# CPE-579 Homework Assignment

Peter Rauscher

May 12, 2023

*"I pledge my honor that I have abided by the Stevens Honor System."*

# Contents

# 1 One-Time Pad

1. Consider the extension of the One-time Pad cipher, where a ciphertext $c$ is possibly longer than the plaintext $m$ (and the key $k$) by one bit, namely, by having the encryption algorithm **Enc** appending to $m \oplus k$, either a 0 with probability 0.65 or a 1 with probability 0.3. Is this extension a perfectly secret cipher and why?

**Solution:** Yes, the extension described above to the one-time pad cipher is still a perfectly secret cipher scheme, because the probability distribution is unaffected. In other words, since the last bit that is added to the ciphertext is not dependent on the plaintext message $m$ in any way, the ciphertext has an equal probability of being any possible value within the ciphertext space $C$.

By definition of the one-time pad cipher, the probability of generating a particular ciphertext $c$ is inversely proportional to its bitlength $t$:

$$\Pr[E_k(m) = c] = \frac{1}{2^t}$$

Now, let us consider the extension to the One-time Pad cipher. Recall the definition of conditional probability:

$$\Pr[A|B] = \frac{\Pr[A \cap B]}{\Pr[B]}$$

Let $m$ be any plaintext message of length $t$, let $k$ be any random key of length $t$, and let $c = m \oplus k$ be the corresponding ciphertext also of length $t$. Let $b$ be the value of the random bit that is appended to $c$.

$$\begin{aligned}
\Pr[E_k(m) = c \mid (b = \{0,1\})] &= \frac{\Pr[E_k(m) = c] \cdot (\Pr[b = 0]) \cup \Pr[b = 1]}{\Pr[b = 0]) \cup \Pr[b = 1]} \\
&= \frac{\frac{1}{2^t} \cdot (0.65 + 0.3)}{(0.65 + 0.3)} \\
&= \frac{\frac{1}{2^t} \cdot 0.95}{0.95} \\
&= \frac{1}{2^t}
\end{aligned}$$

Since the probability of obtaining the ciphertext $c$ with either value of the appended bit $b$ is the same as the probability of obtaining the ciphertext $c$ without any appended bit (the normal one-time pad cipher), we can conclude that the message space $M$ and the ciphertext space $C$ are independently random variables, and that the probability distribution $D_C$ does not depend on the plaintext. Therefore, the perfect secrecy is maintained in this extension.

2. Based on your answer in the previous question, prove or refute the following statement:

An encryption scheme with message space $M$ is perfectly secret if and only if for every probability distribution $D_M$ over $M$ and every $c_0, c_1 \in C$ we have

$$Pr[C = c_0] = Pr[C = c_1].$$

**Solution:** Consider the extension to the one-time pad cipher that we covered in the last problem. Let $c_0$ be any ciphertext that has had a 0 bit appended to it, and let $c_1$ be any ciphertext that has had a 1 bit appended to it. We know by definition that $c_0 \neq c_1$ and $c_0, c_1 \in C$.

The probability of appending a 0 to the ciphertext was given as 0.65 and the probability of appending a 1 to the ciphertext was given as 0.3. Thus,

$$Pr[C = c_0] = Pr[C = c_1]$$
$$0.65 \neq 0.3$$

However, we have proven in the solution to the last question that perfect secrecy is maintained in this extension. Therefore, the claim that "an encryption scheme is perfect if and only if for every probability distribution $D_M$ over $M$ and every $c_0, c_1 \in C$ we have $Pr[C = c_0] = Pr[C = c_1]$" does not hold as the probability distributions differ yet perfect secrecy is still maintained. Thus, the statement is proven false by contradiction.

3. Consider the following encryption scheme. The message space is

$$M = \{m \in \{0,1\}^l \mid \text{the last bit of } m \text{ is } 0\},$$

algorithm **Gen** chooses a uniform key from $\{0,1\}^{l-1}$, algorithm $\mathbf{Enc}_k(m)$ returns ciphertext $m \oplus (k||0)$, and algorithm $\mathbf{Dec}_k(c)$ returns $c \oplus (k||0)$. Is this a perfectly secret cipher and why?

