

Multiple New Formulas for Cipher Performance Computing

Youssef Harmouch¹, Rachid Elkouch¹, Hussain Ben-azza²

(Corresponding author: Youssef Harmouch)

Department of Mathematics, Computing and Networks, National Institute of Posts and Telecommunications¹

10100, Allal El Fassi Avenue, Rabat, Morocco

(Email:harmouch@inpt.ac.ma)

Department of Industrial and Production Engineering, Moulay Ismail University²

National High School of Arts and Trades, Mekns, Morocco

(Received Apr. 03, 2017; revised and accepted July 17, 2017)

Abstract

Cryptography is a science that focuses on changing the readable information to unrecognizable and useless data to any unauthorized person. This solution presents the main core of network security, therefore the risk analysis for using a cipher turn out to be an obligation. Until now, the only platform for providing each cipher resistance is the cryptanalysis study. This cryptanalysis can make it hard to compare ciphers because each one is vulnerable to a different kind of attack that is often very different from others. Our contribution in this paper is to develop new risk analysis formulas to offer a theoretical background for both the cipher designer and the simple users. Those formulas will help to suggest a fair platform for measuring risk, safety, complexity and cost, in order to determine a quantifiable value for performance to each cipher. This can lead to a fair comparison in a fair scale.

Keywords: Complexity & Cost; Quantifiable-Value; Risk Analysis; Security-Level; Security Performance

1 Introduction

Although, humans today constantly depend on computer technology in their life, they continue to have a hard follow to security aspects between different technologies. This is caused by the tiny ability to compare, to contrast, and to make quantifiable statements about security systems. This means that having a fixed global model for information security is extremely valuable for having a basis to determine where to put limited resources, pay attention, and how to best secure systems.

However, risk analysis (quantitative or qualitative [16, 19]) remains a difficult problem, since computer security is a multidimensional attribute (confidentiality, availability, integrity, non-rejection, accountability, authenticity, reliability of IT systems, etc.). Moreover, these dimensions

are not necessarily commensurate properties. For example, an online newspaper will be primarily interested in the integrity of their information while a financial stock exchange network may define their security as real-time availability and information privacy [14, 23]. This means that, the many facets of the attribute must all be identified and adequately addressed. Furthermore, the security attributes are terms of qualities, thus measuring such quality terms need a unique identification for their interpretations meaning [20, 24]. Besides, the attributes can be interdependent. The first thing is to identify a set of security-related attributes that are important to the use of the system. This leads to decide whether the system security must be represented as a vector or as a single value.

In large systems, risk analysis becomes a very painful task. The remaining solution to use the decomposition method to develop simple, small and stand alone components of the system. Therefore, in order to better measure risk analysis of the systems, it is necessary to seek for the common ground between all the systems and their components. This common ground is the security protocols or algorithms. Since this latter is based on computer, mathematical and/or logical operations, the scale for risk analysis should be changed from **macro scale** (network application, software, threats, hardware, protocols, etc.) to **micro scale** (cipher and algorithm).

Thus, this research studies the cipher risk analysis upon **MLO** (**M**athematical or/and **C**omputer **L**ogical **O**peration). This paper proposes new risk formulas that represent an improvement in security quantification. By calculating the security level offered by each MLO, the cryptograph can easily choose which MLO to use and where to place it, so as to increase the complication of the cipher/algorithm within the developing phase before applying any cryptanalysis studies. This paper provides a fair comparison of **security**, **risk** and utilizing **cost** for several ciphers based on their MLO while respecting each

cipher properties. These properties englobe inner structures, key space, round number, complexity, successful cryptanalysis attacks, *etc.* to provide a value that determines the **safety degree** of each cipher, in order to put a fair platforms where ciphers can be judged and compared precisely.

The paper is organized as follows. In Section 2, we introduce some concepts of cryptography, then we define the different structures utilized by ciphers followed with an introduction of the most knowing cryptanalysis attacks, after that, we introduce the ciphers used in this study. In Section 3, we investigate the study and the development of the new formulas for risk analysis, while in Section 4, we focus on results and discussion. We conclude the paper in Section 5.

2 State of Art and Motivation

Because information privacy has become a major concern for both users and companies, cryptography is considered as a standard for providing information trust, security, electronic financial transactions, controlling access to resources and stopping non-authorized persons from obtaining critical or private information. It must be mentioned that the strength of the cryptography algorithm depends on the length of the key, secrecy of the key, the complexity of the process and how they all work together [18].

Ciphers differ with their construction structure. This leads to different types of comportment. These structures can be organized as (for symmetric cryptography):

- Permutation network: is when a cipher uses a permutation box (P-box). This latter is used to permute or transpose data across plaintext, retaining diffusion while transposing [9].
- Substitution network: is when a cipher uses a substitution box (S-box). It is used to obscure the linearity between the key and the ciphertext [7,8].
- Substitution permutation network: is when a cipher uses both S-box and P-box in its encryption function.
- Feistel Network: is when a cipher uses a Feistel scheme. It is a technique used in the construction of block cipher-based algorithms and mechanisms [13]. If the two blocks (left and right) are not of equal length, then the scheme is called unbalanced Feistel scheme.
- Lai-Massey scheme: is when a cipher uses a Lai-Massey scheme [26].

During the cipher design phase, the cryptograph applies one or more structures, to determine the security level offered by the cipher besides its behavior. Those structures in addition to nonlinear functions are an important functionality that each cipher must have in order to put confusion and diffusion alongside, to prevent the finding

of any linear link between plaintext and ciphertext so as to increase the complexity of breaking the cipher.

Cryptanalysis tests the weakness of the cryptosystem by trying to break it without any knowledge of the key used. The most popular attack is the brute force where the cyber criminal tries every possible key to break ciphers; therefore, the only way to resist is by enlarging the key space to make it infeasible [21].

Thus, the question that needs to be asked is, “Which cipher is the best in security term and how can we measure its safety?” The ability to compare and/or to make quantifiable statements about system security is extremely valuable, since it offers a basis to determine how to best secure the systems. Besides, the complete understanding of a subject cannot be done with neither measurements nor quantifying value as written by Lord Kelvin in 1883: “When you can measure what you are speaking about and express it in numbers you know something about it, but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind” [23].

The development of the theory of measurements and quantifying cipher is the main motivation for this research. However, it is a difficult problem due to The large number of structure and operation that cipher utilizes. Moreover, in this research, the only trust measurements used is based over probabilities.

Furthermore, this research alongside with cryptanalysis tries to rate each cipher based on inner structure, key space, successful attacks, *etc.* in order to have a fair platform for comparison and does not try to replace the cryptanalysis, which is still the only science that tests the cipher weakness.

In this work, we try to answer the following questions: if a user has two ciphers, A1 and A2 while knowing that the best successful cryptanalysis attack for A1 resp. A2 is B1 resp. B2 and the two attacks have the same successful rate, what is the best cipher to choose? If A1 is a bit quicker than A2 and B2 is a little less successful than B1, what is the most optimal cipher for a system/network? Another question, if a user has multiple ciphers A1, A2, ..., An, what is the order of the most suitable cipher to her/his own system/network? i.e. what is the order of ciphers which offers an acceptable resistance (not always the best) to cryptanalysis attacks, and which is the most suitable to the system/network (real time application, Full HD conferencing, high throughput network, data-center file encryption, *etc.*)?

Even if risk analysis proposed in this paper can be used for any type of encryption (symmetric, asymmetric, bloc, stream), this paper focuses on symmetric block ciphers like [12] “AES, Blowfish, Camellia, CAST-128/256, DES/3DES, GOST, IDEA, MARS, RC2, RC5, RC6, Serpent, SHACAL2, SHARK, SKIPJACK, Three-way, Twofish, and XTEA”. Each cited cipher will be revised respecting each one’s properties; such as key length, block length and the mode of operation.

