BOOMERANG ATTACKS ON KATAN REVISITED

LI YU

UNIVERSITI SAINS MALAYSIA

2022

BOOMERANG ATTACKS ON KATAN REVISITED

by

LI YU

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

January 2022

DECLARATION

Name: LI YU

Matric No: P-COM0067/21

School: School of Computer Sciences

Thesis Title: Boomerang Attacks on KATAN Revisited

I hereby declare that this thesis I have submitted to .... on .... is my own work. I have stated all references used for the completion of my thesis.

I agree to prepare electronic copies of the said thesis to the external examiner or internal examiner for the determination of amount of words used or to check on plagiarism should a request be made.

I make this declaration with the believe that what is stated in this declaration is true and the thesis as forwarded is free from plagiarism as provided under Rule 6 of the Universities and University Colleges (Amendment) Act 2008, University Science Malaysia Rules (Student Discipline) 1999.

I conscientiously believe and agree that the University can take disciplinary actions against me under Rule 48 of the Act should my thesis be found to be the work or ideas of other persons.

Students Signature: .... Date: ....

Acknowledgement of receipt by: .... Date: ....

ACKNOWLEDGEMENT

This thesis would not have been a success without ......

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION

CHAPTER 2 – LITERATURE REVIEW

CHAPTER 3 – METHODOLOGY

CHAPTER 4 – RESULTS AND DISCUSSION

CHAPTER 5 – CONCLUSION AND FUTURE WORK

APPENDICES

LIST OF PUBLICATIONS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

3DES	Triple Data Encryption Standard
3GPP	3rd Generation Partnership Project
AE	Authenticated Encryption
AEAD	Authenticated Encryption with Associated Data
AES	Advanced Encryption Standard
ARX	Addition Rotation Xor
AS	Active S-box
ASE	Average Shannon Entropy
CBC	Cipher Block Chaining
CCML	Coupled Chaotic Map Lattice
CFB	Cipher Feedback
CPU	Central Processing Unit
CTR	Counter
DES	Data Encryption Standard
DDR	Data-Dependent Rotation
ECB	Electronic Codebook
FFT	Fast Fourier Transform
FIPS	Federal Information Processing
FxP	Fixed Point
FSM	Finite State Machine
GCC	GNU Compiler Collection
GCM	Galois Counter Mode
GFS	Generalized Feistel Structure
GK	Generator Kernel
GOP	Gray Ordering Number
GPU	Graphics Processing Unit
GSM	Global System for Mobile Communications
HDD	Hard Disk Drive

IEC	International Electrotechnical Commission
ISO	International Organization for Standardization
IV	Initialization Vector
KFB	Key Feedback
LCG	Linear Congruential Generator
LCM	Least Common Multiple
LFSR	Linear Feedback Shift Register
LSB	Least Significant Bit
MAC	Message Authentication Code
Mbps	Megabits Per Second
MDS	Maximum Distance Separable
MitM	Meet-in-the-Middle
MSB	Most Significant Bit
MLD	Maximum-Likelihood Decoding
MQ	Multivariate Equation
NESSIE	New European Schemes for Signatures, Integrity and Encryption
NIST	National Institute of Standards and Technology
NLF	Nonlinear Filter

LIST OF SYMBOLS

$|x|$ Bit length of x.

$\oplus$ Bitwise XOR.

$\wedge$ Bitwise AND.

$x||y$ Bitwise concatenation of x and y.

$\boxplus$ Modular addition.

$(x << y)$ / $(x >> y)$ Bitwise left/right shift of x by y bits.

$\lfloor x \rfloor$ Floor function to map a real number x to the previous largest integer.

SERANGAN BOOMERANG KE ATAS KATAN DILAWATI SEMULA

ABSTRAK

Abstrak Bahasa Malaysia ditulis di sini.

BOOMERANG ATTACKS ON KATAN REVISITED

ABSTRACT

English abstract here.

CHAPTER 1

INTRODUCTION

1.1 Background

The boomerang attack is currently the most popular cryptanalysis method, it is based on the differential method, and was proposed by Wagner (1999). It is a chosen plaintext attack and is used to analysis symmetric cryptography, compare to differential cryptanalysis, which provides a more efficient method to analyse more complex ciphers. Here is a brief introduction to the principle. Figure 1.1 shows the design of basic boomerang analysis. The boomerang attack attempts to generate a quartet structure at an intermediate value halfway through the cipher.

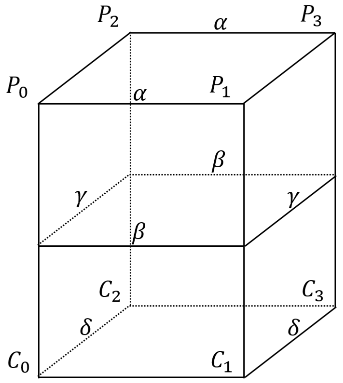


Figure 1.1: Boomerang Attack Model

The boomerang attack is widely used due to its power. However, Murphy (2011) raised concerns about the validity of the boomerang attack results that not all S-BOX ciphers boomerang distinguishers are reliable, for S-BOX based ciphers, two independently chosen differential trails can be incompatible, thus the probability of finding a right quartet can be zero and the same phenomenon was observed by Biryukov, De Cannière, and Dellkrantz (2003) as the middle round S-box trick. In order to solve this problem, Cid, Huang, Peyrin, Sasaki, and Song (2018) proposed a new cryptanalysis tool called Boomerang Connectivity Table (BTC). Compared to the Difference

Distributed Table (DDT), BTC can find better differential trails, the table [1] from the paper demonstrates the advantage. BTC has better efficiency than DDT, but it has only been used in practice on the S-BOX cipher so far, Cid et al. (2018) point out BTC may also be valid for modular addition ciphers.

With the development of IoT, more and more low-end devices are used, such as Smart Locks, there are large holdings of these devices, leading to an increasing need to provide security. Because these devices have the low computing power and run in a complex environment, several lightweight block ciphers were proposed. The KATAN family is based on modular addition, not S-BOX and was proposed by De Cannière, Dunkelman, and Knežević (2009). Although it is subjected to many different types of analysis, it still provides security. The KATAN family contains six ciphers divide into two flavors, and all block ciphers share the 80-bits key size. KATAN is composed of three block ciphers, with 32,48, or 64-bits block size. Chen, Teh, Su, Samsudin, and Fang (2016) use an extended boomerang framework in related-key setting, to achieve the best results by far for KATAN48/64 in differential setting.

1.2 Problem Statement

As described in the previous section, the cryptanalysis based on boomerang has achieved good results Chen et al. (2016) on the KATAN cipher so far, but as Murphy (2011) point out, distinguishers based on boomerang analysis are not necessarily reliable. In other words, BTC Cid et al. (2018) can achieve a better result on S-BOX ciphers. So, there are some questions that need to be discovered:

1. Whether BTC can apply non-S-BOX ciphers?
2. How to improve the reliability of the result of the boomerang analysis of KATAN?

