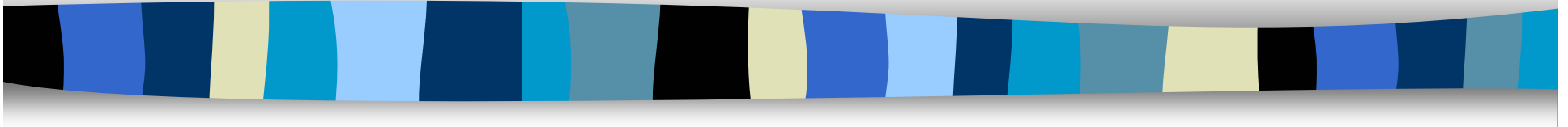


Cryptography CS 555

Lecture 3



Department of Computer Sciences
Purdue University

Symmetric Ciphers (Review)

- A Cipher ($K, P, C, \mathbf{K}, \mathbf{E}, \mathbf{D}$)
 - K : the key space
 - P : the plaintext space
 - C : the ciphertext space
 - \mathbf{K} : the key generation function
 - $\mathbf{E}: K \times P \rightarrow C$: the encryption function
 - $\mathbf{D}: K \times C \rightarrow P$: the decryption function
 - s.t. $\forall k \in K \forall M \in P \mathbf{D}_k[\mathbf{E}_k[M]] = M$
 - $\forall k \in K \mathbf{E}_k[] : P \rightarrow C$ must be a one-to-one function

Cryptanalysis (Review)

- Goals:
 - recover encryption key
 - decrypt one message
- Adversarial models:
 - ciphertext only,
 - known plaintext,
 - (adaptive) chosen plaintext,
 - (adaptive) chosen ciphertext

Classical Ciphers (review)

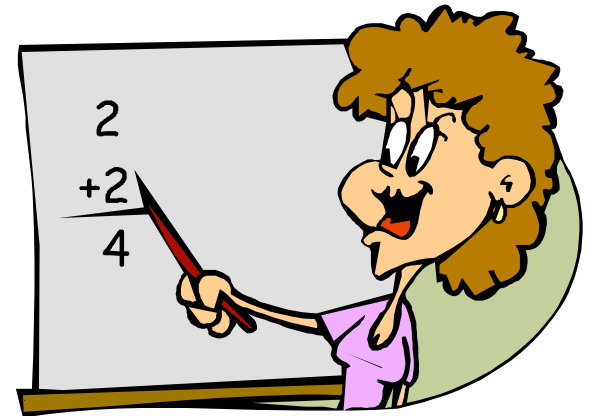
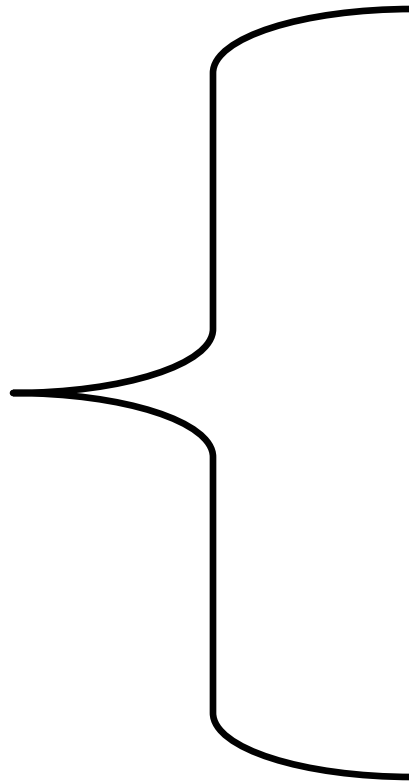
- Shift cipher
- Substitution cipher
- Vigenere cipher
 - Kasiski test
 - Index of coincidence
- Rotor machines
- Shift cipher
- One-time pad

Lecture Outline

- Elements of probability theory
- Perfect secrecy
- One-time pad
- Stream ciphers
 - RC4
 - LFSR

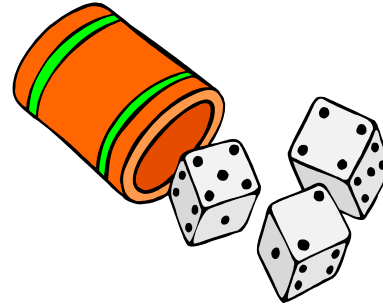


Begin Math



Elements of Probability Theory

A random experiment has an unpredictable outcome.