3 Risk Analysis

Risk analysis is a technique used to identify and assess factors that may put at risk the safety of security based upon a cipher. This technique helps to define the optimal cipher in order to reduce the probability for these factors from occurring. Therefore, in this section we define multiple indexes factor to help studying every cipher either in the design phase or in comparison with those that already exist. These factors are called **index of safety**, **index of risk**, **complexity** and **cipher cost** and will be defined in the following paragraph.

3.1 Index of Safety (IS)

IS defines the level of security factor that a cipher offers to users, that is to say, this factor studies the robustness of the cipher structure. It consists of round number (R), key-block index (K/B) that defines the length of key per length of data block, and the structure type index (S) such as Feistel, P-box, S-box, etc. multiplied by their factors and the number of uses in one round. As so and before defining IS, several definitions must be provided:

Definition 1. We define ρ as the break-probability for a structure or operation used in encryption process. i.e. ρ is equal to probability of extracting the plaintext from the ciphertext after applying a structure or an operation to the plaintext.

The following Table 1 shows ρ for different encryption operations where “ m ” is the block length in bits and “ ξ ” defines the modulus.

Table 1: Break probability for structure type or operation

Structure type	Break Probability
AND	$1/2^{2m}$
OR	$1/2^{2m}$
XOR	$1/2^m$
Concatenation	1
Modular addition	$1/\xi^{\frac{m}{\log_2 \xi}}$
Modular subtraction	$1/\xi^{\frac{m}{\log_2 \xi}}$
Modular multiplication	$1/\xi^{\frac{2m}{\log_2 \xi}}$
Modular exponentiation	$1/\xi^{\frac{2m}{\log_2 \xi}}$
Left or Right rotation	$1/(m - 1)$
NOT operation	1
Conditional NOT operation	$1/2^m$
Permutation box	$1/2^{m-1}$
Substitution box	$1/(2^m - 1)$
balanced Feistel schema	$1/2^{\frac{m}{2}}$
Lai-Massey schema	$1/2^{2m}$

Since each structure presents a different bit operation and is linearly correlated to both “bit number and operation type”, the structure/operation resistance can be measured through ρ .

Note 1. The proof of the results listed in Table 1 is presented in Appendix-A.

Definition 2. We define the measurement of the resistance factor S for a structure or operation by:

$$S = -\log_2 \rho \tag{1}$$

where ρ is the break probability (see Table 2 below).

Since the “1/2” is common in all type of structure due to binary representation, it does not provide any utility for comparison. Remark that, the only valuable information in ρ expression is “ m ” or “ ξ ”. As so, in Equation (1), we use “ $-\log_2$ ” to remove the “1/2” and get the useful information (m, ξ) for the future study.

Table 2: Resistance factor for structure type or operation

Structure type	Resistance Factor
AND	$2m$
OR	$2m$
XOR	m
Concatenation	0
Modular addition	m
Modular subtraction	m
Modular multiplication	$2m$
Modular exponentiation	$2m$
Left or Right rotation	$\log_2(m - 1) \approx \log_2 m$
NOT operation	0
Conditional NOT operation	m
Permutation box	$m - 1$
Substitution box	$\log_2(2^m - 1) \approx m$
balanced Feistel schema	$m/2$
Lai-Massey schema	m

The resistance factor enlarge the scale from $[0, 1]$ for ρ to $[0, +\infty[$ with valuable conservation of bit information “ m ”. therefore, the best resistance factor unity is “bit”.

Definition 3. We define the **Block-Round Function BRF** as the block contains one or more successive functions that have the same round number.

Example 1. Let us define A, B and C as a function or a bit operation and “ $r1$ ”, “ $r2$ ” as the round number ($r1 \neq r2$). According to the following algorithms, we define BRF for each algorithm as showed in following Table 3.

Definition 4. We define the resistance factor efficiency **ES** for one BRF by:

$$ES = \frac{1}{\lambda} \sum_{i=1}^{\lambda} S_i \tag{2}$$

where S_i is the resistance factor for the structure or operation number “ i ” and λ is the total number of structures or operations in one BRF block.

Using Equation (1) in Equation (2) gives

Table 3: BRF definition example

Description	Algorithm.1	Algorithm.2	Algorithm.3	Algorithm.4
Algorithm Body	for i from 0 to r1 repeat {A;B} end repeat C	for i from 0 to r1 repeat {A;B;C} end repeat	for i from 0 to r1 repeat {A;B} end repeat for i from 0 to r2 repeat {C} end repeat	for i from 0 to r1 repeat {A;B;C} end repeat for i from 0 to r2 repeat {A;B;C} end repeat
Number of BRF	2	1	2	2
BRF function	BRF1={A, B} BRF2={C}	BRF1={A, B, C}	BRF1={A, B} BRF2={C}	BRF1={A, B, C} BRF2={A, B, C}

$$ES = \frac{1}{\lambda} \sum_{i=1}^{\lambda} S_i = -\frac{1}{\lambda} \sum_{i=1}^{\lambda} \log_2 \rho_i$$

where ρ_i is the break probability for the structure or operation number “ i ”.

Note 2. ES indicates the efficacy for a BRF .i.e. ES indicates the level of security at which a function or operation affects the rest of the functions in a BRF, and so, ES helps to measure the security offered through a BRF. ES and S have the same unit “bits”.

Definition 5. To determines the potential security-level offered by a cipher, we define the key-block factor KB for a cipher by:

$$KB = \alpha/\beta \tag{3}$$

where α is the bit number of the key and β is the bit number of the plaintext block.

Note 3. Equation (3) do not consider the number or the length of the sub-keys. According to that, and as an example AES-128 has $KB=1$, AES-192 has $KB=1.5$ and AES-256 has $KB=2$.

Definition 6. We define the total resistance factor TS for a BRF number “ i ” by:

$$TS_i = R_i \times ES_i \tag{4}$$

where R_i is the round number for the BRF number “ i ”.

Note 4. Since ES unity is “bit” and R is just a number, we propose a new unity for TS called RB “Round Bit”.

Generally, a cipher is a composition of one or more BRFs. Hence, the cipher total resistance factor is obtained by adding all its BRFs total resistance factor. Thus, Equation (4) gives:

$$\begin{aligned}
 TS &= \sum_{i=1}^{\mu} TS_i = \sum_{i=1}^{\mu} r_i \times ES_i = \sum_{i=1}^{\mu} r_i \times \frac{1}{\lambda_i} \sum_{i=1}^{\lambda_i} S_i \tag{5} \\
 &= -\sum_{i=1}^{\mu} r_i \times \frac{1}{\lambda_i} \sum_{i=1}^{\lambda_i} \log_2 \rho_i
 \end{aligned}$$

where “ μ ” is the BRF number in the cipher encryption process, “ λ_i ” is the number of structures or operations

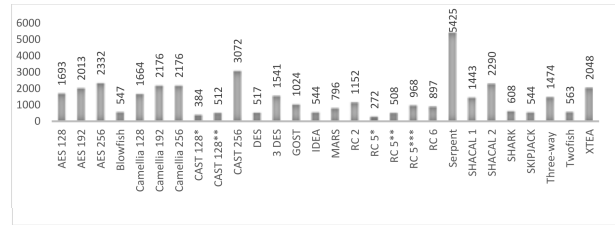


Figure 1: TS measurement for several ciphers

for the BRF number “ i ” and “ r_i ” is the round number for the BRF number “ i ”.

Figure 1 illustrates the TS calculated for all the studied ciphers, while respecting the plaintext block length of each cipher. Figure 1 uses the equations taken from the following Table 4. The equations in Table 4 are calculated using Equation (5).

Figure 1 shows that the serpent presents the highest TS factor followed by CAST-256 and AES-256, this result means that serpent uses more complicated and complex inner structure and/or more round numbers, which yields to more ciphertext-complexity and thus increasing the resistance-probability to cryptanalysis attacks. Note that, it is not always the highest TS that is the more secure, because TS does not provide any information about successful cryptanalysis attacks applied over the cipher and its success rate.