Then, the following research questions will be answered in this work:

1. What strategies can make BTC apply to ciphers based on modular operation, such as KATAN?
2. How can achieve better results of boomerang analysis of KATAN than previous research using BTC?

1.3 Research Motivation

With the rise of edge computing and the Internet of things, many devices need to save data locally, which also leads to high security for such devices. To improve the security of these devices without reducing performance, several lightweight cipers were proposed. KATAN is one example of them. It has withstood various cryptanalysis since it was proposed and still provides high security. This work uses it as a target cipher because it widely used in RFID devices, which represent a huge commercial value, and its analysis can further guarantee the security of the assets it protects. On the other hand, KATAN is a cipher based on modular addition, which can help this work to verify the adaptation of BTC on non-S-BOX ciphers.

This work chose the boomerang attack as the cryptanalysis method because the efficiency of the boomerang attack was demonstrated in several research projects. This helped this work reduce the likelihood of failure, and due to the boomerang attack has become a research hotspot, many strategies which improve the efficiency of the attack were proposed. In conclusion, this work can achieve the better result by using the boomerang attack.

1.4 Research Scope and Objectives

1.4.1 Research Scope

This study consists of two parts. First, in cryptography, this work focus on block ciphers, which are symmetric-key cryptosystem. Second, in cryptanalysis, this work focus on boomerang attack, which is an adaption of differential cryptanalysis. In short, the scope of this study is that boomerang attack on non-S-BOX block ciphers.

1.4.2 Research Objectives

According to this study research questions, the objectives of this research are summarized as follows:

1. BTC be successfully applied on KATAN ciphers.
2. To improve the result of boomerang attack on KATAN ciphers by using BTC.

Two objectives should be enough, and each objective should be measurable.

1.5 Research Methodology

This study has six steps, are shown in Figure 1.2.

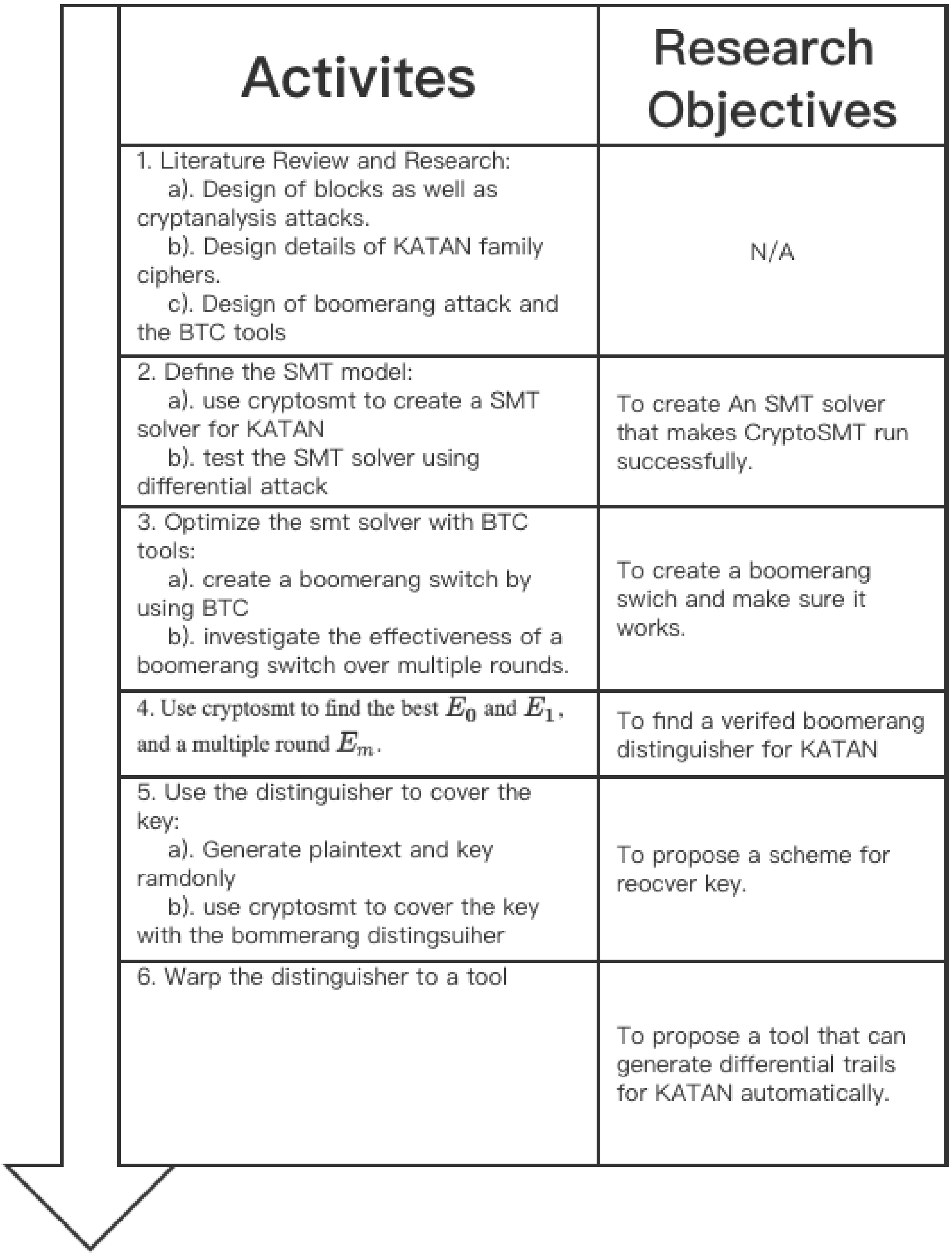
1.6 Research Contributions

This study gets some contributions to cryptanalysis.

First, Cid et al. (2018) point out the BTC tools may be applied on ciphers based on modular addition operation, but it is not practiced. This study successfully demonstrates the BTC tools are not only used on ciphers based on S-BOX and useful to ciphers based on modular addition operations. The contribution can help cryptanalysts to improve their distinguisher of attack to ciphers based on modular operations.

This work also improves the probability of the boomerang distinguisher of KATAN. For KATAN48/64, this work was able to greatly improve upon the previous results to achieve the best result by for. The rules reflected in this distinguisher can help designers design ciphers better.

Last, this work proposed a tool that can generate distinguishers automatically,



Activites	Research Objectives
1. Literature Review and Research: a). Design of blocks as well as cryptanalysis attacks. b). Design details of KATAN family ciphers. c). Design of boomerang attack and the BTC tools	N/A
2. Define the SMT model: a). use cryptosmt to create a SMT solver for KATAN b). test the SMT solver using differential attack	To create An SMT solver that makes CryptoSMT run successfully.
3. Optimize the smt solver with BTC tools: a). create a boomerang switch by using BTC b). investigate the effectiveness of a boomerang switch over multiple rounds.	To create a boomerang swich and make sure it works.
4. Use cryptosmt to find the best E_0 and E_1, and a multiple round E_m.	To find a verifed boomerang distinguisher for KATAN
5. Use the distinguisher to cover the key: a). Generate plaintext and key ramdonly b). use cryptosmt to cover the key with the bommerang distingsuiher	To propose a scheme for reocver key.
6. Warp the distinguisher to a tool	To propose a tool that can generate differential trails for KATAN automatically.