Definition

The **sample space (S)** of a random phenomenon is the set of all outcomes for a given experiment.

Definition

The event (E) is a subset of a sample space, an event is any collection of outcomes.

Probability Distribution

Basic Axioms of Probability:

If E is an event, $P(E)$ is the probability that event E occurs then

(a) $0 \leq P(A) \leq 1$ for any set A .

(b) $P(S) = 1$, where S is the sample space.

(c) If E_1, E_2, \dots, E_n is a sequence of mutually exclusive events, that is $E_i \cap E_j = \emptyset$, for all $i \neq j$ we have:

$$P(E_1 \cup E_2 \cup \dots \cup E_n) = \sum_{i=1}^n P(E_i)$$

Probability

More properties

If E is an event and $P(E)$ is the probability that the event E occurs then

- $P(\hat{E}) = 1 - P(E)$ where \hat{E} is the complimentary event of E
- If outcomes in S are equally like, then $P(E) = |E| / |S|$ (where $| |$ denotes the cardinality of the set)

Example

Random throw of a pair of dice.

What is the probability that the sum is 4?

Solution: Each dice can take six different values $\{1,2,3,4,5,6\}$. The number of possible events (value of the pair of dice) is 36, therefore each event occurs with probability $1/36$.

Examine the sum: $4 = 1+3 = 2+2 = 3+1$

The probability that the sum is 4 is $3/36$.

What is the probability that the sum is 11?

Random Variable

Definition

A **random variable** is a function that assigns a number to each outcome of a random experiment.

Elements of Probability Theory

Definitions

Assume X and Y are two random variables, then we define:

- **joint probability**: $P[x, y]$ is the probability that X takes value x and Y takes value y .
- **conditional probability**: $P[x|y]$ is the probability that X takes on the value x given that Y takes value y .
- **independent random variables**: X and Y are said to be independent if $P[x,y]=P[x]P[y]$, for all $x \in X$ and all $y \in Y$.

Elements of Probability Theory

Bayes' Theorem

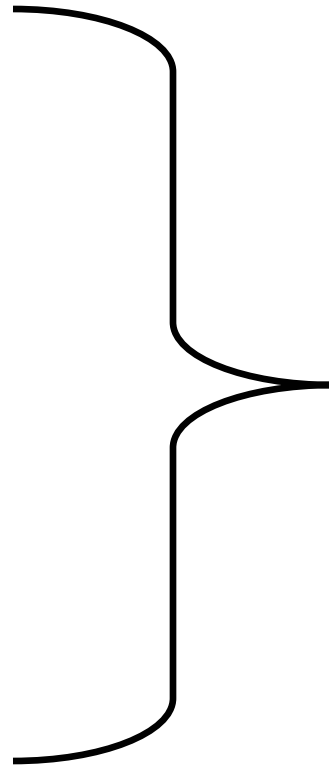
If $P[y] > 0$ then

$$P[x | y] = \frac{P[x]P[y | x]}{P[y]}$$

Corollary

X and Y are independent random variables
iff $P[x|y] = P[x]$, for all $x \in X$ and all $y \in Y$.

End Math



Shannon (Information-Theoretic) Security

- Basic Idea: CT should provide no “information” about PT
- Definition: Encryption scheme is *Shannon secure* if for *every distribution* of PT
 - $\forall C \forall M \Pr [PT=M] = \Pr [PT=M \mid CT=C]$
- We also say such a scheme has perfect secrecy.