Note 5. In Figure 1, we use the terminology:

- *CAST-128*/CAST-128*** is for key length from 40 to 80/80 to 128 bits.
- *RC5*/RC5**/RC5**** is for plaintext length 32/64/128 bits.

Definition 7. We define the index of security for a cipher by:

$$IS = \log_{10}(KB \times TS) \tag{6}$$

Note 6. The goal of the logarithm scale used in Equation (6) is to reduce the vast values obtained from computed IS. Moreover we define a new unity of IS called **SC** “Security per Cipher”.

Accordingly, Table 5 englobes IS measurement for several ciphers which can be observed graphically in the following Figure 2.

Table 4: Total resistance factor for several ciphers

Cipher	TS (n : Length of the plaintext block in bits)
AES 128	$(129n - 31)/12$
AES 192	$(189n - 37)/12$
AES 256	$(219n - 43)/12$
Blowfish	$(17n + \log_2(n) - 1)/2$
Camellia 128	$13n$
Camellia 192	$17n$
Camellia 256	$17n$
CAST 128	$6n$ if $40\text{bits} \leq \text{keysize} < 80\text{bits}$ $8n$ if $80\text{bits} \leq \text{keysize} < 128\text{bits}$
CAST 256	$24n$
DES	$8n + \log_2(n) - 1$
3DES	$24n + \log_2(n) - 1$
GOST	$16n$
IDEA	$\frac{17}{2}n$
MARS	$(41n + \log_2(n) + 56\log_2(13) + 14)/7$
RC 2	$18n$
RC 5	$n + [1 \rightarrow 255](\log_2(n) - 1 + n/2)$
RC 6	$(47n + 40\log_2(n) - 20)/7$
Serpent	$\frac{1}{3}(127n + 1 + \log_2(3) + \log_2(5) + 2\log_2(7) + \log_2(11) + \log_2(13))$
SHACAL 1	$(35n + 40\log_2(n) - 120)/4$
SHACAL 2	$(35n + 40\log_2(n) - 120)/4$
SHARK	$19n/2$
SKIPJACK	$8(n + \log_2(n) - 2)$
Three-way	$(77n - 11)/5$
Twofish	$(25n + 32\log_2(n - 4) + 3\log_2(n) - 67)/6$
XTEA	$32n$

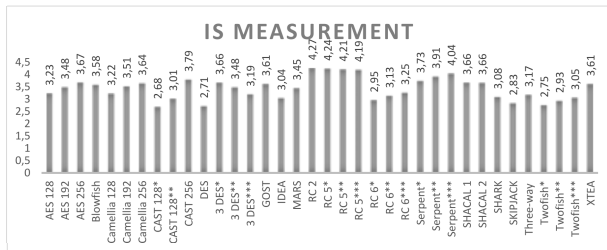


Figure 2: IS measurement for several ciphers

Just like TS, IS provides a scale to measure the possibility of being more secure to cryptanalysis attacks by collecting different information (TS, key). This collection focuses on the complexity of the cipher body and the key space without any examination to cryptanalysis attack neither to the nature of the cipher body. This is why in Figure 2, RC2 and RC5 shows more IS value than Serpent, and 3DES shows more IS value than AES-192.

It is clearly observed that IS does not provide all information for best cipher rating, that is why another factor is needed. This factor will measure the risk of using the cipher by studying its best-known successful cryptanalysis attacks. This measure is explained in the following section under the name of **Index of Risk (IR)**.

3.2 Index of Risk (IR)

As each cipher has a different structure, many different cryptanalysis attacks are invented and developed including: “Linear cryptanalysis [17], Differential cryptanalysis [1], Differential-linear cryptanalysis, Impossible differential cryptanalysis, Truncated differential cryptanalysis [5], Integral cryptanalysis, Higher-order differential cryptanalysis [25], Meet-in-the-middle [2], Slide attack [4], Boomerang Attack [22], Related Key Attack [3], Mod n [15], XSL [6], Frequency analysis [11], The index of coincidence, Chi-square test [10], etc.”

The major differences between those attacks above make it difficult to fairly judge and compare cipher. As so, in order to create a credible scale, we define a new term called the Index of Risk IR.

IR defines the measure of the risk of using a cipher. It combines the success rates of the most successful cryptanalysis attacks and the security index that the ciphers offers. However before defining IR, several definitions must be mentioned.

Definition 8. We define **BA** as the best success rate factor for a multiple cryptanalysis attacks by:

$$BA = 1 - \frac{\min_{i \in [0, \tau-1]} \log_2(CCA_i)}{\text{key length in bits}} \quad (7)$$

where τ presents the number of cryptanalysis attacks while CCA_i is the computational complexity of the attack number “i”.

CCA_i divided by the key-length in Equation (7) presents the success rate factor or the percentage rate for a successful cryptanalysis attack. To show this, we compute BA based on data taken from Table 2 in the paper [12]. Figure 3 contains the computing result:

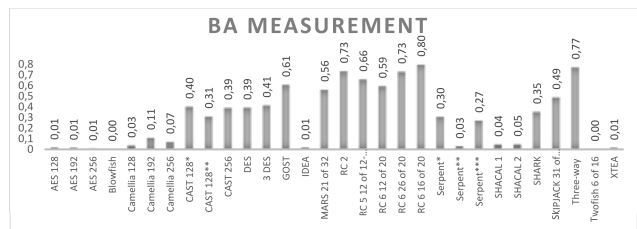


Figure 3: BA measurement for several ciphers

Definition 9. We define the index of risk (IR) for a cipher as:

$$IR = 100 \times \frac{BA}{IS} \quad (8)$$

Equation (8) takes into consideration two factors: the success rate for cryptanalysis attacks and the measured index of security across the body of the cipher. The number 100 is just a coefficient to enlarge the scale since dividing the rate for a successful cryptanalysis attack by the safety of the cipher body gives results always less than “1”. Figure 4 shows the calculated IR.

Table 5: IS calculating for several ciphers per key length and plaintext block length

Cipher	Key Length (bits)	Block Length (bits)	TS (RB)	KB	IS for multiple case (SC)	max IS (SC)
AES 128	128	128	1693, 42	1	3, 22876383	3, 22876383
AES 192	192	128	2012, 92	1, 5	3, 479917055	3, 47991705
AES 256	256	128	2332, 42	2	3, 668836132	3, 66883613
Blowfish	32 → 447	64	546, 5	0, 5 → 6, 98	(2, 43656, 3, 5817)	3, 58171772
Camellia 128	128	128	1664	1	3, 221153322	3, 22115332
Camellia 192	192	128	2176	1, 5	3, 51375015	3, 51375015
Camellia 256	256	128	2176	2	3, 638688887	3, 63868889
CAST 128*	40 → 80	64	384	0, 63 → 1, 25	(2, 38021, 2, 6812)	2, 68124124
CAST 128**	80 → 128	64	512	1, 25 → 2	(2, 80618, 3, 0103)	3, 01029996
CAST 256	138 → 256	128	3072	1, 08 → 2	(3, 52009, 3, 7885)	3, 78845121
DES	64	64	517	1	2, 713490543	2, 71349054
3 DES*	192	64	1541	3	3, 664923893	3, 66492389
3 DES**	124	64	1541	1, 9375	3, 47504435	3, 47504435
3 DES***	64	64	1541	1	3, 187802639	3, 18780264
GOST	256	64	1024	4	3, 612359948	3, 61235995
IDEA	128	64	544	2	3, 036628895	3, 0366289
MARS	128 → 448	128	796 → 318	1 → 3, 5	2, 90109 → 3, 4452	3, 44515447
RC 2	8 → 1024	64	1152	0, 13 → 16	2, 15836 → 4, 2656	4, 26557246
RC 5*	8 → 2040	32	272	0, 25 → 63, 8	1, 83251 → 4, 239	4, 23904909
RC 5**	8 → 2040	64	508	0, 13 → 31, 9	1, 80277 → 4, 2093	4, 20931391
RC 5***	8 → 2040	128	968	0, 06 → 15, 9	1, 78176 → 4, 1883	4, 18829556
RC 6	128	128	896, 571	1	2, 952584895	2, 95258489
RC 6	192	128	896, 571	1, 5	3, 128676154	3, 12867615
RC 6	256	128	896, 571	2	3, 253614891	3, 25361489
Serpent*	128	128	5425, 09	1	3, 734406852	3, 73440685
Serpent**	192	128	5425, 09	1, 5	3, 910498111	3, 91049811
Serpent***	256	128	5425, 09	2	4, 035436848	4, 03543685
SHACAL 1	128 → 512	160	1443, 22	0, 8 → 3, 2	3, 06242 → 3, 6645	3, 6644823
SHACAL 2	128 → 512	256	2290	0, 5 → 2	3, 05881 → 3, 6609	3, 66086548
SHARK	128	64	608	2	3, 084933575	3, 08493357
SKIPJACK	80	64	544	1, 25	2, 832508913	2, 83250891
Three-way	96	96	1474	1	3, 168497484	3, 16849748
Twofish*	128	128	562, 756	1	2, 750319913	2, 75031991
Twofish**	192	128	562, 756	1, 5	2, 926411172	2, 92641117
Twofish***	256	128	562, 756	2	3, 051349909	3, 05134991
XTEA	128	64	2048	2	3, 612359948	3, 61235995