**Solution:** To determine whether this encryption scheme is perfectly secret, we need to show that for any two plaintext messages $m_1, m_2 \in M$ of the same length, and any ciphertext $c$, the probability that the ciphertext is generated from $m_1$ and the probability that it is generated from $m_2$ is the same.

Let $m_1 = m_1'||0$ and $m_2 = m_2'||0$ be two plaintext messages of length $l$ (with $m_1', m_2' \in \{0,1\}^{l-1}$). Let $k$ be a uniformly chosen key from $\{0,1\}^{l-1}$, and let $c = m \oplus (k||0)$ be the corresponding ciphertext.

Then, we have:

$$\Pr[\mathbf{Enc}_k(m_1) = c] = \Pr[m_1 \oplus (k||0) = c] = \Pr[m_1' \oplus k = c']$$

where $c' = c_1'||0$ is obtained by removing the last bit of $c$.

Similarly, we have:

$$\Pr[\mathbf{Enc}_k(m_2) = c] = \Pr[m_2' \oplus k = c']$$

Since $k$ is chosen uniformly at random from $\{0,1\}^{l-1}$, both $\Pr[m_1' \oplus k = c']$ and $\Pr[m_2' \oplus k = c']$ are equal to $2^{-l+1}$.

Therefore, we have shown that for any two plaintext messages $m_1, m_2 \in M$ of the same length, and any ciphertext $c$, the probability that the ciphertext is generated from $m_1$ and the probability that it is generated from $m_2$ is the same. Essentially, although the key chosen is always one bit shorter than the length of the message, it does not result in any leak of meaningful data from the plaintext because the last bit is always 0, and the attacker knows this fact as well. The attacker cannot gain any information on the rest of the message from this bit. Hence, this encryption scheme is perfectly secret.

4. When the one-time pad encryption algorithm runs with the all-zero key $k = 0^l$, the resulted ciphertext $c$ equals the plaintext $m$. Consider the extension that purposely avoids this situation by disallowing **Gen** choosing the all-zero key - that is, $k$ is chosen uniformly at random from the set of nonzero binary strings of length $l$. Is this extension a perfectly secret cipher and why?

**Solution:** The extension of the one-time pad encryption scheme that removes the all-zero key is no longer perfectly secret. Although the attacker may inadvertently view the plaintext message when viewing the ciphertext in the case the all-zero key is chosen during **Gen** of the one-time pad, he will have no way of knowing that the ciphertext he is viewing is in fact equivalent to the plaintext message. By definition of perfect secrecy, the probability that the ciphertext is the plaintext is the same probability that it is any other text within the ciphertext space $C$. In other words,

$$\Pr[c = m] = \Pr[c = c'] \text{ for any } c' \in C \text{ within the classic one-time pad encryption scheme}$$

If anything, "extending" the one-time pad by removing the all-zero key from the keyspace only serves to *help* the attacker. After all, since he is aware that the key space excludes the all-zero key, he can be certain that any ciphertext he views is **NOT** the plaintext message, and thus the size of the message space he has to exhaust is slightly smaller. Consider an example using the proposed extension:

Let the length of the plaintext message $m$ be $t = 3$. The size of the message space $M$ is $2^t = 8$. The ciphertext space now excludes the true value of $m$, so the size of $C$ is $2^t - 1 = 7$. The attacker intercepts the ciphertext message $c = 010$, and wants to learn the plaintext. Knowing the encryption algorithm well, the attacker understands this **cannot** be the value of $m$, and tries - naively - to guess the value of $m$. The probability that the attacker guesses the correct message is now $\frac{1}{2^t-1} = \frac{1}{7}$. Had the all-zero key not been excluded, the probability he guesses the correct message would have been $\frac{1}{2^t} = \frac{1}{8}$. Because the attacker can remove the ciphertext he intercepts from the list of possible plaintext messages, his chances of breaking the encryption scheme by random chance go up significantly.