Figure 1.2: Research Framework

which can help improve the efficiency of attacking KATAN. Then summarize the contributions of this work as follows:

- A new strategy that can use BTC on ciphers based on modular addition operation.

- Improve the probability of the distinguisher for KATAN.

- A new cryptanalysis tool that can generate distinguishers for KATAN automatically.

1.7 Thesis Outline

Chapter 1 has provided an overview of the research including the introduction for KATAN and boomerang attack, and what is this study going to solve.

Next, Chapter 2 provides an in-depth review of prior work in the fields including block ciphers, KATAN, differential attack, boomerang attack and BTC.

Chapter 3 discusses the methodology involved in this research, we demonstrate the boomerang distinguisher search and key recovery attack on the KATAN family by using BTC. This is followed by Chapter 4 which analyses the findings and results, the BTC tools improve the probability of finding the right quartet.

Finally, this thesis concludes in Chapter 6 with a summary of finding and future works.

CHAPTER 2
LITERATURE REVIEW

2.1 Overview

The research analyzes previous cryptanalysis work, including the design of block ciphers, cryptanalysis-related knowledge, including KATAN and Boomerang related literature needed for the research.

2.2 Block ciphers and Cryptanalysis

This section describes the principle of symmetric-key encryption and introduces two symmetric-key ciphers. Then there is a review of cryptanalysis methods, and the details of these methods are described. This section provides help for understanding cryptanalysis.

2.2.1 Block Cipher Designs

Block cipher is an important part of symmetric-key encryption. The design concept of block cipher originated from Shannon (1949). Its public research began with the publication of the DES algorithm Pub (1999) in the late 1970s. The rapid development of block cipher theory and application Benefited from the AES Dworkin et al. (2001) program in the United States in the late 1990s and the NESSIE program in Europe in the early 2000s.

In the literature Shannon (1949), from the perspective of resisting statistical attacks, Shannon proposed the ”confusion” and ”diffusion” criteria for designing encryption algorithms. Each element in the original text is rearranged in obfuscation, whereas in diffusion, each element is changed to be mapped to numerous elements in the ciphertext. This criterion is still one of the important principles to be followed in the design of block ciphers. In the AES plan and the NESSIE plan, the cryptography community has conducted extensive and in-depth research on the design and analysis theory of block ciphers, and the theory of block ciphers is perfect. In the SHA-3 project Dworkin et al. (2015), more than half of the Hash functions adopt the design concept of block ciphers, so block ciphers are becoming more and more important.

The design of block ciphers usually follows the following two principles: the security principle and the implementation principle. The security principles include the principle of confusion, the principle of proliferation, and the principle of resistance to existing attacks. Implementation principles include software implementation principles and hardware implementation principles. Usually, the designed algorithm conforms to the above principles by iterative means: one method is to construct an iterative function with strong cryptographic properties, so that the number of iterations can be reduced; the other method is to construct an iterative function with relatively weak cryptographic properties, but The number of iterations is relatively high. In practical construction, the latter is usually adopted, that is, functions with weak cryptographic properties are iterated for many times to satisfy the security principle and the realization principle.

The mathematic model of block cipher, $\mathbb{F}_2$ presents a binary field, $\mathbb{F}_2^n$ and $\mathbb{F}_2^m$ present n and m dimensional vector spaces, respectively. $S_K \in \mathbb{F}_2^m$ is a key space, then the ciphers can present two reflections.

$$\begin{aligned} E &: F_2^n \times S_k \to F_2^n, \\ D &: F_2^n \times S_k \to F_2^n. \end{aligned} \tag{2.1}$$

The basic structure of commonly used block ciphers will be described below.

2.2.1(a) Feistel Structure

The IBM researchers that developed the DES algorithm also came up with the Feistel structure, which eventually gained popularity. Referring to Figure 2.1, the encryption procedure for the r-round Feistel structure cipher with a $2n$ block length is as follows: Given a civilization P with $2n$ bits, divide it into two left and right n bits, where L_0 is the left n bits and R_0 is the right n bits, and then, in accordance with (2.2), repeat the procedure r times.

$$\begin{cases} L_i = R_{i-1} \\ R_i = L_{i-1} \oplus F(R_{i-1}, K_i) \end{cases} \quad i = 1, 2, \ldots, r \tag{2.2}$$

Here, $\oplus$ denotes an XOR operation, $F : \mathbb{F}_2^n \times \mathbb{F}_2^m \to \mathbb{F}_2^n$ stands for "round," and $K_1, K_2, \ldots, K_r$ According to the key expansion method, K_r is the round key produced by the seed key K, and m is the round key's length. The final round typically does not require "left and right exchange," meaning that the ciphertext is $C = R_r L_r$. The slower diffusion effect is a drawback of Feistel.

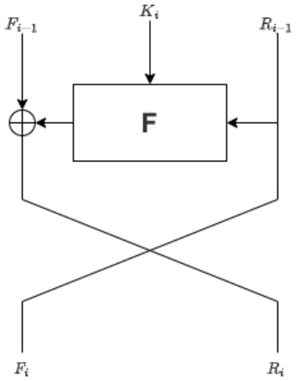


Figure 2.1: Feistel Structure

2.2.1(b) SPN Structure

A reversible linear transformation P and a reversible nonlinear function S, controlled by a round key, typically make up each round of the SPN structure. The S-transform layer serves as the obfuscation layer in this construction, and the P-transform layer serves as the diffusion layer. The SPN structure has a faster diffusion effect than the Feistel structure, and the designer can use this structure to provide algorithms resistant to differential cryptanalysis and linear ciphers with a provable security for analysis.

$$\begin{cases} Y &= S(X_{i-1}, K_i) \\ X_i &= P(Y) \end{cases} i = 1, 2, \ldots, r \tag{2.3}$$

The encryption procedure for the r-round SPN structure cipher with an n-block length is as follows: Let $P = X_0$, where P is the plaintext with n bits, and follow (2.3) to carry out the same procedure for r rounds.

In SPN structured ciphers, the last round of P transformation is usually replaced by key addition.

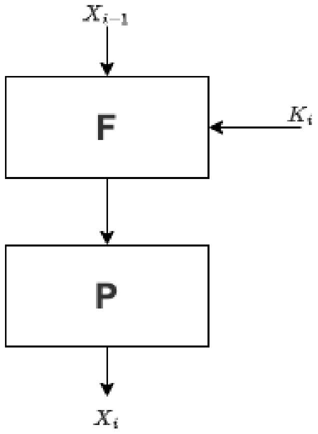


Figure 2.2: SPN Structure

2.2.1(c) Lai-Massey Structure

When developing the IDEA algorithm, Lai and Massey proposed a framework they called the Lai-Massey structure. Most of the time, the Lai-Massey structure has the benefit of consistent encryption and decryption.

