THE FALL OF RC4 AFTER A LONG HISTORY.

Alexandra Poier and Peter Lorenz
Graz University of Technology

In the Beginning ...

In the winter semester 2013, the course "Einführung in das wissenschaftliche Arbeiten" takes place at IAIK - Institute for Applied Information Processing and at Communications Graz University of Technology. In this course, we get the chance to write an extended abstract about "Symmetric Cryptography" and we can deepen in this topic. After skimming through pdf's for hours, we decided to focus our work on RC4 and AES cryptography standards.

1. INTRODUCTION

Rivest Cipher 4 (RC4) is a type of symmetric cryptography, which is the most common spread cryptographic method. It is used in many different network devices and software programs, e. g. Wireless, LAN and HTTP, generally state that is implemented in the SSL (Secure Sockets Layer) and TLS (Transport Layer Security) protocol. The RC4 is a stream cipher algorithm which is very fast and simply to implement. In 1987, RC4 was designed and it is still used, although there are still more secure cryptographic solutions such as the Advanced Encryption Standard (AES). The AES was introduced by Joan Daemen and Vincent Rijmen in year 2000, it has not been asserted yet.

The first part deals with the question: "Why is it in use?" We give a short overview of symmetric cryptography in order to show the range of symmetric cryptography. Compared to the other symmetric cryptographic methods, RC4 is a simple algorithm and can be implemented fast. "The heart of RC4 is its exceptionally simple and extremely efficient pseudo-random generator." ["RSA", 2001] Nowadays, people are not aware of weaknesses of this method, but it is still an accepted standard, for example in banking, proved security software are essential, and expensive to replace with an newer one.

In this work, we present the weaknesses of RC4. According to Frank Thornton [Chandra et al., 2009], it must be noted that there are no known practical attacks on RC4 as long as it is used properly. The WEP protocol (meanwhile replaced by WPA) used RC4, and this is often cited as a reason why the protocol failed, but the problem was in how WEP generated keys in combination with RC4. In fact, there are a lot of reports on attacks on RC4, which we can divide into two groups. [Mete Akgün, 2008] First group based on the weaknesses of the pseudo random generation algorithm (PRGA). The second group is based on the weaknesses of the key-scheduling algorithm (KSA). We explain them more precisely in this chapter. The main problem is the initial value (IV) which is not random at all and thus easily predictable.

In the last section, we describe how to transform the AES into the RC4. AES is a block cipher in contrast to RC4 that is a stream cipher. A block cipher is a deterministic algorithm operating on fixed-length groups of bits, called blocks, with an unvarying transformation that is specified by a symmetric key. It is not allowed to edit a part of a block, because it affects the entire block. Thus, single characters cannot be swapped by an attacker.

A stream cipher is a symmetric key cipher where plain text digits are combined with a pseudo random cipher digit stream (key stream). It works only one character by another at a time. Therefore, it is a lower error propagation, each symbol is affected by only itself. According to Florian Mendel, a cryptography specialist at Graz University of Technology, said that it is not difficult converting a block into a stream cipher. Microsoft has already announced to get rid of RC4 in 2016. [Leyden, 2013] This change of the biggest corporation will affect many other fields, for example the security for banks.

Organization. The next section describes some fundamental knowledge about cryptography in general. In section 3 we define the significance and weaknesses of RC4, and give a short description in subsection 4.1. In subsection 4.2 we give an overview of the known attacks on RC4 and give some examples. Section 5 discusses how AES will replace RC4 and what impacts will occur on all networks (Wireless, LAN, ...). Finally, subsection 5.1 presents the block cipher AES and last but not least we describe how to transform a block into a stream cipher in subsection 5.2. Finally, we will conclude the paper in section 6.

2. FUNDAMENTAL KNOWLEDGE

In this section, we introduce the related background knowledge of symmetric cryptography to facilitate the understanding of this paper.

Cryptography is the study of mathematical techniques related to aspects of information security such as confidentiality, data integrity, entity authentication, and data ori- gin authentication. Cryptography is not the only means of providing information security, but rather one set of techniques. [Praphul Chandra et al., 1997]

There are two different types of cryptographic algorithms:

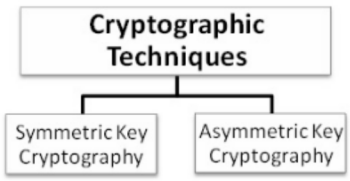

Fig. 1. Cryptographic Algorithms

Every cryptographic system which is used, is a symmetric one. This has to be this way, because the transmitter as well as the receiver need to share the same piece of information in a secure but also secret way. On the one hand it is possible to use the same encryption key, but on the other hand they can also use one of a pair of related keys, who are easily computed from each other. Both, the sender as well as the receiver are using this key system to protect each other from the uncertainty to an unauthorized receiver, for example a hacker.

