CRYPTANALYSIS

Art that involves deciphering and analyzing ciphers

-5th December, 2022

Meshal Cheema
19i1977
CS-F

Abeeha Fatima
19i0742
CS-F

Hajira Uzair
19i0737
CS-F

Abstract

This research presents the sustainable study of cryptanalysis and its classification. In this study, we examined the most popular cryptanalysis techniques which are growing immensely popular among all the cyber inquisitors . After analyzing the definitions, explanations, mathematical relations and theories put forward in the approbation of the cryptanalysis types, we contrived our own solution keeping in view the all the challenges like Key leakage, software faults, operating system flaws, side-channel assaults, phishing attacks, and social engineering are all examples of cyberthreats. This research is focused on providing what type of attacks are possible and to provide an eclectic concept that cryptanalysis aims to achieve and how it is advanced over the years by the investigators with respect to its fields, which sheds light on how you can compromise security measures to get access to the encrypted texts.

Categories and subject descriptors

Information Security, Cryptography, Cryptology, Data Encryption Standard

Keywords

ciphertext, plaintext, keys, encryption, decryption, attack, integral, differential, algebraic, linear, elementary, S-box, complexities

Problem Statement

Security is becoming more crucial as computers become more widespread and complicated. Cryptographic techniques and protocols are the main components of systems that protect network transmissions and data storage. The management, creation, and distribution of the keys used by the cryptographic algorithms have a big impact on how secure those systems are. Even if a cryptographic technique is perfect in theory and practice, its power will be meaningless if the key exchange processes are not properly managed.

Companies may use cryptanalysis to look for security holes. Law establishing companies are looking for ways to unlock encrypted files containing possibly crucial evidence and to be utilized by government organizations to decipher encrypted chats.

Introduction

The art of cryptanalysis entails breaking down and examining ciphers. An effort is made to produce the plaintext or, better yet, the key from the opponent's ciphertext.
It is also a relatively new area of modern science that has experienced significant growth recently, an acceptance made possible by the advancement of modern cryptography. In this context, the word "attack" means the following: In plain English, a (passive) assault on a cryptosystem is any method that first gathers some data regarding the plaintexts and the cipher texts that correspond to them under some (unknown) key.

Mathematically, we can say
Let's say that the encryption algorithm say E consists of a set of E_K functions that convert a sample of plain texts P to an encrypted sample of C ciphers, with E_K functions being indexed by key K. Similar to the encryption algorithm, we have the decryption algorithm say D is a space of D_K functions where D_K(E_K(P)) = P for every plaintext P. Stabilize the output function say O (for encrypted messages), the input functions I (for plain texts), and K (for keys) for n variables. Fix the plaintext and key distribution as well as the encryption mechanism E. An algorithm say X with a pair I, K of inputs and one output O, such that there is probability P of computing h = O(P 1,...,P n), is an attack on E using K assuming I giving O with probability P.

Attacks
The following attacks were investigated for this research as we examined the popular attacks oriented towards three major targets

Text-only cypher attack: Where F is constant in mathematical notation, the attack is ciphertext alone. Decrypting ciphertext when all that is available is the cipher text itself is an effort (i.e., there is no known plaintext or key connected with the ciphertext). Given simply some information G(EK(P1),...,E K(Pn)) about n cipher texts, the attacker must have some likelihood of obtaining some knowledge H(P 1,...,P n) about the plain texts.

Known plaintext attack: The perpetrator is aware of or is able to infer the plaintext for some sections of the encryption text. The remaining cipher text blocks must be decrypted.utilising this knowledge. This could be accomplished by learning the encryption key used on the data or using a shortcut. The known plaintext attack has the mathematical formula F(P_1,P_2) = With respect to P 1, G(C_1,C_2) equals (C_1,C_2), and H(P_1,P_2) only depends on P_2. In other words, the known plaintext attack is given two cipher texts C_1 and C_2, and one decryption P_1.should reveal details regarding the other P_2 decryption.

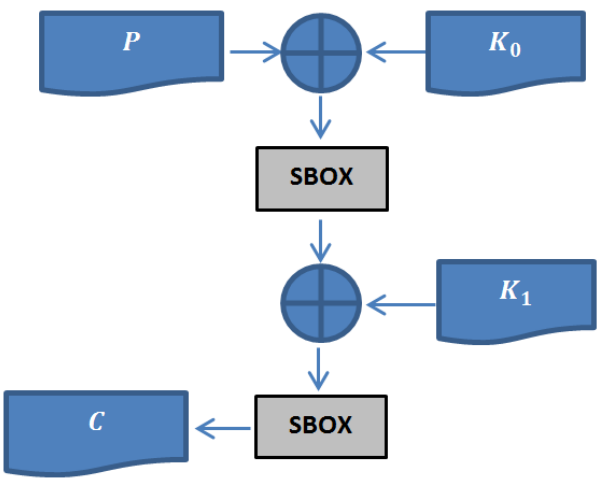
Chosen plaintext attack: Finding the encryption key is the goal of the selected plaintext attack.