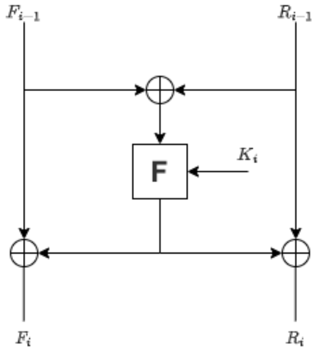


Figure 2.3: Lai-Massey Structure

For the r-round Lai-Massey structure cipher with a block length of 2n, the encryption process is as follows: Given a 2n-bit plaintext P, first divide it into two n-bit parts on the left and right, and denote L_0 as the left n bits of P, and R_0 as the right n bits of P, then $P = L_0R_0$. Then according to (2.4), r rounds of exactly the same operations are performed.

$$\begin{cases} T & = F(L_{i-1} \oplus R_{i-1}, K_i) \\ L_i & = L_{i-1} \oplus T \qquad\qquad i = 1, 2, \ldots, r \\ R_i & = R_{i-1} \oplus T \end{cases} \tag{2.4}$$

2.2.1(d) Summary

The overall structure is an important feature of block cipher algorithms. Different structures have a great impact on the selection of the round function and the performance on various platforms. In addition to the above three mainstream structures, the overall structure also includes generalized (non-)equilibrium Feistel structure, MISTY structure and the mixed use of various structures. In addition, the round functions of many cryptographic algorithms adopt different structures. For example, the Camellia algorithm adopts the Feistel structure as a whole, but the round function adopts the SPN structure; the FOX algorithm adopts the Lai-Massey structure as a whole, and the round function adopts the SPS structure. The structure used to design an algorithm mainly depends on the performance requirements of the algorithm, the construction of sub-modules, and the security of the overall structure.

2.2.2 KATAN family cipher

KATAN family cipher is a hardware oriented block ciphers and was proposed in CHES 2009 De Cannière et al. (2009). KATAN family has three variants of KATAN are KATAN32, KATAN48 and KATAN64. Each various contains six block ciphers divide into two flavors, in the first flavor is composed of three block ciphers, with 32,48 or 64 bit block size, and in the second flavor contains the other three ciphers with same block size. They are use 80-bit key size. KATAN use an extending key algorithm to make 80-bit key to 508bit sub-key. Suppose a key k is 80-bit, and k_i present i-th bit in K, the sub-key is given by:

$$sk_i = \begin{cases} k_i & for \quad i = 0\ldots 79 \\ k_{i-80} \oplus k_{i-61} \oplus k_{i-50} \oplus k_{i-13} & Otherwise \end{cases} \tag{2.5}$$

The round function of KATAN, the plaintext be divided two part and be loaded to two register L_1 and L_2, then the update processes are shown as follows:

$$\begin{aligned} f_a(L_1) &= L_1[x_1] \oplus L_1[x_2] \oplus (L_1[x_3]) \cdot L_1[x_4]) \oplus (L_1[x_5] \cdot IR) \oplus k_a \\ f_b(L_2) &= L_2[y_1] \oplus L_2[y_2] \oplus (L_2[y_3]) \cdot L_2[y_4]) \oplus (L_2[y_5] \cdot L_2[y_6]) \oplus k_b \\ L_1[i] &= L_1[i-1](i \leq i \leq |L_1|), L_1[0] = f_b(L_2), \\ L_2[i] &= L_2[i-1](i \leq i \leq |L_2|), L_2[0] = f_b(L_1), \end{aligned} \tag{2.6}$$

where $\oplus$ and $\cdot$ are bitwise XOR and AND operations, respectively, and $L[x]$ denotes the xth bit of L, IR is the round constant value defined in the specification, and ka and kb are two subkey bits. For round i, k_a and k_b correspond to $sk_{2(i-1)}$ and $sk_{2(i-1)+1}$. The parameters of KATAN family are shown in TABLE 2.1.

Table 2.1: Parameters of KATAN family

algorithm	$\|L_1\|$	$\|L_2\|$	x_1	x_2	x_3	x_4	x_5	y_1	y_2	y_3	y_4	y_5	y_6
KATAN32	13	19	12	7	8	5	3	18	7	12	10	8	3
KATAN48	19	29	18	12	15	7	6	28	19	21	13	15	6
KATAN64	25	39	24	15	20	11	9	38	25	33	21	14	9

2.2.3 Cryptanalysis

There are two basic ways to measure the security of a cryptographic algorithm: one is actual security, and the other is unconditional security (theoretical security). The actual security is evaluated according to the amount of computation required to decipher the cryptosystem. For example, the security of the RSA system is based on the difficulty of decomposing large integers, but with the development of computer technology, RSA may also be vulnerable. Theoretical security is independent of an adversary's computing power and time, and any effort to decipher an algorithm will be no better than random selection.

Most of the cryptanalysis methods proposed in the papers follow Kerckhoffs's principle: The principle holds that a cryptosystem should be secure, even if everything about the system, except the key, is public knowledge. According to the Kerckhoffs's principle, the security of a cryptographic algorithm should depend on the secrecy of the key, not the secrecy of the algorithm itself.

Many literatures propose different cryptanalysis methods to analyze the target algorithms they specify. What is important is whether these algorithms have extremely high accuracy, whether they undermine practical security or theoretical security, in which threats to practical security are very serious issues. In the following, we provide an overview of the cryptanalysis of block ciphers.

The method of cryptanalysis is divided into brute force attack, the security of the algorithm is studied based on the mathematical method, and the security of the algorithm and the security of the algorithm under different usage modes are studied in combination with the physical realization method.

According to the different environments, password attacks can be divided into the following four types:

- Ciphertext-only attack: A cryptanalyst has one or more ciphertexts encrypted with the same key, and analyzes these decrypted ciphertexts to obtain the plaintext or key.
- Known-Plaintext Attack: A cryptanalyst has some plaintexts and ciphertexts of these plaintexts encrypted with the same key, and recovers the key by analyzing these known plaintexts and the corresponding ciphertexts.
- Chosen-plaintext attack: The cryptanalyst can choose the plaintext he wants at will and encrypt it, and recover the key according to the selected plaintext and the corresponding ciphertext.
- Chosen ciphertext attack: The cryptanalyst can freely select the ciphertext he wants and decrypt it, and recover the key according to the chosen ciphertext and the corresponding plaintext.

Then, the research on analytical ciphers is usually based on three different foundations.