Symmetric cryptography is also called “shared key” or “shared secret” encryption, which means that both, the transmitter, as well as the receiver are knowing and using the same key, like Figure 2 shows:

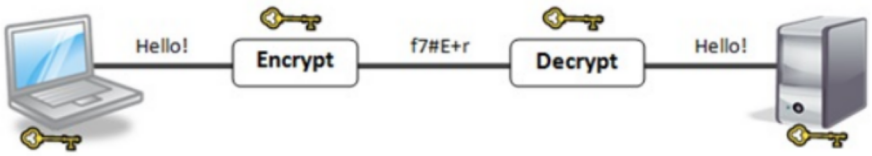

Fig. 2. Symmetric Cryptography

Furthermore there are existing asymmetric encryption systems, where the keys of the transmitter and receiver are not similar, as showed in figure 3. It is nearly impossible to figure out at least one of the characters of the other ones key. Asymmetric key encryption systems are the reason which makes it possible to verify messages who have to be shown to anyone outstanding or allow the transmitter whose key has been compromised to communicate in privacy to a receiver whose key has been kept secret. This is not possible if you use only symmetric cryptography.

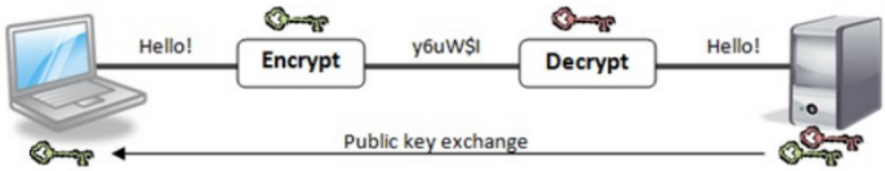

Fig. 3. Asymmetric Cryptography

2.1 Conclusions

In general, cryptography is used to encrypt sensitive data, to ensure data integrity and authentication of messages. To achieve these goals, symmetric or asymmetric encryption can be used. Symmetric cryptography uses the same security key, in contrast to asymmetric cryptography, which uses two different keys as shown in Figure 2 and 3.

3. THE SIGNIFANCE AND WEAKNESSES

In this section, we describe the weaknesses on RC4. It is interesting that there are already appeared several reports on different attacks. At first, we give a short description how the RC4 is designed and afterwards we show the typical attacks on this stream cipher. Subsequently, there are some examples.

3.1 Description

The RC4 is a stream cipher, which is shown on graphic Figure 4. It means that it generates a pseudorandom stream of bits, which is called keystream. It can be used, e. g. for encryption or decryption a plain text.

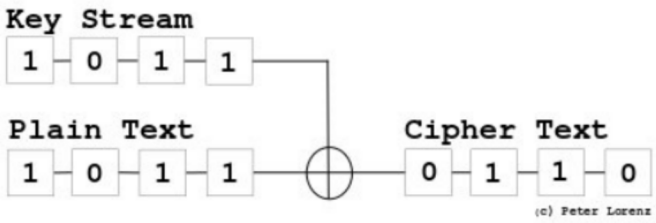

Fig. 4. Stream Cipher

The algorithm is distinguished into two stages:

(1) Initializiation:
KSA (Key-scheduling algorithm) is used to initialize the permutation with the length of the key that is variable and mostly between 0 and 255 bytes to generate a pseudo-random stream, which is XORed with the plaintext to give the ciphers.