Differential cryptanalysis

Differential cryptanalysis is categorized under the chosen plaintext attack on secret-key block ciphers that depends on regularly iterating a cryptographically unsound function, For example, the 16-round Data Encryption Standard (DES), to encrypt data. It is shown that the effectiveness of such attacks on an r-round cypher depends on the presence of (r-1)-round differentials with high probabilities, where an i-round differential is defined as a couple (alpha, beta) such that a pair of different plaintexts with difference can produce a pair of i-th round outputs with difference, for an appropriate notion of "difference". When an r-round cypher is not susceptible to such assaults, the probabilities of such differentials can be utilized to establish a lower bound on the complexity of a differential cryptanalysis attack. Due to the importance of "Markov cyphers" in differential cryptanalysis, the term is introduced for iterated cyphers.

The sequence of differences at each round output produces a Markov chain when the round subkeys of an iterated cypher are independent and the cypher is Markov. It is shown that the Proposed Encryption Standard (PES) of Lai and Massey, an 8-round iterated cipher, and the PES mini-version with block lengths of 8, 16, and 32 bits are Markov ciphers for the correct definition of "difference". The investigators named Biham and Shamir reached the conclusion that DES is a Markov cipher. We can see PES(8) and PES(16) are resistant to differential cryptanalysis after a sizable number of rounds. The likelihood of the most likely 7-round disparity is roughly 258 according to a thorough cryptanalysis of the full-size PES.

It is demonstrated that all 264 feasible encryptions are necessary for this type of cryptanalysis attack on PES(64). His decryption of PES led to the suggestion that the transition probability matrices of Markov ciphers should not be symmetric. This new design criterion is satisfied by a slight update to PES that nevertheless adheres to all of the original design concepts. It is detailed and demonstrated that this updated encryption, known as Improved PES (IPES), has a strong resistance to cryptanalysis type using differential equations.



Their mathematical properties very much depend upon the structures of the S-boxes used while doing encryption, so the attacker analyzes following differentials (Δx, Δy)

$\Delta y = S(x \oplus \Delta x) \oplus S(x)$

where ⊕ denotes exclusive OR for each such S-box *S*

Algebraic Cryptanalysis

Algebraic cryptanalysis is the process of solving the code through solving polynomial equations. It is based on two steps

- Conversion to system of polynomial equations
- Getting solution for secret key of the cipher by solving the equations

Many use cases have unique criteria for symmetric key ciphers with simple algebraic structures, such as lightweight implementation, efficient masking countermeasures, and low multiplicative complexity.

Minimizing the logic area of the circuit is one technique to achieve lightweight implementation. To save space, round functions with a low algebraic degree and sparse algebraic representation are chosen. The stream cipher Trivium is an excellent example. Another advantage of low-degree and sparse primitives is that they reduce the expense of implementing side-channel attack countermeasures.

Algebraic attacks are basically to infer secret keys by solving nonlinear equations involving message, cipher text and key bit.

In this attack conversion is done once. And for solving the equation most frequent methods are linearization and Gröbner basis algorithms.

Higher Order Differential and Cube Attacks deal with multivariable polynomial ciphers. However, for real-world ciphers, the polynomial can be exceedingly difficult. Consider a reduced form of the polynomial by summing over several subsets of the input space. As a result, the simplified polynomial is much easier to manage than the original complex polynomial.

First we introduce the first and higher order derivative of Boolean functions. Let $F(x)$ be a function from $\{0,1\}^n$ into $\{0,1\}^m$. For any $v \in \{0,1\}^n$, the *derivative* of F with respect to v is the function

$$D_v F(x) = F(x) \oplus F(x \oplus v).$$

For any k-dimensional subspace $V \subset \{0,1\}^n$ and for any basis $\{v_0, \cdots, v_{k-1}\}$ of V, the *k-th order derivative* of F with respect to V is the function defined by

$$D_V F(x) = D_{v_0} D_{v_1} \cdots D_{v_{k-1}} F(x) = \bigoplus_{u \in V} F(x \oplus u), \ \forall x \in \{0,1\}^n.$$