2.2.3(a) Brute-force

Brute-force attacks are the simplest and least effective, they generally include four methods:

- Exhaustive key search: Under a ciphertext-only attack, the attacker uses all possible keys to continue decrypting one or more ciphertexts until a meaningful plaintext is obtained. This method can decipher any block cipher in theory, but its efficiency is the lowest. In practical cryptanalysis, it is usually used in combination with other analysis methods.
- Dictionary attack: The attacker collects pairs of plaintext and ciphertext and arranges them into a "dictionary". When seeing a ciphertext, the attacker checks whether it exists in the dictionary, and if so, finds the corresponding plaintext.
- Table lookup attack: This method is a chosen plaintext attack. The attacker uses all possible keys to encrypt the same plaintext, and stores the key and the corresponding ciphertext. When obtaining the plaintext and ciphertext, the attacker only needs to Find the corresponding key in the storage table.
- Time-tradeoff attack: This is a chosen-plaintext attack method, proposed by Hellmma, by using a combination of exhaustive key search and table lookup attack.

Although these methods are not very useful in practice, the complexity of exhaustive key search is used as an indicator to measure the efficiency of an attacking algorithm.

2.2.3(b) Based on the Security of Algorithms Based on Mathematical Methods

This method mainly includes two aspects: one is to study how to distinguish cryptographic algorithms from random permutation. In cryptanalysis, for the same indicator, first calculate its value under random permutation, and then calculate the corresponding value of a cryptographic algorithm in a certain cryptographic algorithm.

If the two values are significantly different, then this metric can distinguish a cryptographic algorithm from random permutation. In cryptanalysis, if for some specific form of plaintext input, the corresponding ciphertext follows a special rule, it is said to find an effective distinguisher of the algorithm. The second is to study how to obtain the key information of the cryptographic algorithm. For iterative block ciphers, the cryptanalyst first finds an efficient distinguisher of the reduced-rounds algorithm, and then verifies the correctness of the distinguisher by guessing some of the round keys. When an effective distinguisher of the cryptographic algorithm is found, there are usually two ways to recover the secret: one is the statistical method, for the guess value of each key, according to certain rules (usually related to the distinguisher) collected plain ciphertext Statistical analysis is performed, and the final value with a clear statistical advantage may be the correct key. The another is to solve by algebraic method. In this method, the transformation corresponding to encryption and decryption is represented by a system of equations. Through certain mathematical methods, the root of the system of equations is solved to obtain the key information. Common math-based methods include:

- differential cryptanalysis
- linear cryptanalysis
- meet-in-the-middle attack
- collision attack
- Square attack
- interpolation attack
- correlated key attack
- boomerang attack

With the development of cryptanalysis, more methods have been proposed.

2.2.3(c) The security of algorithms is related to they way it is implemented in hardware

In the traditional research on the security of algorithms based on mathematical methods, the encryption or decryption process of the algorithm is generally regarded as a transformation with secret parameters, and the key information is inferred only by obtaining the input and output of the transformation. At the end of the 20th century, a new attack method appeared in the cryptography world. In addition to the traditional mathematical method, this attack method also combines the information differences represented by certain physical parameters such as time, energy, electromagnetic, temperature, sound, etc. to infer information about the key. This attack method combined with physical implementation is generally called side channel attack. At present, common side-channel attacks mainly include timing attacks, energy analysis, fault attacks, electromagnetic attacks, and cache attacks.

2.3 Cryptanalysis Method

This section describes details of several attacks.

2.3.1 Differential Attack

The differential attack is one most effective methods to attack iterative block ciphers, and it is an important index for evaluating the security of block ciphers. Biham and Shamir (1991) proposed this analysis method. It finds the difference between specific pairs of plaintext and ciphertext to distinguish block ciphers from random permutations and then recovers the key.

We assume a block cipher, $E\{0,1\}^n \times \{0,1\}^l \rightarrow \{0,1\}^n$, and n, l present length of a block and length of the key, respectively. For $k \in \{0,1\}^l$, the permutation be presented $E_k(\cdot) = E(\cdot, k)$ on $\{0,1\}^n$. The encryption function $E_k(\cdot)$ is consist of r subencryption functions which use sub-key $k_i, i = 1,2,3,4...,r$ to encrypt. The encyrption equation as 2.7.

$$E_k(x) = F_{k_r} \circ F_{k_{r-1}} \circ \cdots \circ F_{k_2} \circ F_{k_1}(x). \tag{2.7}$$

Next, Y_{i-1}, Y_i present the input and output on i-th round, in other words, $Y_i = F_{k_i}(Y_{i-1})$. $X, Z = Y_r$ present the input(plaintext) of the first round and the output of the last round, respectively. There are three important definitions. First, **Difference**, the difference of $X, X^* \in \{0,1\}^n$ is $\Delta X = X \oplus X^*$. Second, **Differential**, presents a propagation from difference α to difference β. Last, **Differential characteristic**, a differential characteristic of i rounds, $\Omega = (\beta_0, \beta_1 \ldots, \beta_{i-1}, \beta_i)$ when the difference of the input pairs (X, X^*) satisfy $X \oplus X^* = \beta_0$, and in i-th round, the middle states (Y_j, Y_j^*) satisfy $Y_j \oplus Y_j^* = \beta_j$. The equation 2.8 presents the probability of a differential,

$$DP(\alpha, \beta) = \underset{X,K}{Prob}\{F(X,K) \oplus F(X \oplus \alpha, K) = \beta\}, \tag{2.8}$$

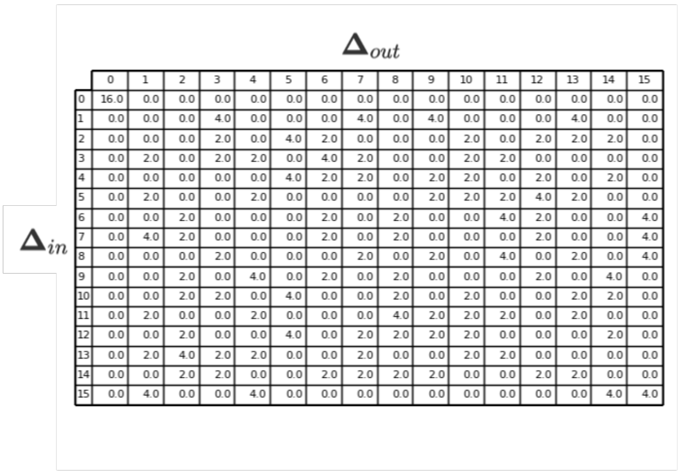
when the number of rounds more than 1 can derive the probability of a differential characteristic $\Omega = (\beta_0, \beta_1, \ldots, \beta_i)$ in following.