3.3 Cipher Cost (CC)

CC defines the cost of using a cipher for a system/network. It depends on IR and complexity. The most developed cipher is normally working only on a fixed block size of plaintext. This takes approximately the same time for encryption/decryption independently of input (ECB mode), thus they are $O(1)$.

Even if we put them into a mode of operation to encrypt a longer plaintext, we usually get an $O(m)$ complexity, where “ m ” is the plaintext size, as we have $O(m)$ blocks of data to encrypt. This $O(m)$ presents the minimum, because each cipher has to encrypt at least each input-bit once, to be reversible, even if different modes of operations have different complexity (Triple-DES usually needs three times the computing power as DES, but still then $O(1)$ or $O(m)$).

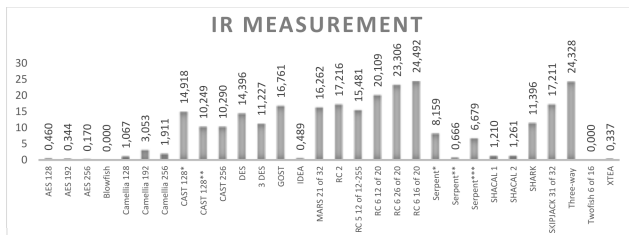


Figure 4: IR measurement for several ciphers

As a result, the uses of O becomes useless to compare the complexity between ciphers. Thus, from now on, we redefine the **complexity** as “**the number of CPU cycle needed to encrypt the plaintext**”. Since the plaintext size differs from cipher to another, we set the plaintext size required for the complexity measurement as one Mega-Byte. Besides, this complexity is linearly linked to the computing power or computing time, as so it allows multiple usages for it.

Furthermore, the complexity must have a **reference point** in order to allow a future comparison. This reference point will be the complexity of encrypting the plaintext with XOR operation, since XOR is the fastest strong easy low-power consumption and simple cryptographic computer operation. Thus, we define the normalized complexity Γ as the ratio between the complexity of the cipher and the complexity of XOR:

$$\Gamma = \frac{\text{Cipher complexity}}{\text{XOR complexity}} \tag{9}$$

Since this paper is interested in putting a quantifiable value to cipher performance, the Γ measurement for ciphers (studied in this paper) from Equation (9) must be standardized (standard score), because we are only interested in choosing the less cipher-complexity compared to others. Thus, Γ becomes $\underline{\Gamma}$:

$$\underline{\Gamma} = \Gamma/\sigma \tag{10}$$

where σ is the standard deviation of Γ for all studied ciphers.

Note 7. *There is no need for the subtraction of Γ by the mean “ μ ” in Equation (10) because it presents just a shift scale by “ $-\mu/\sigma$ ”.*

Table 6 shows the measurement of $\underline{\Gamma}$ for all studied ciphers, this measurement is illustrated in Figure 5.

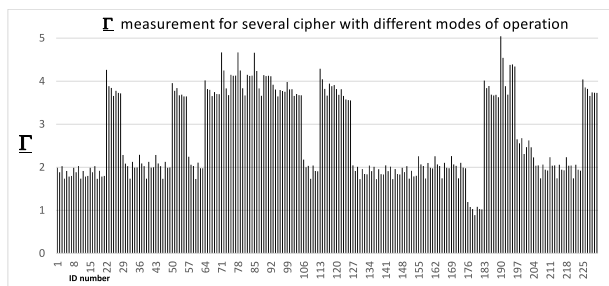


Figure 5: $\underline{\Gamma}$ measurement for several ciphers with different modes of operation

The experimental environment for the complexity measurement was a C++ code application developed in Microsoft Visual studio 2010 for Windows 7 desktop and GCC 4.8.2 for Centos7 for Linux OS. The ciphers used in this study are taken from two version of an open source library called Crypto++ (cryptopp5.6.2 and cryptopp5.6.3). The test was running under different machine

(from Intel core2 until Intel core i5). As observed from our experimental results, the changing of OS affects the complexity in about 10%, while the changing of library affects less than 4%. The most important change in complexity was when changing the tested machines (up to 70%). $\underline{\Gamma}$ decreases the difference between results in less than 0,3%. This tiny difference makes the values in Table 6 a trustful result to calculate the cost of using every studied ciphers.

Definition 10. *We define the cipher cost (CC) by:*

$$CC = IR \times \underline{\Gamma} \tag{11}$$

4 Results & Discussion

The CC study in Equation (11) englobes the recognition of many parameters (safety, speed, resistance, risk...) for ciphers in different modes of operation. This helps to provide a good platform to compare ciphers with many considered variables as sizes of data blocks, key size, type of cipher, complexity, round number, successful cryptanalysis attacks...

CC, IS, complexity ..., and IR present a theoretical and logical MLO formulas for studying risk analysis. These formulas will help to obtain quantifiable values, so as to support either a cipher designer or a normal user to choose the most optimal ciphers to his/her network/system. Since each parameter has a different definition and interpretation, we define each parameter unity as following (see Table 7).

Figure 6 illustrates the data presented in Table 8. It shows the CC values for all studied ciphers with their different operation modes.

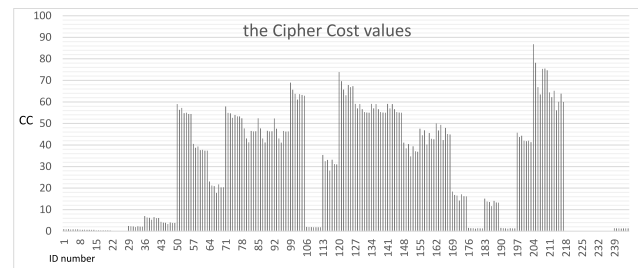


Figure 6: CC measurement for several ciphers with different mode of operation

Figure 6 shows that CC is linearly related to the applied mode of operation (CBC-CTS, CBC, CFB-FIPS, CFB, CTR, ECB and OFB) and the used cipher. For example, Camellia, MARS, RC2 and SKIPJACK show less cost in FIPS than CTR as opposed to AES, DES, GOST, IDEA, RC5/6, Three-Way and XTEA that show more cost in FIPS than CTR. In addition, we notice that Twofish with 128 bits in key has less complexity than Twofish with 192/256 bits in key and the three have the same cost. This is due to the lack of a successful cryptanalysis attack which also makes for instance both

Table 7: The unity and signification for each parameters

Parameter	Unity	Signification
Break Probability ρ	-	
Resistance Factor S	Bit	
Resistance Factor Efficacy ES	Bit	
Key-Block Factor KB	-	
Total Resistance Factor for Cipher Structure TS	RB	Round bit
Index of Security IS	SC	Security per cipher
W	-	
Best Success Rate Factor BA	-	
Index of Risk IR	$RC = SC^{-1}$	Risk per cipher
Complexity Γ	Cycle	CPU cycle
Normalize Complexity $\underline{\Gamma}$	Xcycle	CPU cycle per XOR
$\underline{\Gamma}$	CP	Complexity per Processor
Cipher Cost CC	PR	Performance cost for risk and complexity

Blowfish and Twofish less costly in use than AES with 256/192/128 bits in key. Furthermore, this absence of risk make Twofish with 128 bits in key more optimal since it requests less time for encryption. After the AES, we notice that SHACAL2 and Serpent with 192 bits in key come next, followed by XTEA, SHARK, IDEA, Camellia with 128 bit in key, Camellia with 256 bits in key, Serpent with 256 bits in key, CAST with 256 bits in key, MARS, RC6 and CAST with 128 bits in key. Finally, in the sorted CC list, we note that the greatest cost for using a cipher was taken by DES, followed by 3DES, RC5, RC2 and SKIPJACK.

