Enc}_k(m)$$

  - Dec (decryption algorithm): takes key k and ciphertext c as input; outputs m.

    $$m := \mathrm{Dec}_k(c)$$

Private-key encryption



Threat models for encryption

- Ciphertext-only attack
  - One ciphertext or many?
- Known-plaintext attack
- Chosen-plaintext attack
- Chosen-ciphertext attack

Goal of secure encryption?

- How would you define what it means for encryption scheme (Gen, Enc, Dec) over message space $\mathcal{M}$ to be secure?
  - Against a (single) ciphertext-only attack

Secure encryption?

- “Impossible for the attacker to learn the key”
  - The key is a *means to an end*, not the end itself
  - Necessary (to some extent) but not sufficient
  - Easy to design an encryption scheme that hides the key completely, but is insecure
  - Can design schemes where most of the key is leaked, but the scheme is still secure?

Secure encryption?

- "Impossible for the attacker to learn any character of the plaintext from the ciphertext"
  - What if the attacker is able to learn (other) partial information about the plaintext?
    - E.g., salary is greater than $75K
  - What if the attacker guesses a character correctly, or happens to know it?

The right definition

- “Regardless of any *prior* information the attacker has about the plaintext, the ciphertext should leak no *additional* information about the plaintext”
  - How to formalize?

Perfect secrecy

Notation

- $\mathcal{K}$ (key space) – set of all possible keys
- $\mathcal{C}$ (ciphertext space) – set of all possible ciphertexts

Probability distributions

- Let M be the random variable denoting the value of the message
  - M ranges over $\mathcal{M}$
  - Context dependent!
  - Reflects the likelihood of different messages being sent, given the attacker's prior knowledge
  - E.g.,

$$\Pr[M = \text{“attack today”}] = 0.7$$
$$\Pr[M = \text{“don't attack”}] = 0.3$$

Probability distributions

- Let K be a random variable denoting the key
  - K ranges over $\mathcal{K}$
- Fix some encryption scheme (Gen, Enc, Dec)
  - Gen defines a probability distribution for K:

    $$\Pr[K = k] = \Pr[\text{Gen outputs key } k]$$

  - Generally the uniform distribution

Probability distributions

- Assume random variables M and K are *independent*
  - I.e., parties don't pick the key based on the message, or the message based on the key
- In general, this assumption holds
- If it doesn't hold, can cause problems

Probability distributions

- Fix some encryption scheme (Gen, Enc, Dec), and some distribution for M
- Consider the following (randomized) experiment:
  1. Generate a key k using Gen
  2. Choose a message m, according to the given distribution
  3. Compute $c \leftarrow Enc_k(m)$
- This defines a distribution on the ciphertext!
- Let C be a random variable denoting the value of the ciphertext in this experiment

Example 1

- Consider the shift cipher
  - So for all $k \in \{0, \ldots, 25\}$, $\Pr[K = k] = 1/26$

- Say $\Pr[M = \text{'a'}] = 0.7$, $\Pr[M = \text{'z'}] = 0.3$

- What is $\Pr[C = \text{'b'}]$?
  - Either M = 'a' and $K = 1$, or M = 'z' and $K = 2$
  - $\Pr[C=\text{'b'}] = \Pr[M=\text{'a'}]\cdot\Pr[K=1] + \Pr[M=\text{'z'}]\cdot\Pr[K=2]$
    $= 0.7 \cdot (1/26) + 0.3 \cdot (1/26)$
    $= 1/26$

A	B	C	D	E	F	G	H	I	J
1	2	3	4	5	6	7	8	9	10
K	L	M	N	O	P	Q	R	S	T
11	12	13	14	15	16	17	18	19	20
U	V	W	X	Y	Z				
21	22	23	24	25	26				

- Consid … ribution on M g …

  Pr[M = ‘one’] = ½, Pr[M = ‘ten’] = ½

- Pr[C = ‘rqh’] = ?

  = Pr[C = ‘rqh’ | M = ‘one’] · Pr[M = ‘one’]
  + Pr[C = ‘rqh’ | M = ‘ten’] · Pr[M = ‘ten’]

  = 1/26 · ½ + 0 · ½ = 1/52

Perfect secrecy (informal)

- "Regardless of any *prior* information the attacker has about the plaintext, the ciphertext should leak no *additional* information about the plaintext"

Perfectly Secret Encryption, page 27 of textbook: "Introduction to Modern Cryptography, 3rd edition," Katz and Lindell

Perfect secrecy (informal)

- Attacker's information about the plaintext = attacker knows the *distribution* of M

- Perfect secrecy: observing the ciphertext should not change the attacker's knowledge about the distribution of M

Perfect secrecy (formal)

- Encryption scheme (Gen, Enc, Dec) with message space $\mathcal{M}$ and ciphertext space $\mathcal{C}$ is *perfectly secret* if for every distribution over $\mathcal{M}$, every m $\in$ $\mathcal{M}$, and every c $\in$ $\mathcal{C}$ with Pr[C=c] > 0, it holds that

$$\Pr[M = m \mid C = c] = \Pr[M = m].$$

- I.e., the distribution of M does not change, even conditioned on observing the ciphertext

Example 3

- Consider the shift cipher, and the distribution $\Pr[M = \text{'one'}] = ½$, $\Pr[M = \text{'ten'}] = ½$
- Take m = 'ten' and c = 'rqh'

- $\Pr[M = \text{'ten'} \mid C = \text{'rqh'}] = ?$

  $= 0$

  $\neq \Pr[M = \text{'ten'}]$

Bayes's theorem

- Pr[A | B] = Pr[B | A] · Pr[A]/Pr[B]

Bayes' theorem may be derived from the definition of conditional probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \text{ if } P(B) \neq 0,$$

where $P(A \cap B)$ is the probability of both A and B being true. Similarly,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}, \text{ if } P(A) \neq 0,$$

Example 4

- Shift cipher;
  Pr[M='hi'] = 0.3,
  Pr[M='no'] = 0.2,
  Pr[M='in']= 0.5

- Pr[M = 'hi' | C = 'xy'] = ?
  = Pr[C = 'xy' | M = 'hi'] · Pr[M = 'hi']/Pr[C = 'xy']

Example 4, continued

- Pr[C = ‘xy’ | M = ‘hi’] = 1/26

- Pr[C = ‘xy’]

  = Pr[C = ‘xy’ | M = ‘hi’] · 0.3 + Pr[C = ‘xy’ | M = ‘no’] · 0.2
    + Pr[C=‘xy’ | M=‘in’] · 0.5
  = (1/26) · 0.3 + (1/26) · 0.2 + 0 · 0.5
  = 1/52

Example 4, continued

- Pr[M = ‘hi’ | C = ‘xy’] = ?

  = Pr[C = ‘xy’ | M = ‘hi’] · Pr[M = ‘hi’]/Pr[C = ‘xy’]

  = (1/26) · 0.3/(1/52)

  = 0.6

  ≠ Pr[M = ‘hi’]

Conclusion

- The shift cipher is not perfectly secret!
  - At least not for 2-character messages

- How to construct a perfectly secret scheme?

One-time pad

- Patented in 1917 by Vernam
  - Recent historical research indicates it was invented (at least) 35 years earlier

- Proven perfectly secret by Shannon (1949)