$$DP(\Omega) = \prod_{j=1}^{i} DP(\beta_{j-1}, \beta_j). \tag{2.9}$$

In Feistel or SPN structure ciphers, the S-box is the only nonlinear part of the round function, so we need to make **Differential Distributed Table(DDT)** if we want to find a differential path with high probability. There is an example of DDT of present cipher is Figure 2.4. The first row present difference of out, and first column present difference of input to S-BOX. Suppose $S : \{0,1\}^n \rightarrow \{0,1\}^n$ an n-bit to n-bit S-BOX,the element of DDT is given by:

$$(\Delta_{in}, \Delta_{out}) = \#\{P_i \oplus P_i^* = \Delta_{in} | S(P_i) \oplus S(P_i^*) = \Delta_{out}\} \tag{2.10}$$

The probability of a pair$(\Delta_{in}, \Delta_{out})$ is $p = \frac{pos(\Delta_{in}, \Delta_{out})}{2^n}$. The differential attack steps are


Δ_{out}
Δ_{in}

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	16.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	4.0	0.0	0.0	0.0	4.0	0.0	4.0	0.0	0.0	0.0	4.0	0.0	0.0
2	0.0	0.0	0.0	2.0	0.0	4.0	2.0	0.0	0.0	0.0	2.0	0.0	2.0	2.0	2.0	0.0
3	0.0	2.0	0.0	2.0	2.0	0.0	4.0	2.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	4.0	2.0	2.0	0.0	2.0	2.0	0.0	2.0	0.0	2.0	0.0
5	0.0	2.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	2.0	2.0	2.0	4.0	2.0	0.0	0.0
6	0.0	0.0	2.0	0.0	0.0	0.0	2.0	0.0	2.0	0.0	0.0	4.0	2.0	0.0	0.0	4.0
7	0.0	4.0	2.0	0.0	0.0	0.0	2.0	0.0	2.0	0.0	0.0	0.0	2.0	0.0	0.0	4.0
8	0.0	0.0	0.0	2.0	0.0	0.0	0.0	2.0	0.0	2.0	0.0	4.0	0.0	2.0	0.0	4.0
9	0.0	0.0	2.0	0.0	4.0	0.0	2.0	0.0	2.0	0.0	0.0	0.0	2.0	0.0	4.0	0.0
10	0.0	0.0	2.0	2.0	0.0	4.0	0.0	0.0	2.0	0.0	2.0	0.0	0.0	2.0	2.0	0.0
11	0.0	2.0	0.0	0.0	2.0	0.0	0.0	0.0	4.0	2.0	2.0	2.0	0.0	2.0	0.0	0.0
12	0.0	0.0	2.0	0.0	0.0	4.0	0.0	2.0	2.0	2.0	2.0	0.0	0.0	0.0	2.0	0.0
13	0.0	2.0	4.0	2.0	2.0	0.0	0.0	2.0	0.0	0.0	2.0	2.0	0.0	0.0	0.0	0.0
14	0.0	0.0	2.0	2.0	0.0	0.0	2.0	2.0	2.0	2.0	0.0	0.0	2.0	2.0	0.0	0.0
15	0.0	4.0	0.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	4.0

Figure 2.4: Differential Distributed Table

summarized as follows:

1. To find a high probability differential characteristic of $r-1$ rounds, the probability is presented as p.

2. To determine the sub-key in the i-th round by the output of the differential characteristic. Assume the length of the key attacked is l bits, to set 2^l counters, each counter corresponding to each possible candidate key.

3. Generate the plaintext X uniformly at random, X^* satisfy $X^* = X \oplus \alpha$. To encrypt X, X^* to get ciphertexts Z, Z^*, respectively. The number of generated plaintext pairs is $m \approx c \times \frac{1}{p}$, c presents a constant number.

4. For each ciphertext pairs, use candidate sub-key to decrypt Z, Z^*, the difference of the decrypted data $\Delta = F_{gk_i}^{-1} \oplus F_{gk_i}^{-1}(Z^*)$, if $\Delta = \beta$, then the counter corresponding to the candidate key is plus 1.

5. The max value of these counter corresponding to the candidate key is correct sub-key.

2.3.2 Linear Cryptanalysis

A Japanese researcher Matsui (1993) proposed a new cryptanalysis method in Eurocrypt 1993. The same year, in Crypto 1993, he published new two linear approximation relations, and he use these relations to crack DES, using 12 workshops and 50 days. There are results of the first experimental analysis of DES in the open literature. Unlike differential cryptanalysis, linear cryptanalysis is a known-plaintext attack method. It distinguishes block ciphers and random permutations by finding effective linear approximation relations between plaintext and ciphertext and recovers the key. Remember a block cipher defined in 2.7. In linear cryptanalysis, assume $a, b \in \{0,1\}^n$ and $a = (a_1, a_2, a_3, \ldots, a_n), b = (b_1, b_2, \ldots, b_n)$, $F(x,k)$ presents the round function, given two linear mask (α, β), the linear expression approximation formula is

$$\alpha \cdot x \oplus \beta \cdot F(x,k). \tag{2.11}$$

In this formula, $\cdot$ present multiplication in a binary field. Assume p presents the probability of the linear expression approximation formula of α and β, then the linear probability can present these:

$$\begin{aligned} Bias: \zeta_F &= p(\alpha,\beta) - \frac{1}{2}, \\ Correlation: Cor_F &= 2p(\alpha,\beta) - 1, \\ Potential: Pot_F &= (p(\alpha,\beta) - \frac{1}{2})^2. \end{aligned} \tag{2.12}$$

. In paper [28,39], if the define the linear expression approximation formula of linear mask (α, β) is

$$LP(\alpha,\beta) = (2 \cdot \underset{X,K}{Prob}\{\alpha \cdot X = \beta \cdot F(X,K)\} - 1)^2 \tag{2.13}$$

, the LP can corresponding to DP of differential cryptanalysis. The steps of linear cryptanalysis as follows:

1. To find r-1 rounds differential approximation formula that bias $zeta(\alpha,\beta)$ relatively large.

2. To determine the sub-key in the i-th round by the output of the differential characteristic. Assume the length of the key attacked is l bits, to set 2^l counters, each counter corresponding to each possible candidate key.

3. Select the plaintext X uniformly at random. To encrypt X to get ciphertext Z, respectively. The number of generated plaintext pairs is $m \approx c \times \frac{1}{\zeta^2}$, c presents a constant number.

4. For each ciphertext Z, decrypt it to get Y_{r-1} using candidate sub-keys, if $\alpha x \cdot X \oplus \beta \cdot Y_{r-1} = 0$, then the corresponding counter is plus 1.

5. The max value of $|\frac{\lambda^i}{m} - \frac{1}{2}|$ of these counter corresponding to the candidate key is correct sub-key.

2.3.3 Boomerang Attack

The boomerang attack is a differential-style attack, it can find a higher probability of differential characteristics than a differential attack and it is a more effective cryptanalysis method of block ciphers. Wagner (1999) proposed this attack upon discovering good differential characteristics for the first four and last four Feistel rounds of COCONUT98. The boomerang attack has more power when it combined with other cryptanalysis method, it has been demonstrated when it is used to break the full-round AES-192/256 Biryukov and Khovratovich (2009) and the full-round KASUMI Biham, Dunkelman, and Keller (2005) in the related-key setting.

