Cryptanalysis of a Class of Ciphers based on Classical Method

Shresht Bhatia
ssb9839@nyu.edu

Nishi Taneja
nmt8281@nyu.edu

Harsh Patel
hrp2019@nyu.edu

I. INTRODUCTION

The following are the team members: Shresht Bhatia, Nishi Taneja, and Harsh Patel submitting their cryptanalysis based on the Index of Coincidence as part of project 1. The work done by the team is divided as follows:

1. Shresht: Implemented the encryption algorithm and worked on some test cases to check its efficiency.
2. Harsh: Worked on implementing the decryption algorithm and checked its efficiency w.r.t. The dictionary is given in the project.
3. Nishi: Understood the working of the encryption algorithm and the problem statement and formulated the decryption algorithm with other team members. Wrote the report.

II. STRATEGY

The Index of Coincidence is a technique of counting, the number of times, a letter or character that appears at the same position in two or more than two subsequences of a large ciphertext. [1]The reason IC is effective is that the coincidences caused in the ciphertext are equivalent to the coincidences caused in the plaintext, meaning the pattern of coincidence seen in two plaintext subsequences will be followed in two subsequences of ciphertext as well. This makes breaking ciphertext easy even without the knowledge of plaintext.

III. CODE

The first part of the implementation required us to find the length of the key which could be found by using the probability method (frequency of each character). The second part was finding the key, which was found using the ciphertext and the length of the key. The third part was to finally get the plaintext which can be found using the key we found the given ciphertext.

IV. CONCLUSION

To conclude, it should be noted that the linearity in the key scheduling algorithm was the crucial factor in letting this strategy work out. If that algorithm was in polynomial factor then it would have been a lot difficult to just find the length of the key used, leave alone guessing the actual key used. Also, for part 1, the length of the key is very small compared to the message length which helped in determining the pattern of the key occurring at regular intervals in the ciphertext. And since the plaintext dictionary was provided, it was easy to guess the plaintext from that.

V. REFERENCES

[1] https://en.wikipedia.org/wiki/Index_of_coincidence
[2] https://en.wikipedia.org/wiki/RC4