Secure RFID System Based on Lightweight Block Cipher Algorithm of Optimized S-Box

HUI Yue-chao and WANG Yi-ming

***Abstract*—As a non-contact automatic identification technology, radio frequency identification (RFID) technology plays an increasingly important role in social life. But the existence of potential safety problems has restricted the widespread application of RFID technology. In order to deal with the design challenges in the calculation of speed, communications capability and storage space to the low-cost radio frequency tags, this paper illustrates a new secure communication model based on the existing RFID security protocols. Through comparing and analyzing of the existing RFID tag security policies appeared in many papers, a new lightweight encryption algorithm (LEA) has been proposed based on the criterion of block cipher, especially the improving DES encryption algorithm. It does not require complex computing and high-strength encryption technology. The core components S-boxes are redesigned and implemented on FPGA in this paper, and the security performance of S-boxes is proved through the analysis of probability and statistics. This algorithm can effectively enhance the security in the low cost RFID system.**

I. Introduction

RADIO Frequency Identification System (RFID) has been widely used in industrial production and daily life in all areas, such as the automatic tariffs, supply chain and library management, tracking of product, automated manufacturing, and building access control. However, several security and privacy concerns have been identified in connection with the use of RFID. There are some worth mentioning problems: tracking or violation of location privacy, physical attacks, denial of service, counterfeiting, spoofing, eavesdropping, traffic analysis,etc. Now days, the security and privacy have become some of the most important restrained factors before RFID systems massive deployment [1].

Typical RFID system always consists of three components called RFID tag (transponder), RFID reader (transceiver), and back-end database [2]. RFID tags attached to products are used to identify the object during production or in uses via radio frequency which may be passive or active.

The security research of RFID system is focused on how to solve information security problems between the tag and reader communication. Low-cost RFID tags are extremely limited in storage space. The cheapest tag's ROM is only 64-128bits.It only can accommodate a unique identifier. Also RFID tag's power supply and computing capacity are very limited. These limitations of the low-cost RFID system require the special requirements in security mechanism designing. The choices of security mechanisms are subject to many restrictions. Sarma pointed out that the designing of low-cost RFID security systems must consider two points [3]: the limited computing resources and other networks or systems that RFID systems often interconnect with. So far, the security research based on RFID system mainly remains in theory at home and abroad, especially for the security issues of safe and efficient in low-cost RFID system. Therefore, the study of RFID security technology has a high application value under the conditions of limited hardware resources.

This paper first provides some design ideas of block cipher, then illustrates a new secure communication model based on the existing RFID security system, which meets the EPCglobal protocol. At last a lightweight block cipher algorithm is proposed based on the improved DES algorithm and features of block cipher. For controlling complexity, the encrypted data and key bits are halved in the algorithm. It uses the method of adding bit selection logic traps to increase the difficulty of deciphering the code when selecting the data in S-boxes. Finally, S-boxes are redesigned by the non-linear principle and the security performance of S-boxes is analyzed.

II. RFID Security Program

RFID security mechanisms are mostly researched based on cryptographic techniques. Today It has come up a variety of security protocols, such as the Hash-Lock protocol, randomized Hash-Lock protocol, distributed RFID inquiry-response authentication protocol, etc.[4][5].These methods make up some RFID security vulnerabilities, enhance system security to a certain extent, but there are still have security vulnerabilities. In the Hash-Lock protocol RFID tag's ID plaintext transmission is still exist, an attacker can easily obtain the ID for forgery, and also easy to track tag.

We have designed a secure authentication protocol for the shortcomings of the current secure programs. This protocol not only able to complete key agreement and key updates in the certification process, but also can reuse the hardware resources in tags. So it can achieve the aim to improve resource utilization efficiency and reduce cost, also can support safety performance. The protocol can resist wire-tapping, synchronization destruction, playback and location tracking threats,etc. The modified protocol is composed of *10* steps, which are depicted in Fig. 1, and described as follows:

HUI Yue-chao and WANG Yi-ming are with School of Electronics and Information Engineering,Soochow University, China （ email ： huiyuechao123@163.com, ymwang@suda.edu.cn）

