CS 7880 Graduate Cryptography September 10, 2017

Lecture 1: Perfect Secrecy and Statistical Authentication

Lecturer: Daniel Wichs *Scribe: Matthew Dippel (Edited)*

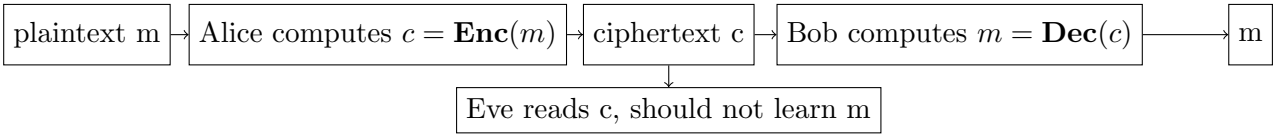
1 Topic Covered

- Definition of perfect secrecy
- One-time pad and optimality (Shannon's theorem)
- Definition of statistical one-time MAC and construction.

2 The Encryption Problem

Consider two persons Alice and Bob, who wish to communicate messages of a sensitive nature. However, an eavesdropper by the name of Eve has the ability to read all the messages that pass between Alice and Bob. Despite this, Alice and Bob would still like to be able to communicate in a way that Eve cannot determine what they are saying to each other.

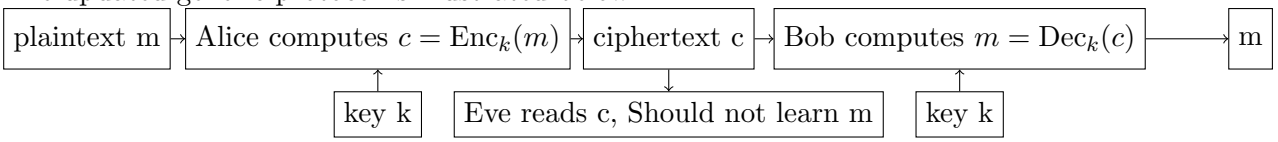
This is done via an encryption scheme, which is a set of functions **Enc** and **Dec** for encryption and decryption of messages:



There are two key properties of the above. First, given only c, Eve should not be able to learn what m is. Second, $\text{Dec}(\text{Enc}(m)) = m$, so that Bob can accurately determine the message that Alice wanted to send. Creating functions **Enc** and **Dec** which satisfy this is a fundamental problem of cryptography. But given the above encryption scheme, knowing what these functions are is enough for anyone to decrypt c. This would make it impossible for practical cryptography to be analyzed in the public, for any scheme which is openly discussed would become useless.

The above problem has a simple solution: first, we assume that the scheme is always public knowledge. This is necessary in order to create cryptography systems which can be used as well as analyzed. Second, our **Enc** and **Dec** functions will be parametrized by a secret key value **k**, which changes the behavior of the functions. By including such a parameter which is chosen at random and shared between the communicating parties, it is possible to both use a cryptography scheme which is public knowledge, as well as analyze the cryptography scheme and prove statements about its security.

The updated generic protocol is illustrated below:



2.1 Formal View

An encryption scheme as we have described so far can be formally described as a tuple of the following objects:

set $\mathcal{K}$	the key space
set $\mathcal{M}$	the message space
set $\mathcal{C}$	the ciphertext space
$\text{Enc} : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$	the encryption function.
$\text{Dec} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$	the decryption function.

We will write $\text{Enc}(k, m)$ as $\text{Enc}_k(m)$ for short, and similarly write $\text{Dec}(k, c)$ as $\text{Dec}_k(c)$. . These functions should satisfy the sanity check that decryption reverses the encryption. We denote this as $\forall k, m$, $\text{Dec}_k(\text{Enc}_k(m)) = m$

To define perfect secrecy, we treat the encryption function as a pseudo random function parametrized by the key choice k. Given any cipher text c, it should be impossible to distinguish with non-trivial probability which message could have produced it, as for any possible message, there exists a key in the key space such that $\text{Enc}_k(m) = c$.

We formalize this as follows: let M be some distribution over the message space $\mathcal{M}$, and K the uniform distribution over the key space $\mathcal{K}$. Since $c = \text{Enc}_k(m)$, then C is also a random variable whose distribution depends on both M and the encryption and decryption functions. Then we have perfect security if, $\forall$ distributions M, $m \in \mathcal{M}$, $c \in \mathcal{C}$:

$$\Pr[M = m] = \Pr[M = m | C = c]$$

Intuitively, this means that conditioned on the value of $\text{Enc}_k(m)$, m is no more or less likely to be any specific message from the message distribution. There are two more equivalent definitions of perfect secrecy. We will list all three below:

1. $\Pr[M = m] = \Pr[M = m | C = c]$
2. M and C are independent random variables.
3. $\forall m, m', c$, $\Pr[\text{Enc}_K(m) = c] = \Pr[\text{Enc}_K(m') = c]$

Note that, whenever we state a probability, it is over the distributions of variables that have not been set.

