CMSC 28400 Introduction to Cryptography **Autumn 2021**

Notes #4: Perfect Secrecy

Instructor: David Cash

These notes cover our first definition of security: *Perfect Secrecy*, which dates back to the work Shannon in the 1950s. The motivation for this approach to security is to (hopefully) escape the attack-then-patch cycle that drove classical cipher development. Abstractly, the plan is to give a mathematical definition of "security" for a cipher, every bit as precise as other definitions, and then *prove* that a cipher meets this definition of security.

What guarantees can this plan give us, ultimately? The proof can only tell us that attacks violating the security definition are impossible. As we will see several times in this course, real attacks may work without violating the definition; They instead violate some underlying assumption about how the definition corresponds to reality. (An extreme example could be an attacker who steals your laptop and gets your key. Clearly all bets are off, even if you had a nice proof.)

Below we start with the security definition, and then prove the One-Time-Pad achieves this definition. In the next set of notes we'll discuss limitations of this definition, and generalize it to something more practical.

4.1 How to Evaluate the Security of Ciphers?

First let's recall the definition of a cipher from the previous notes, for use below.

Definition 4.1. *Let $\mathcal{K}, \mathcal{M}, \mathcal{C}$ be non-empty sets. A function*

$$E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$$

is called a cipher with key-space $\mathcal{K}$, message-space $\mathcal{M}$, and ciphertext-space $\mathcal{C}$ *if for every $K \in \mathcal{K}$, the function $E(K, \cdot)$ is one-to-one.*

For such a cipher, we define

$$E^{-1} : \mathcal{K} \times \mathcal{C} \to \mathcal{M}$$

by letting $E^{-1}(K, \cdot)$ be the inverse of $E(K, \cdot)$ for each $K \in \mathcal{K}$. (More precisely, $E^{-1}(k, c)$ is only defined when there exists $m \in \mathcal{M}$ such that $E(k, m) = c$.)

Suppose E is a cipher. What does it mean for E to be "secure"? We could come up with list of attacks that it should resist. We could say that E should resist single-character frequency analysis, or bigram frequency analysis, or more and more clever attacks as we learn about them. Or, we could ask more abstractly that "no one should be able to recover an encrypted message", or even "no one should be able to recover any part of an encrypted message", though as a cipher designer it would be hard to interpret such a goal. Someone could convince us that we hadn't achieved those goals by presenting an attack, but how could someone possibly convince us that they *had* achieved such poorly-stated goals?

The approach to answering this question taken in the next subsection is to give a mathematical definition of security which compresses a long list of security properties into one simple requirement.

If one deems the definition suitable for an application, then the proof is enough to convince you that a cipher "is secure."

Before we dive in, it is worth noting that the approach of enumerating possible attacks is fundamental to modern cryptography. As we will see later, we can't give definitions and proofs for everything. But our philosophy will be to use definitions and proofs as a tool to help drive our cipher design and limit security risk.

4.1.1 The Definition of Perfect Secrecy

Let's just lay the definition out and then discuss it, first at a purely syntactic level, then at an intuitive level.

Definition 4.2. *Let $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ be a cipher and $\mathbf{K}$ be a uniform random variable on $\mathcal{K}$. We say that E is* perfectly secret *if for all $m_0, m_1 \in \mathcal{M}$, and $c \in \mathcal{C}$,*

$$\Pr[E(\mathbf{K}, m_0) = c] = \Pr[E(\mathbf{K}, m_1) = c].$$

In the probabilities, $\mathbf{K}$ is the only random variable. So for instance the left probability is determined by how often m_0 will encrypt to c with a randomly-chosen key; The other probability is the same except with m_1. But in both probabilities, m_0, m_1, and c are *fixed* and not random.

Intuitively, this definition says that which message you are encrypting should not affect the distribution of the ciphertext you get. This is great for security, because it means an attacker looking at the ciphertext gets "no information" about what the message was.

Exercise 4.1. *Show that the substitution cipher is not perfectly secret. (Hint: Considering a ciphertext like $c =$ `AA` makes this more clear, but any ciphertext will work.)*

We can try to understand this definition by giving an equivalent version. The following definition is somewhat more complicated technically, but may be more clear. It is equivalent to perfectly secrecy, in the sense that any cipher E that is perfectly secret meets this definition, and vice versa.