(2) Operation:
In this stage, Pseudo-random generation algorithm (PRGA) modifies the outputs of the keystream. The PRGA increments `i` seeks the ith element of `S` and adds that to `j`, and consequently exchanges the values of `S[i] and S[j]`. Then it uses the sum `S[i] + S[j] modulo 256` as an index to call a next element of `S`, which is transfered to the keyvalue `K` and finally XORed with the following byte of the message to produce the ciphers. Every element of `S` is swapped with another elment every 256 iterations at least once.

```
i = 0;
j = 0;
while GeneratingOutput:
{
    i = (i + 1) % 255;
    j = (j + S[i]) % 255;
    swap values of S[i] and S[j];
    K = S[(S[i] + S[j]) % 255];
    output K;
}
```

3.2 Attacks

In this section, we will introduce the most recent attacks on RC4. Stream ciphers are generally studied with respect to a number of common attack models. These attack models are considered good models to study the security of stream ciphers in, but they do not cover all possible attacks.

3.2.1 *Finney.* In year 1994, H. Finney [Carmeli, 1994] made an important discovery. He found out that there is an cycle of RC4 that can't happen, because it will never enter - or with other words: he has shown that determined states can occur, regardless of the key. In his "class of states" he summarizes all the states for which $j = i + 1$ and $s[j] = 1$. There are about n^{-2} possibilities that this problem occurs, which is called "Finney's forbidden states".

According to the Book "RC4 Stream Cipher and Its Variants"[.....], we found out that the "Finney cycle", refers to a cycle which is formed in the state space. Every state in this cycle has the same relations and the round of PRGA (Pseudo Random Generation Algorithm), the part of the RC4 process that outputs a streaming key based on the key scheduling algorithms pseudo random state array, is reversible. That's where the problem starts, because there is no possibility to move from this state to a place in a different state, which is called "Finney state".

All in all you can say that Finney states are preserved, if RC4 is in an certain state, prior and future states are the same states too, otherwise it would not work. However there is also a good news: The initialization in RC4 Pseudo Random Generation Algorithm cannot occur in normal operation.

3.2.2 *Itsik Mantin and Adi Shamir.* At the time when WEP has not been replaced by WPA, the most serious weakness in RC4 has been discovered by the cryptographic specialists Itsik Mantin and Adi Shamir [Mantin, 2005a] in 2001. They found out that the possibility that zero occurs at the second round is twice as large as expected. This weakness can be exploit by only one attack at broadcast applications.

Mantin and Shamir have established that RC4 is completely insecure, if it is used in combination with WEP which uses a fixed security key that goes hand in hand with a IV modifier that is already known. So if they know this IV modifiers they can easily encrypt different messages. If the IVs are generated by a multibyte counter, in the way that the last byte grows most rabidly, an attacker can search for them. That's not easy, because there are around 1.000.000 packets. So if the counter starts from zero, he has to "frisk", but it will work.

Nevertheless, it is possible that the counter starts not at zero, but at another point. In this case the attacker just has to use another strategy. He could, for instance, assume the first few bytes of the secret key and search IVs that fix the permutation. When the attacker uses the first two bytes of the secret key, there are approximately 4 settings for every starting IV byte of the two remaining ones. Those are fixing the permutation which is essential to regain a specific key byte. So if the attacker has around one million packets and additional 2^{16} percent more work (referring to their paper), it is still possible for him to regain the key.

In the industry they have already tried to extend the length of the WEP key byte individually, but that doesn't protect them from further attacks, they are just more complicated.
Nowadays this is no longer a problem, because WEP has already replaced by WPA like mentioned before. Since WPA has been introduced the problem doesn't exist anymore.

3.2.3 *Grosul and Wallach.* Grosul and Wallach [Wallach, 2010] observed that, in the near of RC4 could arise collisions, if the length of the key is close to the bursting 256 bytes in the first key pairs which collides. The hamming distance one means that both keys are different in their position.
In other colliding keys, pairs with hamming distance three were also found. Those findings are also created at "near colliding keys" of RC4, where output states are generated with slight hamming distances after key scheduling algorithm.
The findings of Grosul and Wallach are important, because that means that the key scheduling algorithm can be controlled.

3.2.4 *Robert J. Jenkins.* In 1994 Robert J. Jenkins [Mantin, 2005b] found a known, but ignored weakness on the stream cipher RC4. He found out, that secret states leads to leaks at the key stream, which is today called "Jenkins' correlation" or "the RC4 glimpse". This matter can be used by attacks on RC4 in several modes.

This new attack is consistent with regard to the part with the first 256 bytes of the key stream. The complexity grows only linearly with the length the key. It can cause an attack by creating 217 short key streams of different IVs in an exemplary setting of the parameters. One other possibility to attack it, if it succeeds IV, the secret root key to crack it. This mounted to a key recovery attack that recovers the secret key stream by analyzing just a single word of 222. So you just have to improve the attack of Fluhrer et al. by winning several IVs.
In addition, the attacker can inject an error for execution of RC4 by generating the internal state and the secret key by analyzing 214 faulty key streams from the key.