We give a small example. Let $f(x_0, x_1, x_2, x_3) = x_0 \oplus x_1 \oplus x_1 x_2 \oplus x_2 x_3$. Then $D_{(1,0,0,0)} f(x) = 1, D_{(0,0,1,0)} f(x) = x_1 \oplus x_3$, and $D_{(1,0,0,0)} D_{(0,0,1,0)} f(x) = 0$. In this example, the algebraic degree of a derivative is smaller than the original function. This is not a coincidence. Actually, it can be shown that the degree of any first-order derivative of a function is strictly less than the degree of the function. This implies that for every subspace V of dimension equals to or larger than $\deg(F)$, we have

$$D_V F(x) = \begin{cases} \text{constant}, & \text{if } \dim(V) = \deg(F); \\ 0, & \text{if } \dim(V) > \deg(F), \end{cases}$$

holds for all $x \in \{0,1\}^n$. The above property is called the *higher order differential property*, which can be used to obtain information on the message or secret key bits.

In above example $D_{(0,0,1,0)} f(x) = x_1 \oplus x_3$ derivative is linear

This kind of attack considers polynomials in which perimeters are under analysis. In it we try to retrieve some coefficients of the polynomial to get the information on the secrets.

Assume a symmetric key primitive consists of an unknown key K, a plaintext/ciphertext pair (X, Y), and an intermediate bit z. Polynomials considered are:

- $y = F(K, x)$ representing the encryption process;
- $z = F_1(K, x)$ from the encryption direction; and/or
- $z = F_2(K, y)$ from the decryption side.

With a large enough number of known or chosen plaintext/ciphertext pairs; one can use polynomial interpolation or linear equations to reconstruct some unknown coefficients of the target polynomials. It is worth noting that the coefficients are key-dependent. The secret key can then be recovered using a variety of strategies.

Linear Cryptanalysis

In a well-known plaintext assault termed linear cryptanalysis, the attacker looks at probabilistic linear relationships (also known as linear approximations) between the parity bits of the plaintext, the cipher text, and the secret key. There are two steps in linear cryptanalysis:

- Construct a linear equation relating plaintext, cipher text and key bits having a large bias, this is whose possibilities of staying are near 0 0r 1 or are relevant to them
- Derive key bits using linear equations in combination with the already known plaintext-ciphertext pairs

For construction of linear equations, an equality expression with 2 binary variables performing XOR function.

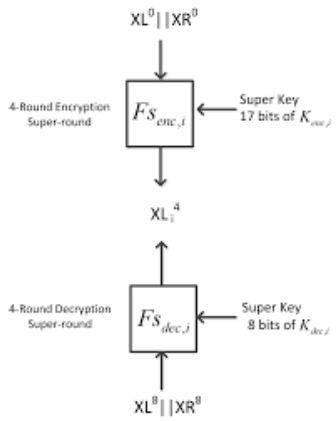
Ideal equation will of form:

$$\mathrm{P}[\alpha 1, \alpha 2 \ldots \alpha \mathrm{a}] \oplus \mathrm{C}[\beta 1, \beta 2 \ldots \beta \mathrm{b}] = \mathrm{K}[\gamma 1, \gamma 2 \ldots \gamma \mathrm{c}]$$

(Where x = 0 or 1; 1≤ a, b≤ n, 1 ≤ c ≤ m, and where the α, β and γ are fixed, specific bit locations) that influence with probability p ≠ 0.5.

Probability is directly proportional to effectiveness of the equation. Because a prospective institution is determined, the method is to evaluate the consequences of the left hand face of the previous equation for an excessive number of plaintext-ciphertext pairs.

For more than half time if answer is 0, then let K [γ1, γ2... γc] = 0, whereas if answer is 1 then K [γ1, γ2... γc] = 1.This gives us a linear equation on the important thing bits.



In a really perfect cipher, any linear equation touching on plaintext, cipher text and key bits could keep with possibility half of. Since the equations treated in linear cryptanalysis will range in probability,

they're greater as they should be known as linear approximations.

Every cipher has a unique process for creating approximations. The S-boxes, the most basic nonlinear component of the substitution-permutation community, the most common type of block cypher (the action of an S-box cannot be represented in a linear equation), are the subject of analysis in this type of block cypher. For small enough S-boxes, it is feasible to list all feasible linear equations affecting the S-input boxes and output bits, calculate their biases, and select the best ones. The remaining operations of the cypher, such as permutation and key blending, must then be added to the linear approximations for S-boxes to produce linear approximations for the entire cypher. This combining step benefits from the piling-up lemma. Additionally, there are methods for improving linear approximations iteratively.