978-1-4244-6700-6/10/$26.00 © 2010 IEEE

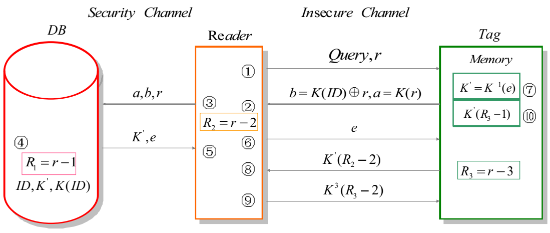

Fig. 1. The basic authentication step of the proposed protocol

TABLE I
THE NOTATION FOR MODIFIED PROTOCOL

$A \to B$	A sends message to B
ID	Tag's id
r	The random number generated by a reader at first.
K	The symmetric key generated by Encryption Algorithm
K'	The new key generated by back-end database
$K(x)$	Encryption of message x with symmetric key K.
$K^{-1}(x)$	Decryption of message x with symmetric key K.
$\oplus$	XOR operation
$-x$	Left shift x bits
R_1	$R_1 = r-1$,which is generated by back-end database
R_2	$R_2 = r-2$, which is generated by reader
R_3	$R_3 = r-3$,which is generated by tag

Step 1. To initiate the communications session, the reader sends an inquiry signal "Query" and produces the random number r from R.N.G which has the same size with tag's "ID"

Step 2. The tag that has received signals from reader encrypts "ID" with the original key "K" in memory and uses an XOR algorithm with the acquired $r: b=K(ID)\oplus r$, this can change the data in transmission to prevent the threat of tracking. Then the tag calculates $a=K(r)$ and updates r , $r=r-1$ (this is R_1). Send a,b to the reader.

Step 3. The reader sends the acquired a,b with own random number r to the back-end database through a secure channel [6].

Step 4. The database conducts an XOR algorithm on the received b and r to produce $K(ID)$, and searches for the corresponding number and key based on $K(ID)$, then makes encryption with the acquired r and key $K: a'=K(r)$. If the received a is identical to the a', which the database contains, the database authenticates the tag, and acquires a new session key K' from the database, also produces a number R_1 , $R_1=r-1$,which has the same bits with K' to ensure the safety of the former, calculates $e=K(ID\oplus K'\oplus R_1)$

Step 5. The back-end database will send K',e to the reader.

Step 6. The reader serves K' as an important key to communicate with tag. At the same time the reader produces the second number $R_2(R_2=r-2)$ and calcu -lates $K'(R_2)$ to send it to the tag with e.

Step 7. The tag decrypts e with the old key K to get $ID\oplus K'\oplus R_1$, and uses an XOR algorithm with ID and $R_1(R_1=r-1)$ to get the new key K' , but it does not immediately repeal the old key, on the contrary it is retained. Because it is in order to prevent the synchronization to destroy between database and tag. It does not repeal the old key K until the final certification is over. Then tag updates r , let $r=r-2$ (R_2).

Step 8. r has been updated R_2 in **Step 7**, so it can directly calculate $K'(R_2-2)$,It is reason to make that operation to prevent replay attacks. Then tag produces a number $R_3(R_3=r-3)$, and sends $K'(R_2-2)$ to reader.

Step 9. When reader receives data, first calculates $K'(R_2-2)$. If it is identical to the received, the reader authenticates the tag once more. Because r in the tag has been updated as R_3 in **Step 6**, then calculates $K'(R_3-2)$ to be transmitted to the tag.

Step 10. RFID tag calculates $K'(R_3-2)$ and compares it with the information that has received, if they are equal, this means the tag makes a successful authentication to reader. It considers the reader is legal. Now the certification is over between reader and tag. At last the database and tag destruct the old key K .

III. LIGHTWEIGHT BLOCK CIPHER ALGORITHM

A. Review of block cipher algorithm

The security is the most important design criterion in Block cipher. Even if the attackers know the internal structure of block cipher, he can not decipher the password. The security of Block cipher depends largely on the keys length, blocks length and algorithm structure, the corresponding criteria for their safety as follows [7][8]:

The design criterion of key length is to enable the algorithm against powerful attacks, so it is necessary to ensure the greatest possible length of the key. In general, the key length is determined by the following factors: the value of the protected data itself, the data protection period, the available resources of an attacker, and the application areas of the cipher. Therefore, in low-cost systems, if the key is too long, it will cause exact costs to the key management and the efficiency of operation will also fall.

The design criterion of block length is that block length must be sufficiently large in order to prevent statistical analysis. $2^n(n=5,6,7)$ bits of the packet are usual.