Figure 1.1 shows the structure of the boomerang attack. Four plaintexts P_1, P_2, P_3, P_4 and with their respective ciphertexts C_0, C_1, C_2, C_3; The $E(\cdot)$ present the encryption operation and decompose the cipher into $E = E_0 \circ E_1$ where E_0 represent the first half of the cipher and E_1 presents last half. There are four differential trails, $\alpha \rightarrow \beta$ for E_0; $r \rightarrow \delta$ for E_1; $\delta \rightarrow r$ for E_1^{-1} and $\beta \rightarrow \alpha$ for E_0^{-1}.

The pair P_0 and P_1 with the trail for E_0 and the pairs P_0, P_2 and P_1, P_3 satisfy the trails for E_1^{-1}, then the pair P_2, P_3 is set up to use the trail $\beta \rightarrow \alpha$ for E_0^{-1}. Consider the intermediate value after half of the rounds, when the previous three trails hold, the formula 2.14 tell why.

$$\begin{aligned} E_0(P_2) \oplus E_0(P_3) &= E_0(P_0) \oplus E_0(P_1) \oplus E_0(P_0) \oplus E_0(P_2) \oplus E_0(P_1) \oplus E_0(P_3) \\ &= E_0(P) \oplus E_0(P_1) \oplus E_1^{-1}(C_0) \oplus E_1^{-1}(C_2) \oplus E_1^{-1}(C_1) \oplus E_1^{-1}(C_3) \\ &= \beta \oplus r \oplus r \\ &= \beta \end{aligned} \tag{2.14}$$

Because the other three differential trails are holding, when the equation is satisfied, a pair of plaintexts P_2, P_3 has the same difference as found in original plaintexts. The idea of boomerang attack is to use such quartet of plaintext and of ciphertext to find key information, and use the information to recover key. In section 4 of Wagner (1999) states that probability of a quartet, assume p_0 is the probability of the differential trail $\alpha \rightarrow \beta$ for E_0 and p_1 is the probability of the differential trail $\delta \rightarrow r$ for E_1. Then the probability p of a right quartet satisfy

$$p \geq p_0^2 p_1^2. \tag{2.15}$$

The boomerang attack has some variants, such as amplified boomerang analysis Kelsey, Kohno, and Schneier (2000) and rectangle analysis Biham, Dunkelman, and Keller (2001). In recent years, a new cryptanalysis method Dunkelman, Keller, Ronen, and Shamir (2020) based on this idea has been proposed, called the retracing boomerang attack , this new attack reduces the complexity of the boomerang method to analyse AES(5-ROUNDS) from 2^{24} to $2^{16.5}$. Also, the boomerang attack can achieve more efficiency with other cryptanalysis methods, such as the related-key method. In Chen et al. (2016), the boomerang attack with a related-key setting achieves the best result for the boomerang to analyse KATAN48/64. But Murphy (2011) published results of several experiments, to demonstrate the boomerang analysis can commonly give highly inaccurate probability values and he shows the boomerang attack is used to analyse DES and AES on some parameters, then the boomerang never comes back.

For the problem, Cid et al. (2018) gives a solution. It proposed new tools to improve the reliability of the sandwich attack, called Boomerang Connectivity Table(BTC). The sandwich attack is a boomerang-style attack depicted in Fig 2.5, it defines $E = E_1 \cdot E_m \cdot E_0$ where E_m is a relatively short operation satisfying some differential propagation among four texts with probability r, then the entire probability is p^2q^2r. Dunkelman, Keller, and Shamir (2010) define r as follows.

$$r = Pr[(x_3 \oplus x_4) = \beta | (x_1 \oplus x_2 = \beta) \wedge (y_1 \oplus y_3 = r) \wedge (y_2 \oplus y_4) = r] \tag{2.16}$$

The BTC can evaluate r in efficacy and easy-to-understand way when E_m is composed of a single S-BOX layer. In this situation, for a given pair of $(\Delta_i, \bigtriangledown_o)$, the probability that a right quartet is generated in each S-BOX in the middle S-BOX layer is given by:

$$\frac{\#\{x \in 0,1^n | S^{-1}(S(x) \oplus \bigtriangledown_o) \oplus S^{-1}(S(x \oplus \Delta_i) \oplus \bigtriangledown_o) = \Delta_i\}}{2^n} \tag{2.17}$$

Where $S : \{0,1\}^n \rightarrow \{0,1\}^n$ is an n-bit to n-bit S-box and S^{-1} is its inverse. When E_m is a single S-BOX layer, the result of Equation2.17 is exactly r in Equation2.16. Cid et al. (2018) demonstrate that the result of using BTC is more than the result only using DDT.

The BTC is not a panacea and can only use ciphers based on S-BOX in practice so far. There is discuss BTC use to ciphers based on ARX in Cid et al. (2018), also points out S-BOX swich not work in ARX.

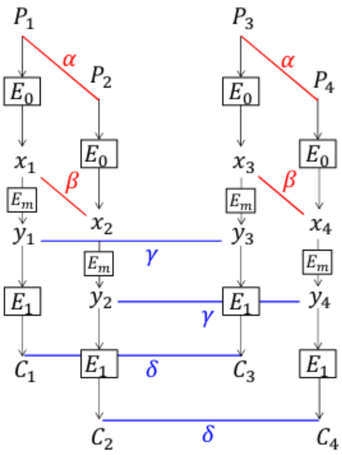


Figure 2.5: Sandwich Attack Model

2.4 Chapter Summary

Section 2.2 discussed the design of block ciphers and introduces three types of design structures, then introduces the basic theory of cryptanalysis and the design of KATAN. Section 2.3 describe the detail of the differential attack, the linear attack and boomerang attack, and discuss the advantages and disadvantages of BTC, this contents of this section are closely related to this study.

CHAPTER 3
METHODOLOGY

3.1 Overview

This chapter describes all processes of this experiment. There are six steps, and each step has a corresponding goal. The steps are shown in Fig1.2.

3.2 Experiment

3.3 Literature Review and Research

In this step, we reviewed the open literature, including block ciphers, KATAN, several attacks and BTC. We went through this step to explore whether BTC can make the results of Boomerang attack more credible, and determine how to improve BTC.

3.4 Define the SMT model

In this step, we analyze KATAN's cryptographic primitives, write Python code that can run on CryptoSMT, and then run differential analysis tests on CryptoSMT to ensure that the code is runnable. This step, we obtain a usable SMT model of KATAN.

3.5 Optimize the smt solver with BTC tools

In this step, we code the smt solver with BTC, and use data that be generated randomly to test and investigate the smt solver over multiple rounds, in the end, we get a boomerang switch.

3.6 Find the best differential trails

In this step, we use the boomerang switch that be created in 3.5, to find the best E_0 and E_1 and a multiple round E_m. In the end, we can crate a verified boomerang distinguisher for KATAN.

