EE4215 Cybersecurity Technology

Assignment 2 Due: Feb 1 (Wed) at 11 am

Computer Exercise: Attack on RC4 (18 points + 4 bonus points)

In this exercise, you are required to implement a practical attack on RC4 implemented in the WEP protocol of the IEEE 802.11 standard. The 802.11 frame includes the WEP-encrypted data. Each frame carries a message m, encrypted with a specific RC4 key. The RC4 key, k, is of 16-byte long, which consists of a 3-byte initialization vector (IV) followed by a 13-byte long-term, private key. The value of IV is increased by 1 after each frame while the private key remains the same for all frames. For example, if the IV of a certain frame is $01\ \ FF\ \ 09$ (in hexadecimal format), then the next frame is $01\ \ FF\ \ 0A$.

Consider the following scenario:

- The message carried by each 802.11 frame is a network-layer datagram, which has a fixed header format. Assume that the first byte of all the messages has a constant value, denoted by $m[0]$.
- The attacker has access to many packets, so he can wait for the use of some specific IV values.

The goal of the attack is to obtain $m[0]$ and the 13-byte private key. It is launched based on the following facts:

i. If $\text{IV} = 01\ \ FF\ \ x$ (for any value x), the first byte of the keystream equals $x + 2$ with a high probability.

ii. If $\text{IV} = 03\ \ FF\ \ x$ (for any value x), the first byte of the keystream equals $x + 6 + k[0]$ with a high probability, where $k[0]$ denotes the first byte of the 13-byte private key.

iii. In general, for $z = 3, 4, \ldots, 15$, if $\text{IV} = z\ \ FF\ \ x$ (for any value x), the first byte of the keystream equals $x + d[z-3] + k[0] + k[1] + \cdots + k[z-3]$ with a high probability, where $d[z-3] = 1 + 2 + \cdots + z$ and $k[i]$ denotes the i-th byte of the private key. (You may verify that fact iii reduces to fact ii when $z = 3$.)

Assume after a long time, the attacker has collected some encrypted data under specific IV. The encrypted data for $\text{IV} = z\ \ FF\ \ x$ is stored in the file named `zFFx.txt`. Totally, there are 15 data files, `01FFx.txt, 02FFx.txt, …, 0FFFx.txt`.

In each data file, the first column represents the values of IV and the second column represents the corresponding first byte of the encrypted data, denoted by $c[0]$. Use the above facts and the provided data files to decode $m[0]$ and the 13-byte private key. You may use any programming language (e.g., python, C, …) to implement the RC4 attack.

Submit your **source code** and **a short report** answering the following questions:

a) (6 points) What is the most probable value (in hexadecimal) for $m[0]$ and what is its frequency of occurrence? (Hint: To decode $m[0]$, using fact i, you need to compute the 256 values of $c[0]$ XOR $(x + 2)$ of the data in `01FFx.txt` and choose the most frequent one as $m[0]$.)

b) (6 points) According to your answer in (a), which network-layer protocol is used? Explain your answer.

c) (6 points) What is the most probable value (in hexadecimal) for $k[0]$ and what is its frequency of occurrence? (Hint: Use fact ii in a similar way as in (a). You also need to use your guessed value of $m[0]$ obtained in (a).)

d) (Bonus: 4 points) What is the value (in hexadecimal) of the 13-byte private key?