Confusion should ensure that the dependence of plain text and cipher text is very complex so that an attacker can not use the features. Proliferation does not make the statistical relationship exist between the simple structures of plain text and cipher text. This relationship also does not exist between different encryption functions. Differential cryptanalysis is one of the most effective ways to attack the cryptosystem [9]. The basic idea is to restore some of the key bits by analyzing the impact of plain text to cipher text. The basic idea of linear cryptanalysis is to decipher cryptosystem by finding the effective linear appro -ximation expression of a given cryptographic alg -orithm [10].In our algorithm choose values in S-boxes by increasing logic trap, and then the values are designed based on non-linear criteria.

B. Lightweight block cipher algorithm designing

In this algorithm the key k is a 28 bits string of symbols 0,1.The plaintext m and ciphertext c are a 32-bit string of 0,1. Set $k = k_1 k_2 \cdots k_{28}$, $m = m_1 m_2 \cdots m_{32}$, $c = c_1 c_2 \cdots c_3$.As a result of the initial permutation and inverse permutation does not affect the encryption, it can be canceled. The encryption process can be expressed as follows: $DES(m) = T_{16} \cdot T_{15} \cdots \cdot T_2 \cdot T_1(m)$ .The encrypt -tion process can be composed of plaintext initialized, iterative process and the program of key generation.

First, change character strings into ASCII codes which are expressed as 8-bit binary strings using 0, 1. Select 32 bits as a block in sequence, so the plaintext can be divided into n blocks. If the last block is not 32 bits enough, it can be added the randomly generated strings of 0,1.But the recipient must clear what are superfluous.

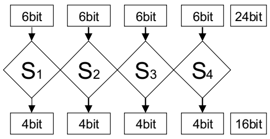

Fig.2. The output through four S-boxes

The iterative process is similar with the DES iterative process in this algorithm. First, the 32-bit block which is to be encrypted is divided as L_0 、 R_0 , they have equal length. The iterative formula: $L_i = R_{i-1}$, $R_i = L_{i-1} \oplus f(R_{i-1}, k_i)$. Then 16-bit R_{i-1} is expanded 24-bit by the E-bit selection table ,the results will make XOR computation with k_i , then the results will be divided into 4 groups, the output through S-boxes is 16-bit (Fig.2).The output is $f(R_{i-1}, k_i)$ after a fixed permutation, Through eight iterations gets $L_8 R_8$,swap $L_8 R_8$ is $R_8 L_8$,this is ciphertext.

Each round uses the 24-bit key $k_i (1 \le i \le 8)$ that derived from the initial key k . k is got from a length of 32-bit string by deleting four $(8th, 16th, 24th, 32nd)$ parity bits. The parity bits can check errors.

IV. S-BOX DESIGNING

A. Methods of Selecting S-box Data

Generally, S-box output data is determined by the sixbits input. The first and last bits choose the row, and the middle four bits choose the column. In our algorithm by adding logic traps to select the data. This way can not only simplify operations but also program conveniently. We can modify logic gates to improve the security performance of the algorithm in a certain period. Specific processes are:

1) In the iterative process after computing $R_{i-1} \oplus E$, the results of 24-bit string are divided into four 4-bit strings. They can be recorded: $B = B_1 B_2 B_3 B_4$.
2) With S_1 S_2 S_3 S_4 four S-boxes, each S-box is a fixed 4×16 matrix. To a 6-bit string, for example: $B_1 = b_1 b_2 b_3 b_4 b_5 b_6$,then output $S_1(B_1)$ is determined as follows: The second and third bits of B_1 can be expressed in binary a number in the range 0 to 3. Let that number be r. $b_1 \oplus b_4$, b_4 , b_5 , b_6 four binary bits can be expressed a number in the range 0 to 15. Let that number be c. Look up in the table the number in the r'th row and c'th column. It is a number in the range 0 to 15 and can be represented by a 4 bit block in binary. Four S-boxes output bit blocks can be record as $D_j = S_j(B_j), 0 \le j \le 3$.
3) At last the 16-bit string $D = D_1 D_2 D_3 D_4$ needs a fixed permutation, the results are $f(R_{i-1}, k_i)$ in the iteration process.

TABLE II gives the method of how to choose the data in four S-boxes.

TABLE II
THE METHOD OF CHOOSING DATA IN FOUR S-BOXES.