Theorem 1 *Definitions 1, 2, and 3 are equivalent*

We will show that $1 \to 2$, $2 \to 3$, and $3 \to 1$.

Lemma 1 *Definition 1 implies Definition 2*

Proof:

$$\Pr[M = m] = \Pr[M = m | C = c] = \frac{\Pr[M = m, C = c]}{\Pr[C = c]}$$

$$\Pr[M = m] \Pr[C = c] = \Pr[M = m, C = c]$$

which is the definition of independence. □

Lemma 2 *Definition 2 implies Definition 3*

Proof: Choose any m from the distribution M and c from C. We have:

$$\Pr[\mathrm{Enc}_K(m) = c] = \Pr[\mathrm{Enc}_K(M) = c | M = m]$$
$$= \Pr[C = c | M = m] = \Pr[C = c]$$

where the conditioning on M can be removed because M and C are independent. Since the choice of m was arbitrary, we can substitute any m', deriving $\Pr[\mathrm{Enc}_K(m) = c] = \Pr[\mathrm{Enc}_K(m') = c]$. □

Lemma 3 *Definition 3 implies Definition 1*

Proof: Consider any specific ciphertext c in the distribution C. We can rewrite the probability $\Pr[C = c]$ as:

$$\Pr[C = c] = \sum_{m'} \Pr\left[C = c, M = m'\right]$$
$$= \sum_{m'} \Pr\left[C = c | M = m'\right] \Pr\left[M = m'\right]$$

We can rewrite the random variable C in terms of M as:

$$= \sum_{m'} \Pr\left[\mathrm{Enc}_K(M) = c | M = m'\right] \Pr\left[M = m'\right]$$

In this, we can replace the random variable M with m' and remove the conditioning:

$$= \sum_{m'} \Pr\left[\mathrm{Enc}_K(m') = c\right] \Pr\left[M = m'\right]$$

From 3, we have that $\Pr[\mathrm{Enc}_K(m') = c]$ takes the same value for any m'. Thus we can pick a specific value m and remove it from the sum:

$$= \Pr[\mathrm{Enc}_K(m) = c] \sum_{m'} \Pr\left[M = m'\right] = \Pr[\mathrm{Enc}_K(m) = c]$$

Where we used the fact that $\sum_{m'} \Pr[M = m'] = 1$. We can now substitute the specific value of m with the random variable M and condition on its value:

$$= \Pr[\text{Enc}_K(M) = c | M = m] = \Pr[C = c | M = m]$$

We have thus shown that $\Pr[C = c] = \Pr[C = c | M = m]$ for any m, c. Using the definition of conditional probability and some rearrangement, this is equivalent to $\Pr[M = m] = \Pr[M = m | C = c]$, as desired. □

3 One Time Pad (OTP)

Having described perfectly secrecy, we now describe an encryption scheme which satisfies it, called the One-Time-Pad, or OTP for short. It is named One Time Pad because it has perfect secrecy for a single message, but loses nearly all security properties if the same key is used twice.

We introduce an additional parameter t, as the number of bits in the keys, messages, and cipher texts. In the OTP, we have our sets and encryption / decryption functions as::

$$\mathcal{K} = \mathcal{M} = \mathcal{C} = \{0,1\}^t$$

$$\text{Enc}_k(m) = k \oplus m$$

$$\text{Dec}_k(c) = k \oplus c$$

where XOR operations are done on each corresponding bit in the vectors.

Note that XOR satisfies the following properties:

$$a \oplus b = b \oplus a$$

$$a \oplus a = 0$$

$$a \oplus 0 = a$$

From these, we can show that $\text{Dec}_k(\text{Enc}_k(m)) = m$.

As it turns out, the set $\{0,1\}^t$ with the operation $\oplus$ is a finite abelian group. Thus we will instead generalize the OTP to abelian groups, and prove that this framework has perfect secrecy.

