OTP and AES: A historical transition between two systems of cryptography

Valdemar Thanner
Kantonsschule Zug
supervised by Mr. Bernhard Keller

August 26, 2016

Contents

Chapter 1

OTP: The One Time Pad

1.1 What is a "One Time Pad"?

When speaking about OTP, it is important to distinguish between its two meanings: On the one hand, it is a technique used to encrypt information. This technique requires one single key, used both to encrypt and decrypt the information. This key is also referred to as a one time pad; therefore, it is important to distinguish between the one time pad (a cryptographical technique) and a one time pad (a key which is used to encrypt and decrypt information).

The One Time Pad is largely derived from the Vernam cipher, which is named after Gilbert Vernam. The Vernam cipher utilized a perforated tape (one of the earliest types of data storage) as the secret key[1]. Each bit of data was stored in the form of a hole punched into the perforated tape.

Figure 1.1: Perforated tape, utilized to store bits as punched holes

However, this system had a vulnerability which the One-Time Pad solved: In Vernam's original method, the perforated tape was not exchanged after it had completed one cycle; instead, it was looped around continuously, often being used to encrypt multiple different messages.

This made the entire system vulnerable. The re-usage of the key meant that the resulting ciphertext suffered from a so-called known-plaintext vulnerability [2]. This means that, if a plaintext and its corresponding ciphertext are captured, the key utilized to generate the ciphertext can be derived from them. This is not an issue if the key is exchanged each time a new message is encrypted. However, if the key of any Vernam cipher machine was compromised in this fashion, any further intercepted ciphertext could be decrypted.

1.2 Method used

In the following section, the utilized method will be clarified through usage of an example. In this example, the message *"cryptography"* will first be encrypted by it's sender, sent to it's intended recipient, and finally decoded by the recipient.

1.2.1 Generation of the random key

In order to encrypt the plaintext, a key must first be generated. This key will be utilized to encrypt the plaintext through the usage of modular addition, turning it into the ciphertext.

This key must fulfil some crucial criteria. Foremost, the length of the key (the amount of characters contained within it) must be equivalent to or greater than the length of the plaintext; otherwise, it is not possible to perform any encryption (using the OTP). Secondly, the key must be generated randomly. This is mainly due to the fact that a randomly generated key makes frequency analysis[3], the form of cryptanalysis most commonly used to break classical ciphers, impossible.

The key consists of numbers. Usually, when the plaintext is made up of Latin letters, the numbers range between 0 and 25. The key can be converted into Latin letters through the same method applied to the plaintext outlined in the following chapter, however, this is not necessary, although the key is often transported in the form of text.

As the message being encrypted in this example has 12 characters, the key must also posses at least 12 characters. For the sake of this example, the key *"sytruifgnihm"* will be utilized.

1.2.2 Modular addition of the key and plaintext

Next, the ciphertext is created through modular addition of the key and the plaintext. This can be applied not only to a message consisting of alphabetical characters, but also to any sequence of bits. If the plaintext consists of a message made up of alphabetical characters, the plaintext and the key are added using arithmetic referred to as *"addition modulo 26"*. The correct mathematical notation for modular arithmetic is $(a + b)\ mod\ c$, where c is referred to as the *modulus*, which is the value that cannot be passed in modular addition. In order to perform modular addition, the variables a and b are first added, after which they are divided by the modulus c, up to an

integer. The resulting remainder r is the final result of the operation.

Before the modular addition of the plaintext and the key can begin, each character (of the plaintext as well as of the secret key, in the case that the key was generated as a string of Latin letters instead of as a sequence of numbers) must be converted to a number, corresponding to it's position in the Latin alphabet:

c	r	y	p	t	o	g	r	a	p	h	y
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
2	17	24	15	19	14	6	17	0	15	7	24

As the key was also generated in the form of characters, it too must be converted to a sequence of numbers:

s	y	t	r	u	i	f	q	n	i	h	m
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
18	24	19	17	20	8	5	16	13	8	7	12

Afterwards, the key and the plaintext are added together utilizing modular addition. Of course, all operations below are performed in *mod 26*. Finally, the resulting numbers are converted to the character to which they correspond in the latin alphabet. The resulting sequence of characters is referred to as the ciphertext.

2+18	17+24	24+19	15+17	19+20	14+8	6+5	17+16	0+13	15+8	7+7	24+12
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
20	15	17	6	13	22	11	7	13	23	14	10
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
u	p	r	g	n	w	l	h	n	x	o	k

1.2.3 decoding of the ciphertext using the key

Assuming the message has been delivered to the intended recipient, who must hold a copy of the secret key (which, as OTP is a symmetrical cryptographical

method, is identical to the one utilized to create the ciphertext), the recipient can now decode the ciphertext in order to view the plaintext.

Since the ciphertext was created through the modular addition of the plaintext and the key, the recipient can utilize modular subtraction in order to view the plaintext. In order to do this, the recipient must subtract the key from the ciphertext in modulo 26, and convert the resulting numbers to Latin characters. However, it is now also necessary for the numbers not to become negative. Fortunatley, this is also made possible by modular arithmetic, as the values simply loop back around from 0 as well. Once again, all below operations are in *mod 26*.

20-18	15-24	17-19	6-17	13-20	22-8	11-5	7-16	13-13	23-8	14-7	10-12
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
2	17	24	15	19	14	6	17	0	15	7	24
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
c	r	y	p	t	o	g	r	a	p	h	y

Now, the message's journey is complete, having passed from plaintext into ciphertext, being transported to its intended recipient in the form of ciphertext, and finally being decoded and read by its recipient.

1.3 Perfect secrecy: Information-theoretical security

1.3.1 definition

1.3.2 Why can only OTP achieve perfect secrecy?

1.4 Issues with OTP

1.4.1 True randomness in generating the key

1.4.2 Secure distribution of the key itself

1.4.3 Secure disposal of a utilized key

Bibliography

[1] G.S Vernam. "Secret Signaling system". Pat. U.S. Patent 1,310,719. 1919.

[2] Gordon Welchman. *The Hut Six Story: Breaking the Enigma Codes*. 1997.

[3] learncryptography.com. *Frequency Analysis*. 2014. URL: `https://learncryptography.com/cryptanalysis/frequency-analysis`.