This result has the advantage of combining theoretical (cryptanalysis attack) and experimental (complexity) results. This combination makes the result valuable and very interesting because a lot of cryptographic studies separate the theoretical background from the experimental results. This separation may cause a loss of information, which makes any comparison between ciphers in their mode of operations less fair and less equitable.

5 Conclusions and Future Work

This article contains new formulas and a definition of risk analysis factors for ciphers. These formulas take into account security factors, risk factors and the ciphers using-cost, while respecting in each cipher its own structure and properties. These parameters include structure, key space, round number, encryption mechanism, complexity and successful cryptanalysis attacks, etc..

These formulas provide a lot of information to allow future comparison in a fair platform, which will help a decision maker to select the most appropriate cipher for its own system with its QOS recommendation. In addition, the ciphers designer can also benefit from these formulas constructed on MLO because it offers the theoretical quantifiable value to test the encryption process before applying any cryptanalysis attack.

Table 8: CC measurement for several cipher with different mode of operation

ID	Cipher/mode	CC	ID	Cipher/mode	CC	ID	Cipher/mode	CC
1	AES128/CBC-CTS	0.913598	83	3DES64/ECB	46.311394	165	RC6256/CFB	49.326927
2	AES128/CBC	0.866341	84	3DES64/OFB	46.349990	166	RC6256/CTR	42.271908
3	AES128/CFB-FIPS	0.931290	85	3DES124/CBC-CTS	52.418960	167	RC6256/ECB	47.942699
4	AES128/CFB	0.797812	86	3DES124/CBC	47.711889	168	RC6256/OFB	45.151879
5	AES128/CTR	0.880670	87	3DES124/CFB-FIPS	43.064941	169	Serpent128/CBC-CTS	44.914185
6	AES128/ECB	0.819800	88	3DES124/CFB	41.170878	170	Serpent128/CBC	45.401617
7	AES128/OFB	0.827240	89	3DES124/CTR	46.575465	171	Serpent128/CFB-FIPS	46.850546
8	AES192/CBC-CTS	0.684291	90	3DES124/ECB	46.273300	172	Serpent128/CFB	46.547380
9	AES192/CBC	0.648608	91	3DES124/OFB	46.351916	173	Serpent128/CTR	44.188476
10	AES192/CFB-FIPS	0.698142	92	3DES196/CBC-CTS	52.313093	174	Serpent128/ECB	47.126104
11	AES192/CFB	0.596957	93	3DES196/CBC	47.585048	175	Serpent128/OFB	46.215067
12	AES192/OFB	0.659049	94	3DES196/CFB-FIPS	43.003079	176	Serpent192/CBC-CTS	46.114726
13	AES192/ECB	0.614101	95	3DES196/CFB	41.119764	177	Serpent192/CBC	45.202350
14	AES192/OFB	0.619291	96	3DES196/CTR	46.498527	178	Serpent192/CFB-FIPS	43.75221
15	AES256/CBC-CTS	0.338464	97	3DES196/ECB	46.224414	179	Serpent192/CFB	43.25249
16	AES256/CBC	0.320875	98	3DES196/OFB	46.322780	180	Serpent192/CTR	41.160307
17	AES256/CFB-FIPS	0.345115	99	GOST/CBC-CTS	68.973789	181	Serpent192/ECB	45.400695
18	AES256/CFB	0.326408	100	GOST/CFB	65.687065	182	Serpent192/OFB	43.32949
19	AES256/CTR	0.326471	101	GOST/CFB-FIPS	63.760146	183	Serpent256/CBC-CTS	43.81338
20	AES256/ECB	0.303837	102	GOST/CFB	61.095157	184	Serpent256/CBC	45.093283
21	AES256/OFB	0.306319	103	GOST/CTR	63.581349	185	Serpent256/CFB-FIPS	43.808402
22	Blowfish/CBC-CTS	0.000000	104	GOST/ECB	62.208584	186	Serpent256/CFB	43.566295
23	Blowfish/CBC	0.000000	105	GOST/OFB	62.895808	187	Serpent256/CTR	41.644703
24	Blowfish/CFB-FIPS	0.000000	106	IDEA/CBC-CTS	1.945411	188	Serpent256/ECB	44.043636
25	Blowfish/CFB	0.000000	107	IDEA/CBC	1.862755	189	Serpent256/OFB	43.300835
26	Blowfish/CTR	0.000000	108	IDEA/CFB-FIPS	1.864797	190	SHACAL2/CBC-CTS	43.216129
27	Blowfish/ECB	0.000000	109	IDEA/CFB	1.786869	191	SHACAL2/CBC	43.502398
28	Blowfish/OFB	0.000000	110	IDEA/CTR	1.809507	192	SHACAL2/CFB-FIPS	43.350775
29	Camellia128/CBC-CTS	2.439061	111	IDEA/ECB	1.797107	193	SHACAL2/CFB	43.297259
30	Camellia128/CBC	2.223550	112	IDEA/OFB	1.794513	194	SHACAL2/CTR	41.20957
31	Camellia128/CFB-FIPS	2.161228	113	MARS/CBC-CTS	35.413582	195	SHACAL2/ECB	43.63086
32	Camellia128/CFB	1.851554	114	MARS/CBC	32.522876	196	SHACAL2/OFB	42.297022
33	Camellia128/CTR	2.268503	115	MARS/CFB-FIPS	33.081898	197	SHARK/CBC-CTS	42.88837
34	Camellia128/ECB	2.124675	116	MARS/CFB	28.138252	198	SHARK/CBC	45.730335
35	Camellia128/OFB	2.128995	117	MARS/CTR	33.176147	199	SHARK/CFB-FIPS	43.737844
36	Camellia128/CFB	2.128995	118	MARS/ECB	31.142896	200	SHARK/CFB	44.303417
37	Camellia192/CBC	6.379568	119	MARS/OFB	31.031431	201	SHARK/CTR	42.002407
38	Camellia192/CFB-FIPS	6.183089	120	RC2/CBC-CTS	73.846996	202	SHARK/ECB	41.805070
39	Camellia192/CFB	5.296762	121	RC2/CBC	69.628161	203	SHARK/OFB	41.339174
40	Camellia192/CTR	6.496791	122	RC2/CFB-FIPS	65.768588	204	SKIPJACK/CBC-CTS	41.355590
41	Camellia192/ECB	6.080542	123	RC2/CFB	63.097860	205	SKIPJACK/CBC	86.800675
42	Camellia192/OFB	6.106682	124	RC2/CTR	67.867250	206	SKIPJACK/CFB-FIPS	78.144263
43	Camellia256/CBC-CTS	4.368407	125	RC2/ECB	66.895077	207	SKIPJACK/CFB	66.851360
44	Camellia256/CBC	3.989085	126	RC2/OFB	67.335425	208	SKIPJACK/CTR	63.444791
45	Camellia256/CFB-FIPS	3.870495	127	RC5*/CBC-CTS	59.065524	209	SKIPJACK/ECB	75.378074
46	Camellia256/CFB	3.310211	128	RC5*/CBC	56.995883	210	SKIPJACK/OFB	75.519345
47	Camellia256/CTR	4.062092	129	RC5*/CFB-FIPS	59.004923	211	ThreeWay/CBC-CTS	74.689386
48	Camellia256/ECB	3.799487	130	RC5*/CFB	56.631110	212	ThreeWay/CBC	64.465559
49	Camellia256/OFB	3.813245	131	RC5*/CTR	55.327618	213	ThreeWay/CFB-FIPS	62.202548
50	CAST128*/CBC-CTS	5.893581	132	RC5*/ECB	55.093507	214	ThreeWay/CFB	65.142521
51	CAST128*/CBC	56.342755	133	RC5**/OFB	55.032004	215	ThreeWay/CTR	56.11431
52	CAST128*/CFB-FIPS	57.251350	134	RC5**/CBC-CTS	59.065524	216	ThreeWay/ECB	60.035728
53	CAST128*/CFB	54.767648	135	RC5**/CBC	56.995883	217	ThreeWay/OFB	63.860083
54	CAST128*/OFB	54.994768	136	RC5**/CFB-FIPS	59.004923	218	Twofish128/CBC-CTS	59.939545
55	CAST128*/ECB	54.410741	137	RC5**/CFB	56.631110	219	Twofish128/CBC	0.000000
56	CAST128**/OFB	54.418252	138	RC5**/CTR	55.327618	220	Twofish128/CFB-FIPS	0.000000
57	CAST128**/CBC-CTS	40.527335	139	RC5**/ECB	55.093507	221	Twofish128/CFB	0.000000
58	CAST128**/CBC	38.706274	140	RC5**/OFB	55.032004	222	Twofish128/CTR	0.000000
59	CAST128**/CFB-FIPS	39.330460	141	RC5**/CBC-CTS	59.065524	223	Twofish128/ECB	0.000000
60	CAST128**/CFB	37.624209	142	RC5**/CBC	56.995883	224	Twofish128/OFB	0.000000
61	CAST128**/CTR	37.780236	143	RC5**/CFB-FIPS	59.004923	225	Twofish192/CBC-CTS	0.000000
62	CAST128**/ECB	37.379022	144	RC5**/CFB	56.631110	226	Twofish192/CBC	0.000000
63	CAST128**/OFB	37.384182	145	RC5**/CTR	55.327618	227	Twofish192/CFB-FIPS	0.000000
64	CAST256/CBC-CTS	23.094559	146	RC5**/ECB	55.093507	228	Twofish192/CFB	0.000000
65	CAST256/CBC	21.170467	147	RC5**/OFB	55.032004	229	Twofish192/CTR	0.000000
66	CAST256/CFB-FIPS	20.882287	148	RC6128/CBC-CTS	41.101655	230	Twofish192/ECB	0.000000
67	CAST256/CFB	17.799309	149	RC6128/CBC	38.453726	231	Twofish192/OFB	0.000000
68	CAST256/CTR	21.689633	150	RC6128/CFB-FIPS	40.461357	232	Twofish256/CBC-CTS	0.000000
69	CAST256/ECB	20.327965	151	RC6128/CFB	34.707866	233	Twofish256/CBC	0.000000
70	CAST256/OFB	20.352091	152	RC6128/CTR	39.363245	234	Twofish256/CFB-FIPS	0.000000
71	DES/CBC-CTS	57.892345	153	RC6128/ECB	37.073234	235	Twofish256/CFB	0.000000
72	DES/CBC	54.911989	154	RC6128/OFB	36.808767	236	Twofish256/CTR	0.000000
73	DES/CFB-FIPS	54.683737	155	RC6192/CBC-CTS	47.569456	237	Twofish256/ECB	0.000000
74	DES/CFB	52.598599	156	RC6192/CBC	44.492275	238	Twofish256/OFB	0.000000
75	DES/CTR	53.955732	157	RC6192/CFB-FIPS	46.860443	239	XTEA/CBC-CTS	0.000000
76	DES/ECB	53.268029	158	RC6192/CFB	40.179859	240	XTEA/CBC	1.363389
77	DES/OFB	53.254812	159	RC6192/CTR	45.589953	241	XTEA/CFB-FIPS	1.300107
78	3DES64/CBC-CTS	52.413608	160	RC6192/ECB	42.889572	242	XTEA/CFB	1.289004
79	3DES64/CBC	47.689617	161	RC6192/OFB	42.689038	243	XTEA/CTR	1.234000
80	3DES64/CFB-FIPS	43.039922	162	RC6256/CBC-CTS	50.047571	244	XTEA/ECB	1.261434
81	3DES64/CFB	41.251006	163	RC6256/CBC	46.789653	245	XTEA/OFB	1.258946
82	3DES64/CTR	45.586229	164	RC6256/CFB-FIPS	46.311394	-	-	-