S-box	**Row(2 bits)**		**Column (4 bits)**			
DES-S	b_1	b_6	b_2	b_3	b_4	b_5
S_1	b_2	b_3	$b_1 \oplus b_4$	b_4	b_5	b_6
S_2	b_1	b_4	b_3	b_2	$b_5 \oplus b_6$	b_6
S_3	b_2	b_5	$b_1 \odot b_4$	b_3	b_4	b_6
S_4	b_1	b_6	b_2	$b_2 \oplus b_3$	b_4	b_5

B. Designing Data in S-boxes

S-box as the only non-linear devices in DES is the key to encrypt. However the criteria for the design of S-box have not been open over the years, we can see a number of criteria from the experience of designing S-boxes [11]:

8) Each line is an integer between 0-15 in S-box.
9) Each S-box is six inputs and four outputs.
10) S-box output is not a simple linear or affine function of the input.
11) If S-box input has one bit change, at least two bits of the output are changed.
12) If only the two bits in the middle of the S-box input are different, then two bits in output must be different at least.

Through *A* Analysis we know the methods of choosing bits in S-box have changed, so we must redesign S-boxes based on the nonlinear principle to meet the encryption requirements, The main nonlinear principle is that one bit input change brings at least two bits output change. First, we can assume that two input bits used choosing row remain unchanged. Then if four input bits used choosing column have one bit change, the corresponding data in S-box must have two bits change. TABLE III shows the selected relation as the example of S_1 box designing. In the table the codes with "□" are before the selected row, and the codes with "__" need not to consider which are same to the selected row. The other three S-boxes' designing is similar.

We can design S$_1$ box as an example. When $b_2b_3 = 00$, if $b_1b_4b_5b_6 = 0100$, so $b_1 \oplus b_4 = 1$、$b_4 = 1$、$b_5 = 0$、$b_6 = 0$，this means that the chosen data in row 0 and column 12. From TABLE Ⅲ can be seen that the data must have at least two different bits from the 0th, 5th and 6th columns. That is to say we have to choose the number in the intersection of the selected non-linear elements in column 0,5,6.

TABLE Ⅲ
THE CODE OF AT LEAST TWO DIFFERENT BITS FROM THE SELECTED COLUMN

Column Serial	Column Code	The code of at least two different bits from the selected column			
0	0000	0001	0010	0100	1000
1	0001	$\boxed{0000}$	0011	0101	1001
2	0010	$\boxed{0000}$	0011	0110	1010
3	0011	$\boxed{0001}$	$\boxed{0010}$	0111	1011
4	0100	$\underline{0100}$	1000	1101	1110
5	0101	$\underline{0101}$	1001	1100	1111
6	0110	$\underline{0110}$	1010	1100	1111
7	0111	$\underline{0111}$	1011	1101	1110
8	1000	$\boxed{0000}$	1001	1010	1100
9	1001	$\boxed{0001}$	$\boxed{1000}$	1011	1101
10	1010	$\boxed{0010}$	$\boxed{1000}$	1011	1110
11	1011	$\boxed{0011}$	$\boxed{1001}$	$\boxed{1010}$	1111
12	1100	$\boxed{0000}$	$\boxed{0101}$	$\boxed{0110}$	$\underline{1100}$
13	1101	$\boxed{0001}$	$\boxed{0100}$	$\boxed{0111}$	$\underline{1101}$
14	1110	$\boxed{0010}$	$\boxed{0100}$	$\boxed{0111}$	$\boxed{1101}$
15	1111	$\boxed{0111}$	$\boxed{0101}$	$\boxed{0110}$	$\underline{1111}$

When $b_2b_3 = 01$, because it has one bit different from $b_2b_3 = 00$,so it must be considered the relevance of the row 0 corresponded with. If $b_1b_4b_5b_6 = 0000$, then $b_1 \oplus b_4 = 0, b_4 = 0, b_5 = 0, b_6 = 0$, it is still the first one, so the value in row 1 and column 0 must be selected from the non-linear correlation table of row 0 and column 0 selected. If $b_1b_4b_5b_6 = 0100$, then the value in row 1 and column 12 must be selected from the non-linear intersection ,it is formed by the values of row 1, column 0,5,6 and row 1, column 12 intersect. Aslo the values that have been selected before must be excluded.

When $b_2b_3 = 10$, they have one bit different from 00,

So the selected method in this row is similar to the second row.

When $b_2b_3 = 11$, it has one different from 01 and 10, so you must consider the relations with column 1 and column 2. When $b_1b_2b_3b_4b_5b_6 = 011101$, that is to say the data in column 3 and row 13 has been chosen. We know form table II that this data must have at lease two bits different from the1th, 4th and 7th row, but also two bits are different from the data in row 1 column 13 and row 2 column 13. So you must select the data from the intersection of the five selected data.

