Cryptography
Lecture 2

Lecture by Dr. Alon Rosen
Typeset by Steven Karas

2019-03-12
Last edited 18:14:25 2019-03-12

Disclaimer These notes are based on the lectures for the course Cryptography, taught by Dr. Alon Rosen at IDC Herzliyah in the spring semester of 2018/2019. Sections may be based on the lecture slides prepared by Dr. Alon Rosen.

1 Recap

1.1 Factorization Complexity

As a clarification, the average case complexity of integer factorization is trivial. However, we care about the average case behavior for a subset of integers in the form $N = PQ$ where both P and Q are prime and around the same size.

The current state of the art works in time $2^{1.92n^{1/3}\log n^{2/3}}$.

2 Agenda

- Probability theory in a nutshell
- Perfect secrecy

3 Probability Theory

Denote a probability space as a countable set S and a function $\Pr : S \to [0,1] \in \mathbb{R}$ such that $\sum_{x \in S} \Pr[x] = 1$

Examples:

- Alice flips 100 fair coins. The probability space is $A = \{0,1\}^{100}$ with $\Pr = 0.5^{100}$
- Bob flips 100 fair coins. The probability space is the same as above
- Carol picks Alice's coin 75% of the time, and Bob's the rest.
- Eve gets $E = A \oplus B$.

$$\Pr[(x, y, z)]$$

3.1 Identities

Complement

$$\Pr[\bar{T}] = 1 - \Pr[T]$$

Union

$$\Pr[T_1 \vee T_2] \underbrace{=}_{T_1 \cap T_2 = \emptyset} \Pr[T_1] + \Pr[T_2]$$

Union Bound

$$\Pr[T_1 \vee T_2] \leq \Pr[T_1] + \Pr[T_2]$$

This is sometimes useful to provide a very rough bound.

Total Probability

$$\Pr[T] \underbrace{=}_{S_i \text{ is pairwise disjoint}} \sum_i \Pr[T \wedge S_i]$$

3.2 Independence

x, y are independent iff:

$$\forall x, y \Pr[X = x \wedge Y = y] = \Pr[X = x] \cdot \Pr[Y = y]$$

XOR Bitwise xor has many useful properties worthy of mention.

From our example before with Alice, Bob, and Eve:

$$E = A \oplus B \Leftrightarrow E \oplus B = A$$
$$E = (E \oplus B) \oplus B$$

which gives us a probability function for (a, b, c):

$$\Pr[A = a \wedge B = b \wedge E = e] \neq \left(\frac{1}{2^{100}}\right)^3$$

3.3 Expectation

$$\mathrm{E}[X] = \sum_x \Pr[X = x] \cdot x$$

Expectation is linear:

$$\mathrm{E}[X + Y] = \mathrm{E}[X] + \mathrm{E}[Y]$$
$$\mathrm{E}[cX] = c\,\mathrm{E}[X]$$

However, the following only holds if X, Y are independent, but not generally:

$$\mathrm{E}[X \cdot Y] = \mathrm{E}[X] \cdot \mathrm{E}[Y]$$

Example: flipping 100 coins Denote the number of heads flipped by Alice as Z_A defined as follows:

$$Z_A^i = \begin{cases} 1 & a_i = 0 \\ 0 & a_i = 1 \end{cases}$$

The expectation:

$$\mathrm{E}[Z_A^i] = \underbrace{\Pr[Z_A^i = 0] \cdot 1}_{\frac{1}{2}} + \underbrace{\Pr[Z_A^i = 1] \cdot 0}_{0} = \frac{1}{2}$$

$$\begin{aligned} \mathrm{E}[Z_A] &= \mathrm{E}\left[\sum_{i=1}^{100} Z_A^i\right] \\ &= \sum_{i=1}^{100} \mathrm{E}[Z_A^i] \\ &= \sum_{i=1}^{100} \frac{1}{2} \\ &= 50 \end{aligned}$$