One-time Pad has Perfect Secrecy

- One time pad: $P = C = K = \{0,1\}^n$
 - \mathbf{K} outputs a uniformly random k , $\mathbf{E}_k[M] = k \oplus M$

Proof: $\forall M_0 \forall C_0 \Pr [PT=M_0 \mid CT=C_0]$

$$= \Pr[PT=M_0, CT=C_0] / \Pr[CT=C_0]$$

$$= \Pr[PT=M_0] \Pr [CT=C_0 \mid PT=M_0]$$

$$/ \sum_{M \in \mathcal{M}} (\Pr[PT=M] \Pr[CT=C \mid PT=M])$$

$$= \Pr[PT=M_0] 1/2^n / \sum_{M \in \mathcal{M}} (\Pr[PT=M] 1/2^n)$$

$$= \Pr[PT=M_0] / \sum_{M \in \mathcal{M}} (\Pr[PT=M])$$

$$= \Pr[PT=M_0]$$

The “Bad News” Theorem

- Thm: Let $\mathcal{P} = \mathcal{C} = \{0,1\}^n$, Any cipher with perfect secrecy has $|\mathcal{K}| \geq 2^n$
 - Proof: Given a CT C , there must exist one key for each PT.
- perfect secrecy \Rightarrow
key-length \geq msg-length

Stream Ciphers

- In OTP, a key is described by a random bit string of length n
- Stream ciphers:
 - Idea: replace “rand” by “pseudo rand”
 - Use Pseudo Random Number Generator
 - PRNG: $\{0,1\}^s \rightarrow \{0,1\}^n$
 - expand a short (e.g., 128) random seed into a long (e.g., 10^6) bit string that “looks random”
 - Secret key is the seed
 - $E_{\text{seed}}[M] = M \oplus \text{PRNG}(\text{seed})$

Properties of PRNG

- No more perfect secrecy
 - security depends on PRNG
- PRNG must be “unpredictable”
 - Given consecutive sequence of bits output (but not seed), next bit must be hard to predict
 - Don't use UNIX rand for crypto!
 - Kerberos V4
- Never reuse any part of an output stream
- Typical stream ciphers are very fast: (10 * DES)
- Used everywhere:
 - SSL (RC4), Cell phones, DVD (LFSR)

RC4

- A proprietary cipher owned by RSA, designed by Ron Rivest in 1987.
- Became public in 1994.
- Simple and effective design.
- Variable key size, byte-oriented stream cipher.
- Widely used (web SSL/TLS, wireless WEP).

The RC4 Cipher: Encryption

- The cipher state consists of
 - an 256-byte array, which contains a permutation of 0 to 255
 - total number of possible states is $256!$, very big number
 - two indexes: i, j

$i = j = 0$

Loop

$i = (i + 1) \pmod{256}$

$j = (j + S[i]) \pmod{256}$

swap($S[i], S[j]$)

$t = (S[i] + S[j]) \pmod{256}$

End Loop

RC4 Initialization

- How to generate the initial permutation from a key k
- Divide k into L bytes

```
for i = 0 to 255 do
    S[i] = i
j = 0
for i = 0 to 255 do
    j = (j + S[i] + k[i mod L]) (mod 256)
    swap (S[i], S[j])
```

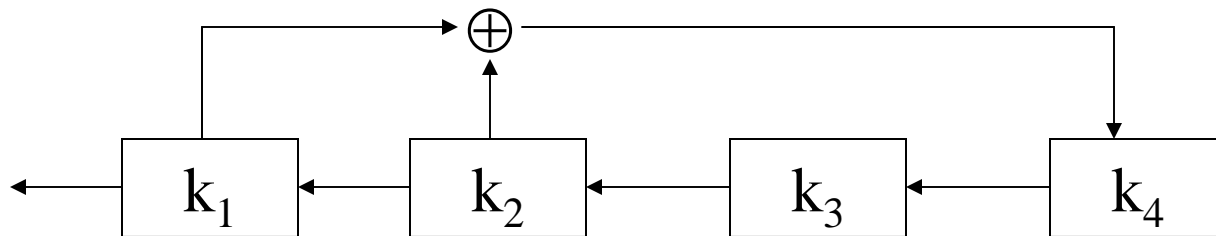
RC4 Cryptanalysis

- Two weaknesses:
 - Problem with init
 - the first byte generated by RC4 leaks information about individual key bytes.
 - best to drop first 256 bytes of output
 - Statistical attack
 - prob. of seeing (0,0) is $1/256^2 + 1/256^3$
 - after seeing $(256^3)^2 = 2^{48} = 10^7$ GB can distinguish RC4 from truly random stream