3.2.5 *Ilya Mirnonov.* Ilya Mirnonov [Takanori Isobe, 2001] constructed RC4 as a "Markov chain", which means that if you know only a limited prehistory of this chain, you are able to predict how the chain will develop in future as if you had known the entire history of this chain.

Mironov strongly recommended to discard the first 12 N bytes of the output stream (at least 3N) to achieve a uniform distribution of the initial permutation of the elements. The attack of Ilya Mironov is able to win each plaintext byte of pure digit texts. For this purpose, there are simply used different keys.

For instance: There is a high probability that only the first 250 bytes of the plaintext are obtained of 234 cipher texts. However, it should be noted, that the attack on RC4 has the advantage of the sequential recovery of plaintext bytes. So if the initial 256/512/768 bytes of the keystream are suppressed in the protocol, the attack does not work anymore. The widely used protocols such as SSL or TLS do use the first bytes of this key stream, which means that the attack will work out in this case.

3.2.6 *Andrew Roos.* Andrew Roos [Roos, 1995] summed up the basis for a simple known plaintext attack. This attack reduces the effective key space which is needed to search by a probability around 50 percent to discover a key to about 11.2 bits. The disadvantage is that quite a few known plaintexts are needed to start the attack.

Roos describes the function in 5 Points:
(a) Collect a large number of known plaintexts (generator sequences)
(b) Discard generator sequences which do not start with "00.00" or "FF.03"
(c) For generator streams from "00.00" search all keys that begin with "00.00.FD"
(d) For generator streams from "FF.03" search all keys that begin with "03.FD.FC"
(e) Carry on until you find a key.

Of course, this attack will only recognize a very small number of keys. However, most generator sequences are discarded without being searched before. For those who have been searched, the search is 2^{24} times smaller than necessary to search the full key space .
To sum it up in Roo's words: "The number of trials to determine a key is much smaller than brute force alone."

3.3 Conclusions

In this chapter, we have given the weaknesses of RC4. The number has been increased rapidly last years and therefore we selected the most significant ones in order to emphasize that it is important to replace the RC4 because of its lack of security.

4. HOW WILL RC4 BE REPLACED IN THE FUTURE?

We have already shown enough weaknesses of the RC4 in this paper and it has been proved that it is an obsolete algorithm. Experts are looking forward to implement the AES, an algorithm standard, which is a block cipher. AES, based on Rijandel, may be the cryptographic method to replace the current RC4. In the following, we explain how to transform a block into a stream cipher.

4.1 AES - The block Cipher

We present the design of AES in order to have fundamental knowledge to understand the transformation in the next chapter.
AES as well as the origin Rijandel are based on the design principle, known as a substitution - permutation networks. It has a fixed block size of 128 bits, which is expanded to a 4x4 matrix.
A word consists of four bytes, that is 32 bits. Therefore, each column of the state array is a word, as is each row. The matrix shows one state block:

$$\begin{matrix} byte_1 & byte_2 & byte_3 & byte_4 \\ byte_5 & byte_6 & byte_7 & byte_8 \\ byte_9 & byte_{10} & byte_{11} & byte_{12} \\ byte_{13} & byte_{14} & byte_{15} & byte_{16} \end{matrix} \quad (1)$$

The key size varies between 128, 192, or 256 bits. This is important, because it specifies the number of repitions and transformation rounds that encrypt the plain text. The final output is called cipher text. The number of cycles are as follows:

(a) 10 cycles of repetition for 128-bit keys.
(b) 12 cycles of repetition for 192-bit keys.
(c) 14 cycles of repetition for 256-bit keys.