In probabilistic terms, this extension to the one-time pad encryption scheme cannot be perfectly secret because it the probability of obtaining any ciphertext $c \in C$ is not the same for all messages and ciphertexts.

## 2 EAV-Security

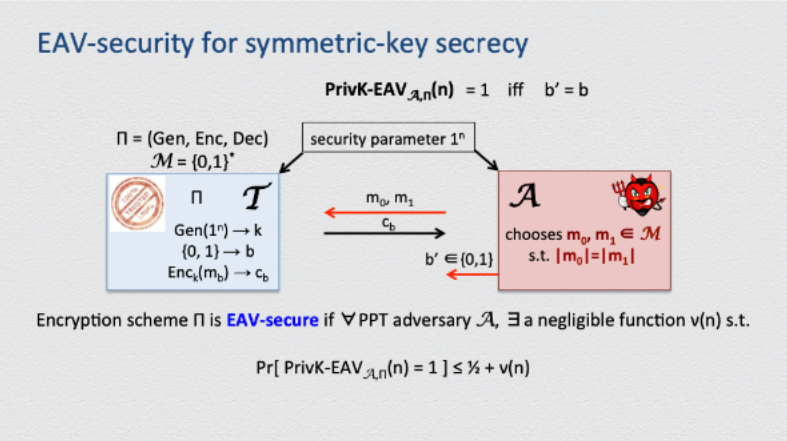


Figure 1: Indistinguishability of encryptions in the presence of an eavesdropper.

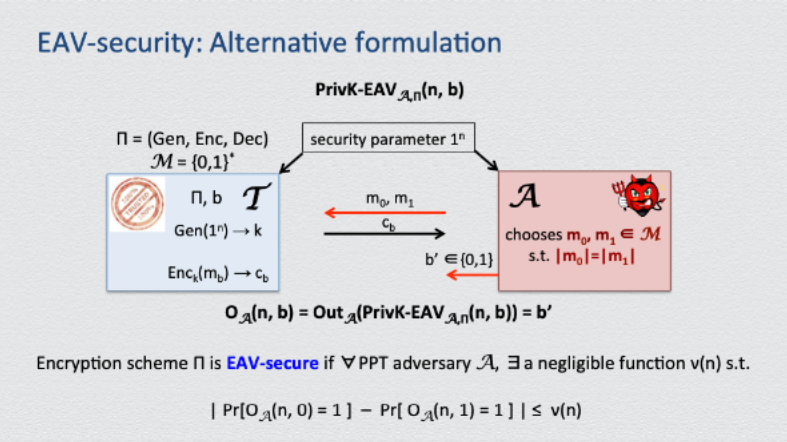


Figure 2: Alternative formulation of EAV-security.

Recall the basic game-based definition for secrecy, EAV-security, and its alternative formulation, depicted in Figures 1 and 2, respectively. In both cases, a PPT adversary $A$ selects two messages in $M$, one of which is encrypted to form the challenge ciphertext $c_b$, that must be determined by $A$.

1. Explain how the definition in Figure 1 intuitively captures the property that any efficient adversary $A$ *can essentially only guess the challenge ciphertext.*

**Solution:** In the definition for EAV-security in figure 1, the attacker correctly guesses the challenge ciphertext $c_b$ when he chooses a $b'$ such that $b = b'$:

$$\textbf{PrivK-EAV}_{A,\Pi}(n) = 1 \text{ if } b = b'$$

We know that the trusted source $T$ generates $b$ by randomly and uniformly picking it from $\{0, 1\} \rightarrow b$, so $b$ can take the value of either 0 with probability $\frac{1}{2}$ or 1 with probability $\frac{1}{2}$:

$$\Pr[b = 1] = \Pr[b = 0] = \frac{1}{2}$$

The attacker is choosing a $b'$ between either 0 or 1 - so his probability of choosing correctly is $\frac{1}{2}$:

$$\begin{aligned}
\Pr[b' = b] &= \Pr[b = 0 \cap b' = 0] \cup \Pr[b = 1 \cap b' = 1] \\
&= (\Pr[b = 0] \cdot \Pr[b' = 0]) + (\Pr[b = 1] \cdot \Pr[b' = 1]) \\
&= (\frac{1}{2} \cdot \frac{1}{2}) + (\frac{1}{2} \cdot \frac{1}{2}) \\
&= \frac{1}{4} + \frac{1}{4} \\
&= \frac{1}{2}
\end{aligned}$$

Since the probability ($\frac{1}{2}$) of correctly choosing $b'$ is the same as the probability of randomly guessing the challenge ciphertext $c_b$ from between two possible messages $m_0$ or $m_1$ ($\frac{1}{2}$), even an efficient attacker can essentially only guess the challenge ciphertext.

2. Explain how the definition in Figure 2 intuitively captures the property that any efficient adversary $A$ *behaves essentially the same no matter what the challenge ciphertext is.*

**Solution:** The alternative formulation in figure 2 is EAV-secure such that:

$$|\Pr[O_A(n, 0) = 1] - \Pr[O_A(n, 1) = 1]| \leq v(n)$$

where $v(n)$ is some negligible function. Here $O_A(n, 0) = 1$ represents the case that the attacker outputs $b' = 1$ when it is playing against $b = 0$, and $O_A(n, 1) = 1$ represents the case that the attacker outputs $b' = 1$ when it is playing against $b = 1$.

Since the difference between the probabilities of these cases are negligible, and there are only two possible outputs for the attacker and two possible outcomes for the game, the same could be said for the case that the attacker outputs $b' = 0$:

$$|\Pr[O_A(n, 0) = 0] - \Pr[O_A(n, 1) = 0]| \leq v(n)$$

In both cases, regardless of the correct value of $b$, the attacker chooses $b' = 1$ or $b' = 0$ with essentially the same probability. Thus, we can conclude that the attacker is playing the game the same way regardless of what the challenge ciphertext is.

3. Show that the alternative formulation of Figure 2 implies EAV-security as defined in Figure 1.

**Solution:** The condition for EAV-security as defined in Figure 1 is:

$$\Pr[\text{PrivK-EAV}_{A,\Pi}(n) = 1] \leq \frac{1}{2} + v(n)$$

Figure 2 gives an alternative formulation of EAV-security, which states that for any adversary $A$, there exists a negligible probability function $v(n)$ such that:

$$\big|\Pr[\mathrm{O}_A(n,0)=1]-\Pr[\mathrm{O}_A(n,1)=1]\big|\leq v(n)$$

Let's consider the case that the attacker chooses $b'=1$. We can rearrange the previous statement using some algebra:

$$\begin{aligned}|\Pr[\mathrm{O}_A(n,0)=1]-\Pr[\mathrm{O}_A(n,1)=1]|&\leq v(n)\\ |\Pr[\mathrm{O}_A(n,1)=1]|&\leq v(n)\pm\Pr[\mathrm{O}_A(n,0)=1]\end{aligned}$$

Recall that $\mathrm{O}_A(n,b)$ represents the guess $b'$ that the attacker makes when playing against the unknown input $b$. Given that $b\in 0,1$ can only take one of two possible values, the probability of obtaining either value is $\frac{1}{2}$:

$$|\Pr[\mathrm{O}_A(n,1)=1]|\leq v(n)\pm\frac{1}{2}$$

Here, $\mathrm{O}_A(n,1)=1$ represents the attacker guessing $b'=b=1$. In Figure 1, the outcome $\text{PrivK-EAV}_{A,\Pi}(n)=1$ only when $b=b'$. Thus, $\mathrm{O}_A(n,1)=1$ and $\text{PrivK-EAV}_{A,\Pi}(n)=1$ are equivalent cases:

$$|\Pr[\text{PrivK-EAV}_{A,\Pi}(n)=1]|\leq v(n)\pm\frac{1}{2}$$

We can drop the absolute value bars, and replace the $\pm$ with simply $+$ since $v(n)$ can be any negligible polynomial probability function, and simply write:

$$\Pr[\text{PrivK-EAV}_{A,\Pi}(n)=1]\leq\frac{1}{2}+v(n)$$

Which is equivalent to the condition for EAV-security as it is defined in Figure 1.