Cipher specification: 1) CAST-128*/CAST-128** is for key length from 40 to 80/80 to 128 bits; 2) RC5*/RC5**/RC5*** is for plaintext length 32/64/128 bits.

Moreover, these formulas are developed so that their value can be taken as a standard, since even when the system or machine, OS, CPU, *etc.* changes, the result is not very much affected (change that does not exceed 0,3%). Our future work will concern two paths:

- The first will focus on obtaining more ciphers or algorithms using-cost measurement.
- The second will concentrate on getting deeper in risk analysis study over cipher.

Acknowledgments

The authors are grateful to the anonymous reviewers for valuable comments.

References

- [1] E. Biham and A. Shamir, "Differential cryptanalysis of DES-like cryptosystems," *Journal of Cryptography*, vol. 4, no. 1, pp. 3–72, 1991.
- [2] A. Biryukov, "Meet-in-the-middle attack," in *Encyclopedia of Cryptography and Security*, pp. 772–773, 2011.
- [3] A. Biryukov, D. Khovratovich, and I. Nikolić, "Distinguisher and related-key attack on the full AES-256," in *Advances in Cryptology (CRYPTO'09)*, pp. 231–249, 2009.
- [4] A. Biryukov and D. Wagner, "Slide attacks," in *International Workshop on Fast Software Encryption*, pp. 245–259, 1999.
- [5] C. Blondeau, *La Cryptanalyse Différentielle Et Ses Généralisations*, Thesis, Université Pierre et Marie Curie-Paris VI, Nov. 2011.
- [6] C. Cid and G. Leurent, "An analysis of the XSL algorithm," in *International Conference on the Theory and Application of Cryptology and Information Security*, pp. 333–352, 2005.
- [7] S. Dey and R. Ghosh, "A review of cryptographic properties of S-Boxes with generation and analysis of Crypto secure S-Boxes," *International Journal of Electronics and Information Engineering*, vol. 8, no. 1, pp. 49–73, 2018.
- [8] I. R. Dragomir and M. Lazăr, "Generating and testing the components of a block cipher," in *8th International Conference on Electronics, Computers and Artificial Intelligence*, pp. 1–4, June 2016.
- [9] A. B. Forouzan, *Data Communications and Networking*, McGraw-Hill Forouzan networking series, McGraw-Hill Higher Education, 2007.
- [10] W. F. Friedman, *The Index of Coincidence and Its Applications in Cryptanalysis*, Aegean Park Press California, 1987.
- [11] B. Gérard, *Statistical Cryptanalyses of Symmetric-Key Algorithms*, Thesis, Université Pierre et Marie Curie-Paris VI, Dec. 2010.
- [12] Y. Harmouch and R. E. Kouch, "A fair comparison between several ciphers in characteristics, safety and speed test," in *Europe and MENA Cooperation Advances in Information and Communication Technologies*, pp. 535–547, 2017.
- [13] V. T. Hoang and P. Rogaway, "On generalized feistel networks," in *Annual Cryptology Conference*, pp. 613–630, 2010.
- [14] M. S. Hwang, C. T. Li, J. J. Shen, Y. P. Chu, "Challenges in e-government and security of information", *Information & Security: An International Journal*, vol. 15, no. 1, pp. 9–20, 2004.
- [15] J. Kelsey, B. Schneier, and D. Wagner, "Mod n cryptanalysis, with applications against RC5P and M6," in *International Workshop on Fast Software Encryption*, pp. 139–155, 1999.
- [16] M. C. Lee, "Information security risk analysis methods and research trends: Ahp and fuzzy comprehensive method," *International Journal of Computer Science and Information Technology*, vol. 6, pp. 29–45, Mar. 2014.
- [17] M. Matsui, "Linear cryptanalysis method for des cipher," in *Workshop on the Theory and Application of Cryptographic Techniques*, pp. 386–397, 1993.
- [18] H. H. Ngo, X. Wu, P. D. Le, C. Wilson, and B. Srinivasan, "Dynamic key cryptography and applications," *International Journal of Network Security*, vol. 10, pp. 161–174, May 2010.
- [19] H. Sato, "A new formula of security risk analysis that takes risk improvement factor into account," in *IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pp. 1243–1248, Oct. 2011.
- [20] N. Shukla and S. Kumar, "A comparative study on information security risk analysis practices," in *Issues and Challenges in Networking, Intelligence and Computing Technologies (ICNICT'12)*, pp. 28–33, Nov. 2012.
- [21] O. Tornea, *Contributions to DNA Cryptography: Applications to Text and Image Secure Transmission*, PhD Thesis, Université Nice Sophia Antipolis, 2013.
- [22] D. Wagner, "The boomerang attack," in *International Workshop on Fast Software Encryption*, pp. 156–170, 1999.
- [23] C. Wang and W. A. Wulf, "Towards a framework for security measurement," in *20th National Information Systems Security Conference, Baltimore*, pp. 522–533, 1997.
- [24] J. Wang, K. Fan, W. Mo, and D. Xu, "A method for information security risk assessment based on the dynamic bayesian network," in *International Conference on Networking and Network Applications*, pp. 279–283, July 2016.
- [25] Y. Yeom, "Integral cryptanalysis and higher order differential attack," *Trends in Mathematics*, vol. 8, no. 1, pp. 101–118, 2005.

