RC4 CIPHER ANALYSIS

Introduction

RC4 is a stream cipher which operates on data in a bit by bit manner. The main objective of RC4 algorithm is to generate a random output keystream from an initial main key and then use this to encrypt the data stream. The algorithm consists of two phases:

1) **Key scheduling Algorithm**
In this phase, the main key is expanded and further used for shuffling of the S array which has values from 0 through 256.

2) **Pseudo Random Number Generation**
After performing certain operations and swapping of indices on the S array, a single byte is generated for encrypting each byte of data. This process continues till we reach the end of the data stream.

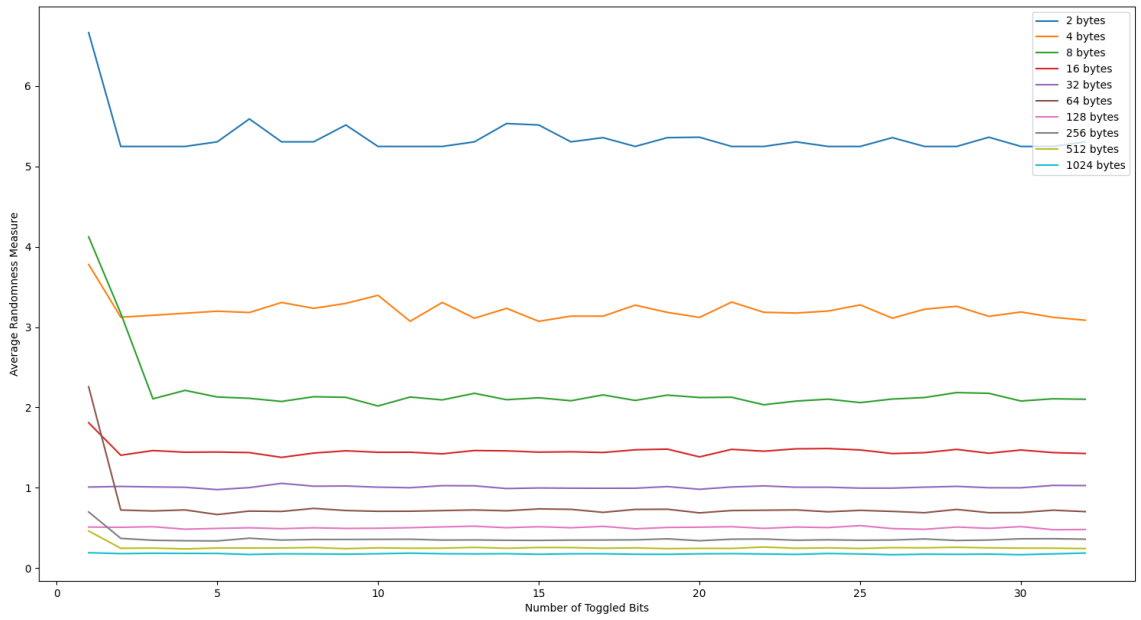
Output Analysis

From the initial main key, a different key is generated by toggling one bit at a random position. These 2 keys give very much different S arrays and output keystream.

Example of 15 bytes output stream generated with my main key:
[39, 196, 131, 88, 89, 123, 15, 19, 151, 240, 173, 0, 175, 196, 248]

Example of 15 bytes output stream generated when main key with one random bit toggled is used:
[13, 149, 224, 20, 197, 242, 246, 45, 147, 201, 103, 78, 15, 30, 161]

When we observe these 2 arrays, it is clear that there is not a single byte which is same in both cases. This randomness is more clearly observed when we toggle more number of bits and compare the output of different keys while varying the number of toggled bits and output size.



From the graph, we notice that for each output size, average randomness value gradually decreases and saturates at a certain value. Also, as we toggle more bits(1 to 32), the average moves closer to zero and nearly stretches out into a straight line. More closer the randomness value is to zero, more unpredictable is the nature of output.

Conclusion

From this experiment, we can understand the strength of the RC4 algorithm though it is very easy to implement. The experiment highlights the “avalanche effect” which is a crucial property of any encryption algorithm, where even the slightest change in input causes extreme variations in output.

Manu K Paul