Let $(G, +)$ be a finite abelian group. Then we define our encryption scheme as:

$$\mathcal{K} = \mathcal{M} = \mathcal{C} = G$$

$$\text{Enc}_k(m) = m + k$$

$$\text{Dec}_k(c) = c - k$$

Theorem 2 *The one-time pad over a finite group $(G, +)$ satisfies perfect secrecy.*

Proof: For any $m, m', c \in G$, we have:

$$\Pr[\text{Enc}_K(m) = c]$$

$$= \Pr[K + m = c]$$

$$= \Pr[K = c - m]$$

Since $c - m$ is a constant value for chosen m and c, we are really considering the probabiltiy that a randomly chosen key is equal to their distance. Since all keys are equally likely and chosen from the key space G, this probabiltiy is equal to:

$$= \frac{1}{|G|}$$

Since the choice of m was arbitrary, we have that $\Pr[\text{Enc}_K(m) = c] = \Pr[\text{Enc}_K(m') = c]$, which is one of our definitions of perfect secrecy. □

Despite achieving perfect secrecy, the One Time Pad is an impractical scheme with many undesirable properties.:

1. The key is as long as the message
2. The key cannot be reused
3. Alice and Bob must share a secret key unknown to Eve.

As it turns out, these are all necessary requirements for any perfectly secret encryption scheme. In particular, we can prove that the key space must be at least as large as the message space for perfect secrecy to be achieved:

Theorem 3 (Shannon 1949) *In any perfect secrecy scheme, we must have that* $|\mathcal{K}| \geq |\mathcal{M}|$.

Proof: Let M be the uniform distribution over $\mathcal{M}$ and c be some ciphertext such that $\Pr[C = c] > 0$.

Consider the set $\mathcal{M}' = \{\text{Dec}_k(c) : k \in \mathcal{K}\}$, which has cardinality at most $|\mathcal{K}|$. If $|\mathcal{K}| < |\mathcal{M}|$, then there exists $m \in \mathcal{M}/\mathcal{M}'$.

We have then that:

$$\Pr[M = m | C = c] = 0$$

, since no key decrypts c to m. Yet we have that $\Pr[M = m] = 1/|\mathcal{M}|$. Thus the requirements of perfect secrecy are violated. □

The above proof is essentially a formalization of a brute force attack. If we tried all possible encryption keys, we could rule out some messages from the message space. However, it is computationally expensive, and does not seem to gain that much information for the attacker. In the future, we will introduce new ways of modeling secrecy that involve a computationally bounded adversary.

4 Authentication and MACs

Our current model of secrecy has assumed an adversary which passively listens to the communication channel. Suppose our adversary was less passive, and interested in actively modifying the message from Alice before it reached Bob. Then the One Time Pad would provide no authenticity, as there is no way for Bob to guess whether or not the message has been tampered with. To address this, we will introduce Message Authentication Codes, or MAC for short.

A MAC scheme can be described by the following tuple:

4.1 Formal View

$\mathcal{K}$	the key space
$\mathcal{M}$	the message space
$\mathcal{T}$	the tag space
$\text{MAC} : \mathcal{K} \times \mathcal{M} \to \mathcal{T}$	the MAC function

The usage is as follows. Alice wishes to send a message m to Bob. She computes $t = \text{MAC}(k, m)$ and sends the pair (m, t) to Bob. Bob receives (m', t'), and checks if $t' = \text{MAC}(k, m')$. If it does, he proceeds under the assumption that m' is the intended message from Alice. If it does not, he concludes that the message $m \neq m'$.

We must now define what propeties we want in a MAC. To do this, we define a "game" where Eve must attempt to forge a message and MAC tag.

Consider the following game: Given a MAC scheme, a uniformly random key k from $\mathcal{K}$ is chosen and kept secret from Eve. Eve is allowed to choose exactly one message $m \in \mathcal{M}$, and is given the value of $t = \text{MAC}(k, m)$. Eve must now come up with a pair (m', t') such that $m \neq m'$. She wins if $t' = \text{MAC}(k, m')$, and loses otherwise.

We say that a MAC is 1-Time Statistically Secure with ε-security if, no matter what Eve's strategy is, the probability that Eve wins the above game is $\leq \varepsilon$.

The first question to ask is whether or not we can achieve $\varepsilon = 0$. The answer is no... given any message, there is a tag that results from running the MAC function on it. If Eve picks a random message and a random tag, there is a non-zero probability that she will pick the correct tag. Thus, we are unable to break the lower bound of $\varepsilon \geq 1/|\mathcal{T}|$.

We will provide a construction which achieves reasonable security. The domain that we will work in will be the integers modulo some prime p. That is, we will have $\mathcal{M} = \mathcal{T} = \mathbb{Z}_p$. Our key space will be all tuples of such integers, $\mathcal{K} = \mathbb{Z}_p \times \mathbb{Z}_p$

Then, we define our MAC functions as:

$$\text{MAC}(k, m) = x * m + y$$

where $k = (x, y)$, and multiplication and addition are done over $\mathbb{Z}_p$.