Keep track of how frequently the approximation holds true over all recognised plaintext-ciphertext combinations for each set of values for the key bits on the right side (referred to as a partial key); call this number T. The most likely set of values for the one key bit is the partial key whose T has the best absolute difference from half of the range of plaintext-ciphertext pairs. This is so that the approximation will be preserved with a high bias, which is still believed to be the case. The importance of the bias is full-size right here, rather than the value of the opportunity itself.

Up until the variety of unknown key bits is sufficiently low, this method can be repeated to obtain guesses about the values of the key bits. At that point, brute force can be used to attack the key bits. Determine how many times the approximation holds true across all regarded plaintext-ciphertext pairings for each set of values for the critical bits on the right-hand side (referred to as a partial key); label this question T.

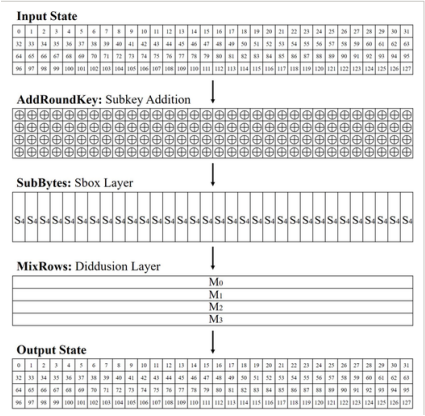
The greatest likely set of values for the one key bit is given as the partial key whose T has the best absolute difference from half of the possible pairs of plaintext and ciphertext. This is because it's usually believed that using the right partial key will result in an excessive bias being preserved in the approximation. In this case, the prejudice is much more significant than the actual probability's size.

This approach can be repeated with different linear approximations, obtaining guesses at values of key bits, until the quantity of unknown key bits is low enough that they may be attacked with brute pressure.

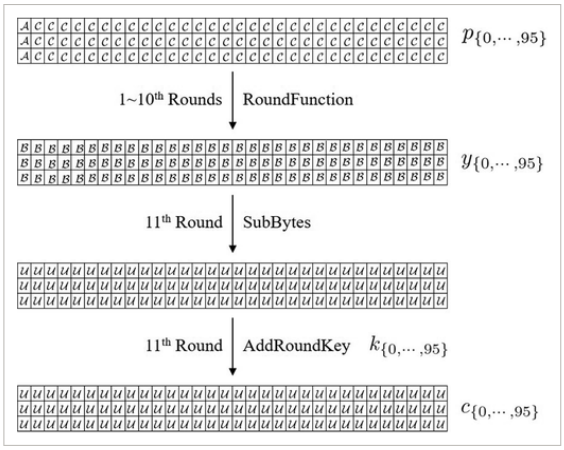
Approach by us

Integral cryptanalysis is a technique for evaluating the security of block ciphers. However, when using the MILP-aided division property to find the integral distinguishers, the calculations become unacceptable in practice since so many initial division property candidates need to be verified. This study makes use of the division property propagation of the S-box to improve and further simplify the optimal integral distinguisher searching algorithm.

The enhanced technique is then applied to provide 8- and 9-round integral distinguishers for uBlock-128 and uBlock-256, as well as 10- and 9-round integral distinguishers for Pyjamask-96 and Pyjamask-128. The authors demonstrate 9- and 11-round key-recovery attacks on uBlock-128 and Pyjamask-96, respectively, using this justification and the partial sums technique. The data complexity is 2 power 143 and 2 power 93.



11-round integral attack



\

References

https://books.google.com/books?hl=en&lr=&id=rWdUPSuLFwAC&oi=fnd&pg=PR9&dq=elementary+cryptanalysis&ots=c-lljvLFgZ&sig=OQ7TnqTt9KR5NA-HvPPjZtvXIcU
https://www.tandfonline.com/doi/pdf/10.1080/00029890.1961.11989689
https://www.tandfonline.com/doi/abs/10.1080/0161-110291890885
https://link.springer.com/chapter/10.1007/978-3-540-77272-9_10
https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/iet-ifs.2019.0624
https://books.google.com.pk/books?hl=en&lr=&id=Zb2RBQAAQBAJ&oi=fnd&pg=PP1&dq=cryptanalysis+and+its+types&ots=ydXkiqj_XZ&sig=6P-dCO-EQBNBw7
http://www.csshl.net/sites/default/files/downloadable/crypto/TEA_Cryptanalysis_-_VRAndem.pdf
https://www.tutorialspoint.com/what-is-linear-cryptanalysis-in-information -security
https://www.esat.kuleuven.be/cosic/blog/algebraic-cryptanalysis