4. In reference to the game in Figure 1, how do CPA-security and CCA-security extend EAV-security?

**Solution:** CPA-security and CCA-security are strengthened extensions of EAV-security as it is defined in Figure 1.

CPA-security, otherwise known as Chosen-Plaintext Attack security, extends EAV-security by allowing the adversary to choose the plaintext messages for encryption. In the CPA-security game, the adversary is allowed to send a plaintext message $m_i$ and receives the ciphertext $c_i$. He can do this as many times as he would like before sending the message pair $m_0$ and $m_1$ that he will have to guess $b'$ for. CPA-security implies that having access to many plaintext-ciphertext conversions will not reveal any information about the plaintext or the encryption scheme to the adversary, thus it requires a non-deterministic encryption scheme. A scheme is again considered CPA-secure if no efficient adversary can win the CPA-security game with probability significantly greater than 1/2.

CCA-security, otherwise known as Chosen-Ciphertext Attack security, extends EAV-security by allowing to decrypt any ciphertext of its choosing, except for the challenge ciphertext. A scheme is considered CCA-secure if no efficient adversary can win the CCA-security game with probability

significantly greater than 1/2. Again, probabilistic (non-deterministic) encryption is required to be CCA-secure.

# 3 On leaking, or hiding, the message length

Consider encryption scheme $\Pi$ = (Gen, Enc, Dec) that is EAV-secure w.r.t. the definition in Figure 1.

1. What does the condition $|m_0| = |m_1|$ capture? Does it weaken or strengthen $\Pi's$ security?

**Solution:** The condition $|m_0| = |m_1|$ means that the adversary who is playing the game can only send two messages $m_0$ and $m_1$ if they share the same length. This condition weakens the security of $\Pi$, as although it protects the *contents* of the message itself, it concedes in letting the adversary know the *length* of the message.

2. Let PrivK-EAV2$_{A,\Pi'}(n)$ be the game in Figure 1 where $A$ is allowed to choose challenge messages of *arbitrary length* for breaking encryption scheme $\Pi'$ = (Gen', Enc', Dec'), and consider the security notion derived by this modified game: Intuitively, $\Pi'$ is EAV2-secure, if no efficient $A$ can determine $c_b$ better than guessing, even when $|m_0| \neq |m_1|$. Show that no EAV2-secure scheme $\Pi'$ exists.

**Solution:** In $\Pi'$, the message space is defined over $M = \{0,1\}^*$, meaning messages can be arbitrary in length. Under this new scheme, attackers could choose two plaintext messages with a large difference in length, and exploit the varying lengths to gain information about which ciphertext they have received.

Assume that the polynomial $p(n)$ is an upper bound on the time spent by $Enc'$ for encrypting a single bit. Let $m_0 \in \{0,1\}$ be a message of a single bit in length, and some random $m_1 \in \{0,1\}^{p(n)+c}$ with $c \geq 1$ be a longer message. Then,

$$c_0 = \text{Enc}(m_0) \text{ where } |c_0| \leq p(n)$$
$$c_1 = \text{Enc}(m_1) \text{ wher } |c_1| > p(n) + 1$$

In this example, the attacker can win the game by choosing $b' = 0$ when $|c_b| \leq p(n)$ and choosing $b' = 1$ otherwise. Thus, such an EAV2-secure scheme $\Pi'$ cannot exists.