During each round, the following operations are applied on the state:

(a) SubBytes: Each byte in the state will be replaced by another one
(b) ShiftRow: Each row in the 4x4 array is shifted to the left for unkown times
(c) MixColumn: A linear transformation on the columns of the state
(d) AddRoundKey: Each byte of the state is bounded to a round key, which is a different key for each round

4.2 Transformation of a block cipher into a stream cipher

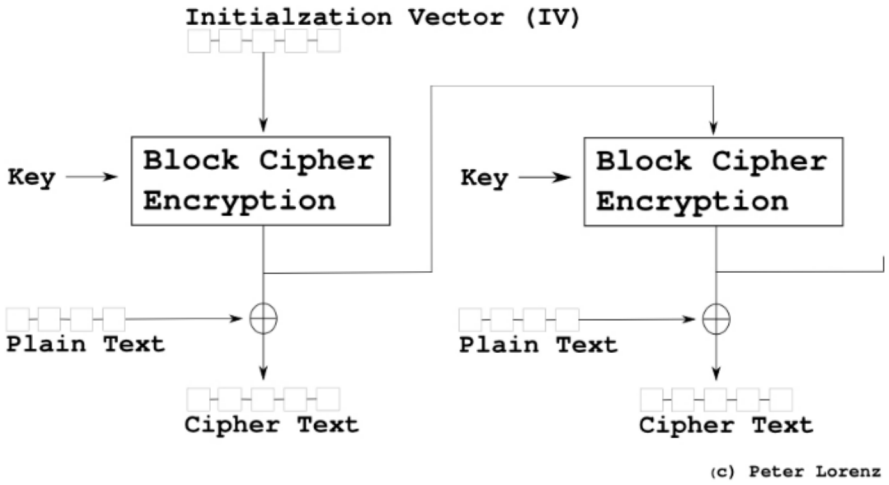

Fig. 5. Outputfeedbackmode

One method to transform a block cipher into a stream cipher is called "Output Feedback Mode" (OFB). OFB is a working model, in which block ciphers could be operated. The operation starts with an official known initialization vector that gets encrypted by the block cipher. Next, the encrypted plaintext is encrypted once more and so on. Within this process, a block algorithm like DES or AES is used to build a synchron stream cipher. The output of the block cipher is linked to the plaintext to build a secret key which creates a stream cipher. The output of the block cipher is fed back to its input, to build up a continuous stream of cipher text blocks.

4.3 Conclusions

In this chapter we have tried to make explicit, that the transformation of block chiphers into stream ciphers is realisable so that there are not any obstacles to replace RC4 with AES.

5. CONCLUSIONS

In the first section we gave a short overview of what we explained in this paper. In the second section we described some fundamental knowledge about cryptography in general. Next, we defined the significance and weaknesses in the third section and gave a short description in subsection 3.1. In subsection 3.2 we gave an overview of the known attacks on RC4 and some examples. In section four we showed how RC4 would be replaced in the future. Finally subsection 5.1 presented the transformation from a block cipher like AES to a stream cipher like RC4 by Output Feedback Mode.

REFERENCES

Alfred J. Menezes, Paul C. Van Oorschot, S. V. 1997. Handbook of Applied Cryptography. CRC Press.

Carmeli, E. B. . Y. 1994. Eﬃcient Reconstruction of RC4 Keys from Internal States. Computer Science Department Technion - Israel Institute of Technology Haifa 3200, Israel.

Leyden, J. 2013. Microsoft, Cisco: RC4 encryption considered harmful, avoid at all costs. The Register Inc.

Mantin, I. 2005a. A Practical Attack on the Fixed RC4 in the WEP Mode. NDS Technologies, Israel.

Mantin, I. 2005b. A Practical Attack on the Fixed RC4 in the WEP Mode. NDS Technologies, Israel.

Mete Akgün, Pinar Kavak, . H. D. 2008. New Results on the Key Scheduling Algorithm of RC4. Springer Inc.

Praphul Chandra, Dan Bensky, T. B. C. H. S. R. J. R. J. F. R. T. S. G. L. S. F. T. C. L. . J. S. W. 2009. Wireless Security: Know it all. Elsevier Inc.

Roos, A. 1995. Weak Keys in RC4 (one Week Later...), Cryptanalysis of RC4 { Preliminary Results. Netfuture: The future of networking.

\RSA". 2001. RSA Laboratiories. RSA Security Response to Weaknesses in Key Scheduling Algorithm of RC4. In Technical Notes and Reports. EMC: Global Site Selector.

Takanori Isobe, Toshihiro Ohigashi, Y. W. . M. M. 2001. Full Plaintext Recovery Attack on Broadcast RC4. Kobe University.

Wallach, G. . 2010. Security and Cryptography for Networks: 7th International Conference. Springer Berlin Heidelberg: Selected Areas in Cryptography.