Theorem 4 *The above MAC has 1-time security with $\varepsilon = 1/p$.*

Proof: Let $K = (X, Y)$ be uniformly random. Then $\forall m, t$ we have:

$$\Pr\left[\mathrm{MAC}(K, m) = t\right] = \Pr\left[X * m + Y = t\right] = 1/p$$

which follows from the system $X * m + Y = t$ having a unique solution Y for every choice of X.

Next, consider some $m' \neq m$, and any t, t'. Consider the probability:

$$\begin{aligned}
&\Pr\left[\mathrm{MAC}(K, m') = t', \mathrm{MAC}(K, m) = t\right] \\
&= \Pr\left[Xm' + Y = t', Xm + Y = t\right] \\
&= \Pr\left[X = x, Y = y\right] = 1/p^2
\end{aligned}$$

where $x = \frac{t-t'}{m-m'}$ and $y = t - xm$. Again, all operations are done over the field $\mathbb{Z}_p$.

Thus from the values we just determined and properties of conditional probability, we have that

$$\Pr\left[\mathrm{MAC}(K, m') = t' | \mathrm{MAC}(K, m) = t\right] = 1/p.$$

Thus we achieve $1/p$-security. □

Like our OTP scheme, this MAC scheme is also impractical for several reasons. First, the key is twice as big as the message. Thus a relatively large amount of secret information needs to be exchanged betwen parties compared to the messages they will exchange. Second, we can only use the key once to authenticate a single message. If the key is used for two different messages, then the key can be exactly determined. We will see later on that we can achieve significant improvements on the key size, but the single message security is inherent to our definition.

APPENDIX

Below is a short collection of some of the assumed knowledge for this course / lecture.

Probability Theory

In Probability Theory, we have a universe $\mathcal{U}$, and a probability function $\mathbf{Pr} : \mathcal{U} \to [0, 1]$. This function satisfies the property that $\sum_{uin\mathcal{U}} \Pr[u] = 1$. A specific probability function, along with a universe, is sometimes denoted as a **distribution** over that universe. Informally, they can be viewed as a set of possible disjoint events, and the relative likelihood of each one occuring.

DEFINITION 1 A probability distribution is denoted as the **uniform distribution** if, for all $u, u' \in \mathcal{U}$, it holds that $\Pr[u] = \Pr[u'] = 1/|\mathcal{U}|$. ♢

Random variables are functions over $\mathcal{U}$, which also induce distributions. Given a random variable X which maps $\mathcal{U}$ into $\mathcal{X}$, we can define a distribution over $\mathcal{X}$ as $\Pr[X = x] = \sum_{u:X(u)=x} \Pr[u]$.

We can also induce distributions over multiple random variables at the same time by introducing an additional random variable that is a tuple of the variables we are interested in. Suppose we had random two variables X, Y, and wanted to express the probability that both $X = x$ and $Y = y$ simultaneously. Then we would introduce a new random variable $Z = (X, Y)$, and examine the distribution:

$$\Pr[X = x, Y = y] = \Pr[Z = (x, y)] = \sum_{u:Z(u)=(x,y)} \Pr[u]$$

We say that two random variables X, Y are independent if $\Pr[X = x, Y = y] = \Pr[X = x] \Pr[Y = y]$ for all choices of x, y.

Example 1 *Consider the uniform distribution over $\{0, 1\}^2$. Let X be the first bit, Y be the second bit, $Z = X + Y$, and $W = X \oplus Y$. Then X and Y are independent. X and Z are not independent. X and W are independent.*

Conditional Probability For two random variables X, Y and outcomes x, y we define the conditional probability:

$$\Pr[X = x | Y = y] = \frac{\Pr[X = x, Y = y]}{\Pr[Y = y]}$$

The LHS is read as "probability X=x, given that Y=y".

Example 2 *Use the same universe and random variables from the last example. Consider the conditional probability $\Pr[X = 1 | Z = 1]$. Then we can calculate this as:*

$$= \frac{\Pr[X = 1, Z = 1]}{\Pr[Z = 1]}$$
$$= \frac{0.25}{0.5} = \frac{1}{2}$$

Events

An event E is a subset of $\mathcal{U}$. We define $\Pr[E] = \sum_{u \in E} \Pr[u]$. Alternatively, we can think of E is a binary random variable, where $E(u) = 1$ if $u \in E$, and 0 otherwise.

Theorem 5 (Union Bound) *For any events E_1, E_2, we have:*

$$\Pr[E_1 \cup E_2] = \Pr[E_1] + \Pr[E_2] - \Pr[E_1 \cap E_2]$$

$$\leq \Pr[E_1] + \Pr[E_2]$$