Linear Feedback Shift Register

- Example:



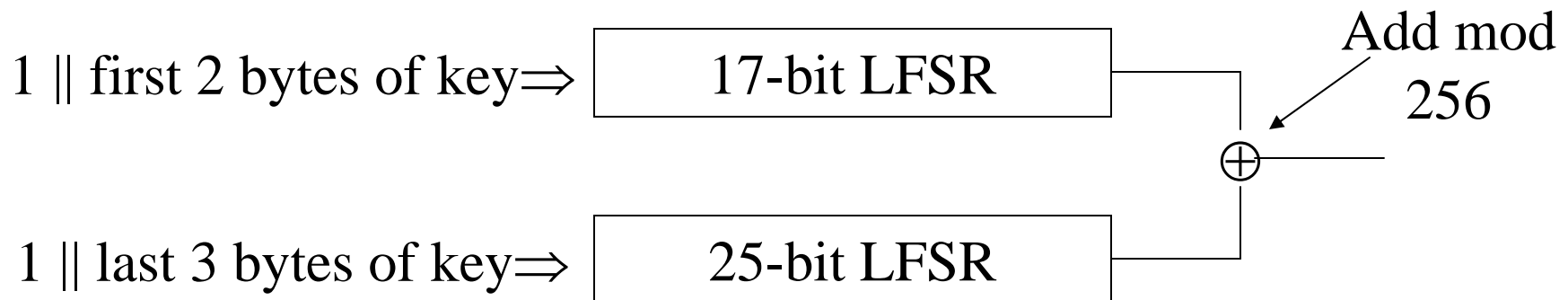
- Starting with 1000, the output stream is
 - 1000 1001 1010 1111 000
- Repeat every $2^4 - 1$
- The seed is the key

Cryptanalysis of the LFSR Stream Cipher

- Vulnerable to know-plaintext attack
 - A LFSR can be described as
$$z_{m+i} = \sum_{j=0}^{m-1} c_j z_{i+j} \text{ mod } 2$$
 - Knowing $2m$ output bits, one can
 - construct m linear equations with m unknown variables c_0, \dots, c_{m-1}
 - recover c_0, \dots, c_{m-1}

CSS Stream Cipher

- Key = 5 bytes = 40 bits
 - brute-force attack is possible
 - more efficient attacks exist



Given short output sequence, can recover the seed in
time 2^{20}

Attacks for the OTP

- Two time pad
 - Given $C_1 = M_1 \oplus k$ and $C_2 = M_2 \oplus k$
 - $C_1 \oplus C_2 = M_1 \oplus M_2$
 - Never reuse pad in one-time pad
 - Same applies to stream cipher
 - RC4 key should never be reused
- Proper Use of RC4:
 - $E_k[M] = 3DES_k[\text{Seed}] \parallel M \oplus \text{RC4}(\text{Seed})$

Exploiting Keystream Reuse

- Poor implementation of RC4 and keystream reuse, allowed WEP to be broken
- Methods to obtain pairs (plaintext, ciphertext):
 - IP fields predictable: login sequences, recognize shared libraries transfer
 - Send email and wait for the user to check it via wireless links
 - Send data to access-points that have access control disabled and observe the encrypted data

OTP & Stream Ciphers are Highly Malleable

- Easy to change ciphertext so that plaintext changes in predictable
 - easy to flip bits

Next Lecture

- Block Cipher & DES
- Recommended reading for next lecture:
Stallings Chapter 3
Stinson Chapter 3
BR Chapter 2