Consider the expectation of the square:

$$\begin{aligned} \mathrm{E}[(Z_A)^2] &= \mathrm{E}\left[\left(\sum_{i=1}^{100} Z_A^i\right)^2\right] \\ &= \mathrm{E}\left[\sum_{i=1}^{100} \left(Z_A^i\right)^2 + \sum_{i \neq j} Z_A^i Z_A^j\right] \\ &= \sum_{i=1}^{100} \mathrm{E}\left[\left(Z_A^i\right)^2\right] + \sum_{i \neq j} \mathrm{E}\left[Z_A^i Z_A^j\right] \\ &= \sum_{i=1}^{100} \mathrm{E}\left[Z_A^i\right] + \sum_{i \neq j} \mathrm{E}\left[Z_A^i Z_A^j\right] \\ &= \dots \end{aligned}$$

3.4 Bounds

3.4.1 Markov Bound

If X is a non-negative random variable.

$$\Pr[X \geq t] \leq \frac{\mathrm{E}[X]}{t}$$

If the expectation is small, this gives us a good bound.

3.4.2 Chernoff Bound

Let $X_1, \dots, X_n$ be independent $\{0, 1\}$ valued random variables with $\Pr[X_i = 1] = \mu \;\; \forall i$

$$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} x_i > \mu + \varepsilon\right] \leq e^{-2\varepsilon^2 n}$$
$$\Pr\left[\frac{1}{n}\sum_{i=1}^{n} x_i > \mu - \varepsilon\right] \leq e^{-2\varepsilon^2 n}$$

Example:

$$\begin{aligned}
\Pr[Z_A \geq 70] &= \Pr\left[\frac{1}{100}\sum_{i=1}^{100} Z_A^i \geq \frac{70}{100}\right] \\
&= \Pr\left[\frac{1}{100}\sum_{i=1}^{100} Z_A^i \geq \underbrace{\frac{50}{100}}_{\mu} + \underbrace{\frac{20}{100}}_{\varepsilon}\right] \\
&< e^{-2\cdot\left(\frac{1}{5}\right)^2\cdot 100} \approx 0.00033
\end{aligned}$$

3.4.3 Chebyshev Bound

Not covered, but lies in between Markov and Chernoff.

3.5 Conditional Probability

Let E, F be events. Consider the probability of E occurring given that F occurs:

$$\Pr[E|F] = \frac{\Pr[E \wedge F]}{\Pr[E]}$$

Example: Consider:

$$\begin{aligned}
\Pr[Z_c \text{ is even}|Z_A \text{ is even}] &= \dots \\
&= \frac{3}{4} + \frac{1}{4}\cdot\frac{1}{2} = \frac{7}{8}
\end{aligned}$$

3.5.1 Bayes' Law

$$\Pr[E|F] = \frac{\Pr[E \wedge F]}{\Pr[E]} = \frac{\Pr[F|E]\Pr[E]}{\Pr[F]}$$

Example: Consider:

$$\begin{aligned}
\Pr[Z_A \text{ is even}|Z_c \text{ is even}] &= \dots \\
&= \frac{3}{4} + \frac{1}{4}\cdot\frac{1}{2} = \frac{7}{8}
\end{aligned}$$

4 Private-key Cryptography

This section is largely the result of Shannon's masters thesis[3] and his later paper[4].

We will present private key encryption defined as two actors Alice and Bob who have an agreed upon key k. Alice encrypts a message m using key k denoted as $E_k(m)$ and sends it over an insecure channel to Bob. Bob then decrypts the message using key k denoted as $D_k(E_k(m)) = m$. The threat model we will consider is a simple eavesdropper Eve who sees all messages over the channel.

4.1 Kerckhoff principle

This is the underlying principle of cryptography through WW2: all security must rely on the secrecy of the key, and nothing else. Specifically, assume that the workings of your system will be discovered by an adversary.

4.2 Syntax

A private key cryptosystem consists of a plaintext space P, a ciphertext space C, and a keyspace K with three algorithms G, E, D:

1. Key generation G is a randomized algorithm that gives a private key $k \in K$, $\;k \leftarrow^R G$.
2. Encryption E_k using key k gives ciphertext c: $c = E_k(m)$. This algorithm may be randomized.
3. Decryption D_k using key k gives the plaintext m: $D_k(E_k(m))$ under the assumption that the key used for decryption is the same as that used for encryption.