Definition 4.3. *Let $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ be a cipher and $\mathbf{K}$ be a uniform random variable on $\mathcal{K}$. We say that E has* independent ciphertexts *if for all random variables $\mathbf{M}$ on $\mathcal{M}$ independent of $\mathbf{K}$, we have that the random variable $\mathbf{C} = E(\mathbf{K}, \mathbf{M})$ is independent of $\mathbf{M}$.*

Here we use a random variable $\mathbf{M}$ to model picking a message in some application-dependent way (or to model the distribution of the message given an attacker's knowledge). For instance, if an attacker knows you sending a payment response to a website, it might know every character of your message except your credit card number. In that case $\mathbf{M}$ could be a fixed request but with random digits for the credit card number.

Note also in the definition that $\mathbf{C}$ is another random variable defined as a function of $\mathbf{K}$ and $\mathbf{M}$. The probability notes review the underlying foundations of what this means mathematically (though, as the notes mention, we are operating without explicitly saying what the sample space is, and will continue to do so in later definitions).

Theorem 1. *A cipher $E : \mathcal{K} \times \mathcal{M} \to \mathcal{C}$ is perfectly secret if and only if it has independent ciphertexts.*

Exercise 4.2. *Prove this theorem. (Hint: There are two directions. You will need to look up the precise definition of independent random variables if you are unsure how to proceed rigorously.)*

4.2 The One-Time Pad is Perfectly Secret

We now carry out the second step in our plan: Constructing a cipher and proving that it is perfectly secret. We'll use the last of our historical ciphers, the one-time pad. Concretely, for each integer $n \geq 1$, define a cipher $\mathrm{OTP}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}^n$ by

$$\mathrm{OTP}_n(k, m) = k \oplus m.$$

These ciphers are perfectly secret, as we now prove.

Theorem 2. *For every integer $n \geq 1$, the cipher OTP_n is perfectly secret.*

Proof. We use the original definition, so we have to show that for every $m_0, m_1 \in \{0,1\}^n$ and $c \in \{0,1\}^n$,

$$\Pr[E(\mathbf{K}, m_0) = c] = \Pr[E(\mathbf{K}, m_1) = c],$$

where $\mathbf{K}$ is uniform on $\{0,1\}^n$. We will do even better, and show that for every $m \in \{0,1\}^n$ and $c \in \{0,1\}^n$,

$$\Pr[E(\mathbf{K}, m) = c] = 2^{-n},$$

which will imply what we need (if the probabilities are all 2^{-n} no matter what m and c are, then certainly the probabilities above are equal). This follows because

$$\Pr[E(\mathbf{K}, m) = c] = \Pr[\mathbf{K} \oplus m = c] = \Pr[\mathbf{K} = c \oplus m] = 2^{-n}.$$

The first two equalities hold because $E(\mathbf{K}, m) = c$, $\mathbf{K} \oplus m = c$, and $\mathbf{K} = c \oplus m$ are all the *same event* (one happens if and only if the others happen; it looks like simple manipulation of the equations, but it is formally justified by them representing the same event). The last equality holds because $\mathbf{K}$ is uniform and $c \oplus m$ is just a fixed element of $\{0,1\}^n$. □

As short as it is, this proof might be tricky if you are not used to manipulating random variables. It is worth referring back to the probability background if you found the equalities in the proof difficult to verify.

This theorem explains why we needed to have multiple encryptions under the same key in order to break a one-time pad cipher: If we only had one ciphertext, then we'd effectively have only a random string, independent of the message. But once more than one ciphertext with the same key is available, attacks like crib-dragging are effective.

Exercise 4.3. *Define a version of the Vigenère cipher with an n-letter key working only on ciphertexts of n letters. Show that this cipher is perfectly secret.*

Exercise 4.4. *Define a cipher $E : \{0,1\} \times \{0,1,2\} \to \{0,1,2\}$ by $E(k, m) = k + m \bmod 3$. Show directly that E is not perfectly secret. (This will also be a consequence of a later theorem.)*