TABLE Ⅳ
THE DATA OF S1 BOX

S$_1$	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	14	4	13	1	5	12	3	11	7	10	0	6	9	2	8	15
1	0	15	3	10	8	2	9	13	11	12	14	5	7	1	6	4
2	11	2	8	15	3	9	10	13	0	7	6	1	5	14	4	12
3	7	4	14	9	6	5	12	0	13	2	1	15	10	11	8	3

TABLE Ⅳ only gives the results of S$_1$ box.The selection method of row and column in S$_1$ box is shown in TABLE Ⅱ.The selection processes of other S-boxes are similar.

V. FPGA IMPLEMENTATION AND SECURITY ANALYSIS

A. *The Implementation of S-box*

The design uses device XC2V400 of Xilinx's low-cost Virtex2 series chip as the core. Verilog language on ISE9.1i [12][13] is used to achieve the design. Fig.3 and Fig.4 show the RTL structure figure (including the top and bottom module chart) and the waveform figure.

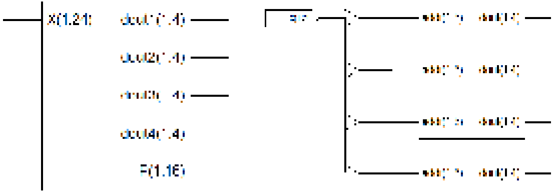

Fig.3. RTL structure figures

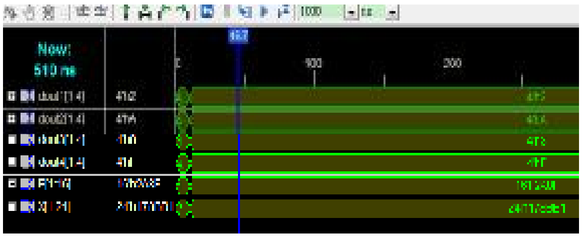

Fig. 4. The simulation waveform

Fig.4 shows that when $B = BBBB = 000101$ _111000_101110_110001, its hexadecimal data is 178B B1H. According to the selection method in Table Ⅱ, the output data can be computed as 0010_1010_1000_ 1111 (here only S$_1$-box's data is given), it is also expressed as 2A8FH.The results are correctly shown in Fig.4.The algorithm can be proved correct and effective which completes the conversion of 24 bits to 16 bits.

B. *Security Analysis*

This four S-boxes' relationship of input and output can be corresponded to four functions[14]. The definition domain of each function is $\{0,1,2,3,\cdots,63\}$, and the range is $\{0,1,2,3,\cdots,15\}$ Because four S-boxes are independent each other, so it can be calculated the joint probability distribution of four S-boxes. According to previous analysis [10], they have similar statistical properties to meet the normal distribution. As a result of the four S-boxes output range is $\{0,1,2,3,\cdots,60\}$, its mean and variance can be computed as $\mu = (0+1+2+\cdots+60)/61 = 30$ and $\sigma = 17.6068$ .Fig.6 gives the probability density curves of the four S-boxes output including actual design curve and theoretical curve, DES algorithm S-box probability density curves and the error probability density curve of the actual S-boxes output compared with the theoretical one. Fig.5 shows the values of four S-boxes output between 60-80 still can be get in the theoretical probability density curve, but according to the above analysis we can see the actual output value of the four S-boxes can not be more than 60. The error formula of our S-boxes output with the theoretical output is function (1) and the error formula of our S-boxes output in DES algorithm is function (2) as follows.

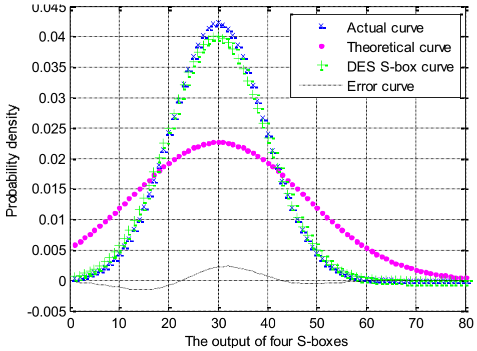

Fig.5. The comparison of four S-boxes output probability density curves

$$f_1(x) = \frac{1}{9.4306 \cdot \sqrt{2\pi}} e^{\frac{-(x-29.9668)^2}{2\cdot 89.0378}} - \frac{1}{9.9473 \cdot \sqrt{2\pi}} e^{\frac{-(x-29.6641)^2}{2\cdot 98.9513}} \qquad (1)$$

$$f_2(x) = \frac{1}{9.4306 \cdot \sqrt{2\pi}} e^{\frac{-(x-29.9668)^2}{2\cdot 89.0378}} - \frac{1}{17.6068 \cdot \sqrt{2\pi}} e^{\frac{-(x-30)^2}{2\cdot 310}} \qquad (2)$$