3. Show that EAV2-security can be achieved by encryption schemes defined over messages up to a given *maximum* length, i.e., by schemes $\Pi'$ such that for $k \in \{0,1\}^n$, algorithm Enc$'_k$ is defined over message space $M' = \{m : |m| \leq l\}$, where $l \triangleq l(n)$ for some given polynomial $l(\cdot)$.

**Solution:** If we assume that $\Pi'$ restricts the message space to messages up to a maximum length $l$, then EAV2-security can be achieved while still allowing the attacker some control over the length of the messages he chooses to send.

Consider the scheme $\Pi'$ = (Gen', Enc', Dec') that has an identical key generation function as $\Pi$ = (Gen, Enc, Dec) as defined in Figure 1 of the last problem, but differs from $\Pi$ in that the encryption function Enc' will first left-pad the message $m$ with bits $1^{(l-|m|-1)}0$ such that the length of the new message $m'$ is the maximum length of the message that is allowed, then encrypt as normal with Enc.

When decrypting, Dec' will decrypt with Dec then remove all leading 1 bits up to and including the first 0 bit.

The definition of EAV2-security is identical to that of EAV-security other than in that messages $m_0$ and $m_1$ in EAV2-secure schemes may be differing in length. We can then use the formal definition for EAV-security to show that our scheme is EAV2-secure, given that we have already constructed $\Pi'$ to accept messages of length up to a maximum of $l$ and that the schemes are otherwise the same. Specifically, we can show that if an attacker $A$ can break the EAV-security of $\Pi'$ with non-negligible probability, then we can construct an attacker $A'$ that can break the EAV-security of $\Pi$ with non-negligible probability, which would contradict the assumption that $\Pi$ is EAV-secure.

- $A'$ receives the challenge ciphertext $c$ from $A$
- $A'$ constructs two plaintext messages $m_0$ and $m_1$ of fixed length $l$, such that $m_0$ is the left-padded version of the shortest possible message in $M$ and $m_1$ is a message of the longest possible length $l$.
- $A'$ sends the two messages $m_0$ and $m_1$ to the trusted source $T'$ of $\Pi'$.
- $T'$ left-pads $m_0$ and $m_1$ with bits as described above up to length $l$, and encrypts one of them (chosen uniformly at random) to produce a ciphertext $c'$, which is returned to $A'$.
- $A'$ returns $c'$ to $A$ as the challenge ciphertext.
- $A$ outputs a guess $b'$ of which plaintext was encrypted in $c'$.
- $A'$ outputs the same guess $b'$ as $A$.

So, $A'$ can break the EAV-security of $\Pi$ with non-negligible probability if $A$ can break the EAV2-security of $\Pi'$ with non-negligible probability.

Assume that $A$ breaks the EAV2-security of $\Pi'$ with non-negligible probability $v(n)$. Then, the probability that $A$ guesses the correct plaintext message in the EAV-security game of $\Pi'$ is greater than $\frac{1}{2} + v'(n)$ for some non-negligible $v'(n)$.

Now consider the EAV-secure scheme of $\Pi$. If $c$ is the challenge ciphertext, then $A'$ sends two messages $m_0$ and $m_1$ to the trusted source $T'$ of $\Pi'$, and receives a ciphertext $c'$ as a response. Since $m_0$ and $m_1$ are the left-padded versions of the shortest and longest messages of length $l$, respectively, we know that $m_0$ and $m_1$ have the same length. Since $m_0$ and $m_1$ have the same length, $c'$ is a valid ciphertext for the EAV-secure scheme $\Pi$.

If $A$ can distinguish between the two messages with non-negligible probability $v(n)$, then $A'$ can distinguish between the two messages in the EAV-security game of $\Pi$ with the same non-negligible probability $v(n)$. This contradicts the assumption that $\Pi$ is EAV-secure (which we are given from the last problem), and hence we conclude that $\Pi'$ is also EAV-secure, and thus EAV-2 secure as it also accepts messages of varying lengths.