[26] A. Yun, J. H. Park, and J. Lee, "On lai-massey and quasi-feistel ciphers," *Designs, Codes and Cryptography*, vol. 58, no. 1, pp. 45-72, 2011.

Appendix-A: The Computation of Break Probability for Each Structure and Operation

Let us consider two random variables X and K with i resp. j is a number while p and q are probabilities. Now, assume that the distribution probability is given by:

$$Pr(X) = \begin{cases} p, & X = i \\ 1 - p, & X \neq i \end{cases} \text{ and } Pr(K) = \begin{cases} q, & K = j \\ 1 - q, & K \neq j \end{cases}$$

X and K are two independent variables, Thus:

$$Pr(X, K) = \begin{cases} pq, & X = i, K = j \\ p(1 - q), & X = i, K \neq j \\ q(1 - p), & X \neq i, K = j \\ (1 - p)(1 - q), & X \neq i, K \neq j \end{cases}$$

■ Left/Right rotation:

Let us consider $f:\{0, 1\}^n \times GF(2^8) \rightarrow \{0, 1\}^n$ f can be described as $f(\xi, \phi) \rightarrow \xi'$ where f is a function, ξ is a binary-vector, ϕ is a number with $dim\xi > \phi$ and ξ' is the binary-vector results. We denote by $dim\xi$ the size of the vector ξ .

Given f , we have $f(\xi, \phi) = \xi \begin{pmatrix} \ll \\ or \\ \gg \end{pmatrix} \phi$ where " \ll "

resp. " \gg " indicates left resp. right rotation. Of note, f is itself invertible with $Pr(f = 1) = 1/\phi$ because ξ' has ϕ possibilities. ϕ is unknown, hence $Pr(f = 1) = \frac{1}{dim\xi - 1}$

■ NOT:

Let us consider $f:\{0, 1\} \rightarrow \{0, 1\}$ where f is the bitwise NOT function. For such function, we have $f(x) = \bar{x}$ with x is a binary variable. The probability of guessing the result is equal to $Pr(f = 1) = p + (1 - p) = 1$, thus, for the general case (binary vector) we have $Pr(f^n) = \prod_{i=1}^n [p + (1 - p)] = 1$.

■ Conditional NOT:

Let us consider $f:\{0, 1\}^2 \rightarrow \{0, 1\}$ where f is the bitwise conditional-not-function. Given f , we write $f(x, k) = y$ with x, k and $y \in \{0, 1\}$. To show y , let us consider that f applies "not" to x if k is true, so we have $f(x, k) = \bar{x}k + x\bar{k}$ where "+" indicates logical addition. As observed, f is equivalent to XOR function, so Conditional Not and XOR have the same probabilities (see below for more details).

■ AND:

Let us consider $f:\{0, 1\}^2 \rightarrow \{0, 1\}$ where f is the bitwise AND function. For such function, we have

$f(X, K) = X \times K$ where " \times " indicates "AND" and $X, K \in \{0, 1\}$. Note that f is not invertible, this implies that even if X is found, K cannot be known (vice versa). Consequently, the only possible case of breaking f is to know both X and K. Therefore $Pr(f = 1) = Pr(X, K) = pq = 1/2^2$. As for the general case (X and K are binary vector) we have $Pr(f^n) = Pr(X, K) = (\prod_{i=1}^n \frac{1}{\#f})^2 = 1/2^{2n}$ with $\#$ denotes the set cardinal.

■ OR:

Let us consider $f:\{0, 1\}^2 \rightarrow \{0, 1\}$ where f is the bitwise OR function. For such function, we have $f(X, K) = X + K$ where "+" indicates "OR" and $X, K \in \{0, 1\}$. Remark that f is not invertible, this implies that even if X is found, K cannot be known (vice versa). Consequently, the only possible case of breaking f is to know both X and K. thus $Pr(f = 1) = Pr(X, K) = pq = 1/2^2$. As for the general case (X and K are binary vector) we have $Pr(f^n) = Pr(X, K) = (\prod_{i=1}^n \frac{1}{\#f})^2 = 1/2^{2n}$.

■ XOR:

Let us consider $f:\{0, 1\}^2 \rightarrow \{0, 1\}$ where f is the bitwise XOR function. For such function, we have $f(X, K) = X + K$ where "+" indicates mod2 addition and $X, K \in \{0, 1\}$. Given these two variables, f can only present one of the following two scenarios: $f:\{0, 1\}^2 \rightarrow \{0\}$ is a linear expression and is equivalent to $X = K$ and $f:\{0, 1\}^2 \rightarrow \{1\}$ is an affine expression and is equivalent to $X \neq K$. Since $Pr(X = K) = Pr(X = 0, K = 0) + Pr(X = 1, K = 1)$ and $Pr(X \neq K) = Pr(X = 0, K = 1) + Pr(X = 1, K = 0)$, we have $Pr(X = K) = pq + (1 - p)(1 - q)$ and $Pr(X \neq K) = p(1 - q) + q(1 - p)$. Moreover, f is invertible. This implies that knowing one variable from those defined above led to know the second. Therefore the probability of breaking f is $Pr(X|K)$. This can be solved by: $Pr(X|K) = \frac{Pr(X \cap K)}{Pr(K)} = \frac{Pr(X) \times Pr(K)}{Pr(K)} = Pr(X) = p = 1/2$. for the general case (X and K are binary vector) we have $Pr(f^n) = Pr(X, K) = \prod_{i=1}^n 1/2 = 1/2^n$.