4.3 History: Shift cipher (Caesarean cipher)

The oldest known cipher. Works by shifting letters in alphabetic order. For example, a becomes b, b becomes c, etc. The key is the number of times to shift the alphabet.

Formally:

$$k \leftarrow^R \{0, \ldots, 25\}$$
$$P = C = \{A, \ldots, Z\}^l \approx \{0, \ldots, 25\}^l$$
$$E_k(m_1, \ldots, m_l) = c_1, \ldots, c_l$$
$$c_i = m_i + k \mod 26$$
$$D_k(c) : m_i = c_i - k \mod 26$$

This system is trivially insecure because the keyspace is too small. However, if the message length is 1, then it is secure.

4.4 Large key size principle

k should be large enough to avoid exhaustive key search.

4.5 History: Substitution cipher

The key is a random permutation of $\{0, \ldots, 25\}$. This gives us a keyspace with $26! \approx 2^8 8$ possible keys.

$$E_k(m_1, \ldots, m_l) = k(m_1), \ldots, k(m_l)$$
$$D_k(c_1, \ldots, c_l) = k^{-1}(c_1), \ldots, k^{-1}(c_l) = m_1, \ldots, m_l$$

This is also insecure due to frequency analysis.

4.6 History: One-time pad (Vernam's cipher)

Let the key be $k \leftarrow^R \{0,1\}^l$. The text space is $m \in \{0,1\}$.

$$c_i = m_i \oplus k_i$$
$$E_k(m) = m \oplus k$$
$$D_k(c) = c \oplus k$$

This system is perfectly secure, yet it leaks the size of the messages.

4.7 Perfect Security

All of the following are insufficient conditions

1. An adversary cannot learn the key from the ciphertext. Insufficient because the ciphertext may be the plaintext.
2. An adversary cannot learn the plaintext from the ciphertext. Insufficient because we need to secure portions of the plaintext. For example, just the content of an encrypted email, without the headers.
3. An adversary cannot learn any symbol of the plaintext. This is insufficient because it still allows the adversary to extract contextual information.
4. An adversary cannot learn any information about the plaintext.

To resolve this, Shannon defined perfect indistinguishability:

4.7.1 Perfect indistinguishability

A cryptosystem G, E, D satisfies perfect indistinguishability[1] if $\forall m_1, m_2 \in P$ and $k \leftarrow^R G$ the random variables $E_k(m_1), E_k(m_2)$ have the same distribution. That is, $\forall c \in C$:

$$\Pr[E_k(m_1) = c] = \Pr[E_k(m_2) = c]$$

4.8 Proof: Insecurity of Shift, Substitution ciphers

For example, consider the two plaintexts: "GANZ" v "BIBI". Because "BIBI" has repeated characters, the ciphertext will have identical characters in the first and third positions, as well as the second and fourth.

4.9 Proof: Security of one-time pads

For a fixed $m \in \{0,1\}^l$ and $c \in \{0,1\}^l$:

$$\Pr_k[E_k(m) = c] = \Pr_k[m \oplus k = c] = \frac{1}{2^l}$$

5 Next week

Mor Weiss, a postdoc at IDC will give the lecture next week. We will cover Shannon security and that the key length must be larger than the message length.

[1]Take note that this still allows some information about the message to be leaked, such as the presence of a message and its length.

References

[1] Thomas H. Cormen, Clifford Stein, Ronald L. Rivest, and Charles E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.

[2] Oded Goldreich. *Foundations of Cryptography: Volume 1, Basic Tools*, volume 1. Cambridge University Press, 2000.

[3] C. E. Shannon. A symbolic analysis of relay and switching circuits. *Transactions of the American Institute of Electrical Engineers*, 57(12):713–723, Dec 1938.

[4] C. E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):656–715, 1949.