The average and the biggest error of our S-boxes output with the theoretical S-boxes output in probability density curve is 1.40×10^{-3} and 1.96 $\times10^{-2}$ .The average and the biggest error of our S-boxes output with the S-boxes output in DES algorithm is 1.88×10^{-3} and 2.3×10^{-3} . The error is basically control under 10^{-3} , so the statistical characteristics are similar with the statistical properties of S-boxes in DES algorithm, they have high security. We can also clearly see from Fig.5 that the vast output will focus on the 30 around, DES algorithm also has this feature, so, when choosing cipher text we should make the output of four S-boxes as far as possible from the 30 to prevent selective cipher text attack.

VI. CONCLCSION

In this paper a new secure communication model is proposed based on many security protocols, and it can resist some attacks. This communication model needs cryptographic techniques. So we research a lightweight encryption algorithm based on the improved DES algorithm to meet the security performance. S-boxes as the only non-linear devices in DES algorithm play a direct impact on the performance of the entire encryption system. In this paper, The S-box design is the core, through its optimization LEA algorithm is come true and the algorithm can be successful implemented on low-cost FPGA, statistical analysis shows that the algorithm features a certain degree of security.

VII. Acknowledgement

This research is sponsored by the natural science fund of China (60872003).

REFERENCES

[1] Ari Juels, "RFID Security and Privacy: A Research Survey" *IEEE Journal On Selected Areas In Communications*, VOL. 24, NO. 2, 2006, pp. 381-393.

[2] Su Mi Lee, Young Ju Hwang, Dong Hoon Lee, and Jong In Lim, "Efficient Authentication for Low-Cost RFID". *Computational Science and Its Applications – ICCSA*, Vol.3480, Springer Berlin, 2005, pp. 619-622.

[3] S.A. Weis, S.E. Sarma, R.L. Rivest, and D.W. Engels, "Security and privacy aspects of low-cost radio frequency identification systems", In Proc. of the First Security in Pervasive Computing, LNCS, Vol. 2802, 2003, pp. 201-212.

[4] Tang F, Wang L. "An Adaptive Active Control for the Modified Chua's Circuit ". Physics Letters A, 2005, 346(5 - 6) ,pp.342 - 346.

[5] Bowong S, Ka *kmeni* F M M, Koina R. "A New Synchronization Principle for a Class of Lur'e Systems with Applications in Secure Communication". International Journal of Bifurcation and Chaos, 2004, 14 (7), pp. 2477-2491.

[6] Weis A Stephen. "Security and privacy in radio frequency identification devices". Cambridge, MA: Master's Thesis of Massachusetts Institute of Technology, 2003.

[7] National Bureau of Standards. FIPS PUB 46, *The Data Encryption Standard*. U.S. Department of Commerce, Jan 1977.

[8] Biham E, Shamir A. *Differential cryptanalysis of DES -like cryptosystems*. Advances in Cryptology–CRYPTO'90 Proceedings.Spribger-Verlag, 19-91, 4(1).

[9] Wu WL, Ma HT, Qing S H. Differential and linear crypt -analysis of AC block cipher. Journal of Software, 2003, 14(3), pp.569-574.

[10] Matsui M. *Linear Cryptanalysis of DES cipher*.Eurocrypt'93, pp. 111-125.

[11] Lai K J. On the design and security of b1ock ciphers. In: ETH Series in Information Processing, Val. IKonstanz: Hartung-GorreVerlag, 1992.

[12] Xilinx, Inc., 2100 Logic Drive, San Jose, California. *Virtex 2.5V FPGA Series Data Sheet*, Oct 1999.

[13] Steve Kelem. Virtex Configuration Architecture Advanced User's Guide. Xilinx, Inc., 2100 Logic Drive, San Jose, California, Jun 1999. Application Note 151.

[14] Kullhack S. *Statistical Method in Cryptanalysis*. New York: Aegean Park Pess, 1976.