■ Concatenation:

Let us consider $f:\{0, 1\}^2 \rightarrow \{0, 1\}$ where f is the concatenation function. For such function, we have $f(X, K) = X \| K$ where " $\|$ " indicates the concatenation-operation and $X, K \in \{0, 1\}$. The probability of guessing X and K from f is equal to $Pr(f) = p + q = 1$, hence for the general case is equal to $Pr(f^n) = Pr(f) = 1$. If X and K were a binary vectors with unequal or unknown size then we will have $Pr(f) = \frac{1}{\#f - 1}$

■ Modular addition:

The result proved in XOR can be generalized to modular addition since XOR is mod 2 addition case. Thus, for f defined as $f:\{0, 1, \dots, \xi - 1\}^2 \rightarrow$

$\{0, 1, \dots, \xi - 1\}$ the break probability is equal to $Pr(f^n) = Pr(X, K) = \prod_{i=1}^n \frac{1}{\#f} = 1/\xi^n$.

■ **Modular subtraction:**

The result proved in modular addition is the same as modular subtraction since “+” and “-” has the same break probabilities, thus for f defined as $f:\{0, 1, \dots, \xi-1\}^2 \rightarrow \{0, 1, \dots, \xi-1\}$ the break probability is equal to $Pr(f^n) = Pr(X, K) = 1/\xi^n$.

■ **Modular multiplication:**

Let us consider $f:GF(\xi)^2 \rightarrow GF(\xi)$ where f is the modular multiplication function and ξ is the modulus. For such function, we have $f(X, K) = X \times K$ where “ \times ” indicates multiplication mod ξ and $X, K \in GF(\xi)$. Notice that f is not itself invertible, it implies that to find X, K should have a modular multiplicative inverse K' . i.e. $K \times K' \equiv 1(mod\xi)$. Consequently, two scenarios are possible: either K admits a modular multiplicative inverse thus K and ξ are coprime or K do not admits a modular multiplicative inverse. These scenarios shows that the found of one variable X or K do not help of guessing the other one. Thus, the only possible case to break f is $Pr(X,K)=pq$. As so, for the general case, we have $Pr(f^n) = (\prod_{i=1}^n \frac{1}{\#f})^2 = 1/\xi^{2n}$.

■ **Modular exponentiation:**

The modular exponentiation is a special case of the modular multiplication where knowing both X and K is the only way to break the operation, thus $Pr(f^n) = 1/\xi^{2n}$.

■ **P-box:**

Let us consider $f:\{0, 1\}^n \rightarrow \{0, 1\}^n$ where f is a permutation function (P-box). Given f , we write $f(x_0, x_1, \dots, x_{n-1})=(x_i, \dots, x_j, \dots, x_k)$ where $i, j, k \in [0, n-1]$. If we consider f as a black box (dynamic P-box) where the linear link between input and output is not known, the breaking probability for f for a binary vector X is $Pr(f) = Pr(X) = \prod_{i=1}^{n-1} \frac{1}{p} = \prod_{i=1}^{n-1} \frac{1}{\#f} = 1/2^{n-1}$. As for the static P-box where the linear link between input and output is exactly known, we have $Pr(f) = Pr(X) = \prod_{i=1}^{n-1} \frac{1}{p(1-p)} = 1$.

■ **S-box:**

Let us consider $f:GF(\xi)^n \rightarrow GF(\xi)^m$ where f is a substitution function and 2 is the modulus. Given f , we write $f(x_0, x_1, \dots, x_{n-1})=(y_0, y_1, \dots, y_{n-1})$. For instance, the AES S-box is written as $f(x_i)=\sum_{u \in GF(2)^n} a_u \prod_{i=1}^n x_i^{u_i}, a_u \in GF(2)^n$.

Thereby, this equation can be denoted as $f(x_i)=(l \circ h)$, where “ l ” indicates the $n \times m$ binary matrix and “ h ” is a function. For example, $h(x)$ in AES is equal

$$\text{to } h(x) = \begin{cases} x^{-1}, & X \neq 0 \\ 0, & X = 0 \end{cases}$$

Thus, as shown by Liam Keliher in “Linear Cryptanalysis of Substitution-Permutation Networks” in

ch.4, the probability for breaking the S-box is $Pr(f)=\frac{1}{2^n-1}$.

■ **Feistel:**

Let us consider $f:\{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}^n$ where f is a Feistel function and $m < n$. Given f , we write $f(X,K)$ with X, K are two binary vector and $f(X,K) = \begin{cases} x_{m+i}, & 1 \leq i \leq m \\ x_{i-m} \oplus G(x_i, k_{i-m}), & m < i \leq n \end{cases}$ with G is a round function.

Since f admits a linear liaison for $n-m$ random binary variables, the security for this structure is built over K , and the only possible case to break f is by guessing K , as so $Pr(f)=pq+q(1-p)=q=\prod_{i=1}^m \frac{1}{2} = 1/2^m$. It must be mention that in the case of $m=n/2$ the Feistel structure is called balanced Feistel function, otherwise, it is called unbalanced Feistel function and the probability turn to be equal to $Pr(f)=\frac{1}{2^{n-m}}$.

■ **Lai-Massey:**

Let us consider $f:\{0, 1\}^n \times \{0, 1\}^m \rightarrow \{0, 1\}^n$ where f is a Lai-Massey function and $m < n$. Given f , we write $f(X,K)$ with X, K are two binary vector and $f(X,K) = \begin{cases} \sigma(x_i + G(x_i - x_{\frac{n}{2}+i}, k_j)) & 1 \leq i \leq \frac{n}{2} \text{ and} \\ x_{\frac{n}{2}+i} + G(x_i - x_{\frac{n}{2}+i}, k_j) & 1 \leq j \leq m. \end{cases}$

G is a round function and σ is an orthomorphism permutation (in mathematical sense, that is, a bijection not a P-box). The Lai-Massey schema differs from Feistel schema, because it modifies both the left half and the right half of the plaintext block. Thus the security for this structure is built over K and P . Therefore the only possible case to break f is by guessing either X or K , as so $Pr(f)=\prod_{i=1}^n \frac{1}{2}[q(1-p) + p(1-q)] = \prod_{i=1}^n \frac{1}{2}[p+q-2pq]$ and since $p=q=1/2 \Rightarrow Pr(f)=\prod_{i=1}^n \frac{1}{2} = 1/2^n$.

Biography

Youssef Harmouch is a Ph.D. student at National Institute of Post & Telecommunication INPT-Rabat Morocco. He started his career in 2012 as a network and telecommunications engineer Specialized in VOIP and information security. Since 2014, he returned to INPT as a PhD candidate working in fields of cryptography within “Multimedia, Signal And Communication Systems” Laboratory, where he focus on cryptographic schema, cipher design, cryptanalysis Study, risk analysis, advanced mathematical theory precisely in algebra, chaos and coding theory.

Rachid Elkouch is a Professor and the Manager of PABX Laboratory, attached to Systems and Communications Department at INPT Since 1981, he started his carrier as a telecommunications engineer, and then in 1989 he got his Master’s degree of Science in Telecommunications from the University of Colorado-Boulder in the USA. In 2005, he obtained his doctorate degree

in applied mathematical from Mohammed Ben Abdellah University-Faculty of Sciences and Techniques of Fez-USMBA. Meanwhile, he continued working as an assistant for the INPT Electricity and electronics laboratory for five years since 1981. Between 1990 and 1992, he became a specialized tutor for a number of technical inspectors. More specifically, it was a technological transfer program within the framework of a convention between France & Morocco. Since 1992, he was in charge of the Engineering Cycle Internship and then officially the Manager of the Internships Service since 1998. He was then promoted to the position of Deputy Director of Internships and International Relations between the INPT and outside companies since 2008. In 2010, he was assigned to the position of Deputy Director of Continuing Education. He is a member of the research team Phare (UFR MDA) from the Laboratory of Computer Science and Mathematics from USMBA, a tutor of 2 modules on the E-platform Miage (Lyon 1): B202-Basics of Telecommunications and B214-Networks Protocols.

Hussain Ben-azza is a Professor at ENSAM-National High School of Arts and Trades, Meknes, Morocco, attached to Department of Industrial and Production Engineering, Moulay Ismail University. He obtained his Ph.D. degree in mathematics and computer science in 1995 from Claude Bernard University Lyon 1, France. His research interests include coding theory, cryptography, wireless communications, but also applications of optimization techniques to industrial engineering.