3.7 Cover the key

In this step, we generate plaintext and a key randomly, and use the data to cover the key with the boomerang distinguisher that be created in 3.6. In this end, we design a scheme for recover key and recover the key successfully.

3.8 Proposed A tools

Last, when we create a tool using Python that includes the distinguisher we proposed in 3.6, it can automatically generate differential paths for KATAN.

3.9 Chapter Summary

This Chapter describes our experimental steps and our desired goals.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Overview

Overview of the chapter.

4.2 Example Section 1

This chapter will discuss what you have achieved.

4.3 Example Section 2

Provide graphs, charts, tables and discuss your findings. Link them to achieving your objectives.

4.4 Chapter Summary

Summarise the chapter. Section 4.2 discussed.... Section 4.3 discussed...

CHAPTER 5
CONCLUSION AND FUTURE WORK

Summarise your findings here.

5.1 Limitations and Future Work

Discuss limitations and future work here.

5.2 Closing Statements

Discuss if your objectives has been achieved and provide a table to summarise. Table 5.1 will have Section links that do not work. You need to edit and put your own information.

Table 5.1: Summary of Findings and Results

Research Questions	**Research Objectives**	**Findings**	**Results**
How can symmetric-key ciphers be secured against statistical-based cryptanalysis?	To develop symmetric encryption encryption schemes that can seamlessly integrate true random numbers to resist statistical-based cryptanalysis.	**Section ??**: AEAD-based scheme **Section ??**: Stream cipher-based scheme	**Section ??**: Theoretical evaluation of security against block cipher cryptanalysis. **Section ??**: Theoretical evaluation of security against stream cipher cryptanalysis.
What strategies can be employed to construct efficient yet easily obtainable TRNGs?	To design chaos-based TRNGs that are highly efficient, statistically secure and non-deterministic by using easily obtainable computer hardware as entropy sources.	**Section ??**: TRNG design process **Section ??**: TRNG evaluation process **Section ??**: CPU-TRNG **Section ??**: GPU-TRNG **Section ??**: Audio-TRNG	**Section ??**: Statistical evaluation: CPU-TRNG. **Section ??**: Statistical evaluation: GPU-TRNG. **Section ??**: Statistical evaluation: Audio-TRNG. **Section ??**: Complexity and throughput analysis of the proposed TRNGs.
How can true random numbers influence an encryption or decryption process while minimizing computational overhead? - What strategies can be employed to maximize the performance of a non-deterministic encryption algorithm based on chaotic maps?	To design a chaos-based authenticated block cipher based on the proposed true random number scheme that has security and performance comparable to current cryptographic standards. - To design a chaos-based stream cipher based on the proposed true random number scheme that has security and performance comparable to current cryptographic standards.	**Section ??**: Chaos-based authenticated block cipher with associated data **Section ??**: Chaos-based stream cipher	**Section ??**: Security evaluation: Chaos-based authenticated block cipher. **Section ??**: Performance analysis: Chaos-based authenticated block cipher. **Section ??**: Security evaluation: Chaos-based stream cipher. **Section ??**: Performance analysis: Chaos-based stream cipher.

REFERENCES

Biham, E., Dunkelman, O., & Keller, N. (2001). The rectangle attack—rectangling the serpent. In *International conference on the theory and applications of cryptographic techniques* (pp. 340–357).

Biham, E., Dunkelman, O., & Keller, N. (2005). A related-key rectangle attack on the full kasumi. In *International conference on the theory and application of cryptology and information security* (pp. 443–461).

Biham, E., & Shamir, A. (1991). Differential cryptanalysis of des-like cryptosystems. *Journal of CRYPTOLOGY*, *4*(1), 3–72.

Biryukov, A., De Cannière, C., & Dellkrantz, G. (2003). Cryptanalysis of safer++. In D. Boneh (Ed.), *Advances in cryptology - crypto 2003* (pp. 195–211). Berlin, Heidelberg: Springer Berlin Heidelberg.

Biryukov, A., & Khovratovich, D. (2009). Related-key cryptanalysis of the full aes-192 and aes-256. In *International conference on the theory and application of cryptology and information security* (pp. 1–18).

Chen, J., Teh, J. S., Su, C., Samsudin, A., & Fang, J. (2016, 07). Improved (related-key) attacks on round-reduced katan-32/48/64 based on the extended boomerang framework. In (Vol. 9723, p. 333-346). doi: 10.1007/978-3-319-40367-0_21

Cid, C., Huang, T., Peyrin, T., Sasaki, Y., & Song, L. (2018). Boomerang connectivity table: A new cryptanalysis tool. In J. B. Nielsen & V. Rijmen (Eds.), *Advances in cryptology – eurocrypt 2018* (pp. 683–714). Cham: Springer International Publishing.

De Cannière, C., Dunkelman, O., & Knežević, M. (2009). Katan and ktantan — a family of small and efficient hardware-oriented block ciphers. In C. Clavier & K. Gaj (Eds.), *Cryptographic hardware and embedded systems - ches 2009* (pp. 272–288). Berlin, Heidelberg: Springer Berlin Heidelberg.

Dunkelman, O., Keller, N., Ronen, E., & Shamir, A. (2020). The retracing boomerang attack. In *Annual international conference on the theory and applications of cryptographic techniques* (pp. 280–309).

Dunkelman, O., Keller, N., & Shamir, A. (2010). A practical-time related-key attack on the kasumi cryptosystem used in gsm and 3g telephony. In *Annual cryptology conference* (pp. 393–410).

Dworkin, M. J., Barker, E. B., Nechvatal, J. R., Foti, J., Bassham, L. E., Roback, E., ... others (2001). Advanced encryption standard (aes).

Dworkin, M. J., et al. (2015). Sha-3 standard: Permutation-based hash and extendable-output functions.

Kelsey, J., Kohno, T., & Schneier, B. (2000). Amplified boomerang attacks against reduced-round mars and serpent. In *International workshop on fast software encryption* (pp. 75–93).

Matsui, M. (1993). Linear cryptanalysis method for des cipher. In *Workshop on the theory and application of of cryptographic techniques* (pp. 386–397).

Murphy, S. (2011). The return of the cryptographic boomerang. *IEEE Transactions on Information Theory*, *57*(4), 2517-2521. doi: 10.1109/TIT.2011.2111091

Pub, F. (1999). Data encryption standard (des). *FIPS PUB*, 46–3.

Shannon, C. E. (1949). Communication theory of secrecy systems. *The Bell System Technical Journal*, *28*(4), 656-715. doi: 10.1002/j.1538-7305.1949.tb00928.x

Wagner, D. (1999). The boomerang attack. In L. Knudsen (Ed.), *Fast software encryption* (pp. 156–170). Berlin, Heidelberg: Springer Berlin Heidelberg.

APPENDICES

APPENDIX A - CHAOS THEORY

You can add additional data here that may not be vital, or too lengthy for the actual reader.