Nonlinear Transformation's Impact Factor of Cryptography at Confusion and Cluster Process

Lan Luo[1]
[1]*Networks and Intelligent Application of Block Cipher lab*
University of Electronic Science Technology of China,
Chengdu, China luolan@uestc.edu.cn

Tao Lu[1,4]
[4]Department of Civil and Structural Engineering,
School of Engineering, Aalto University, Aalto,
Espoo Finland tao.lu@aalto.fi

Zehui Qu[1,2]
[2]*Dipartimento di Informatica*
Università di Pisa, Pisa, Italy
zehui.qu@gmail.com

Qionghai Dai[1,5]
[5]*School of Information Science & Technology*
Automation Department, Tsinghua University, China
qhdai@tsinghua.edu.cn

Yalan Ye[1,3]
[3]*School of Computer Science and Technology*
University of Electronic Science Technology of China,
Chengdu, China yalanye@uestc.edu.cn

***Abstract*—For investigating the nonlinear character at the confusion and cluster process of cryptography, the nonlinear transformation's impact factor has been introduced. According to sorting result by different years, the amount of online cryptography is accounted. And then the nonlinear indexes, which are related to the outcome of amount reflecting the nonlinear character directly, are researched at the confusion process and the cluster process respectively. The nonlinear transformation's impact factor which includes both the private-key cryptography and public-key cryptography is studied with known nonlinear index at naïve Bayesian model, which combines with the networks environment fused in protocol whenever confusion or cluster process. To any networks environment, higher the nonlinear transformation's impact factor is, more popular the used cryptography is because more amounts of kinds of level users requiring stronger secure cryptography. So, the impact factor of the nonlinear transformation is a kind of cryptography's label indicating the suitable to application environment by suitable crowd. Contrarily, the extent of secure can measure up the nonlinear character of cryptography precisely.**

Keywords-nonlinear transformation; confusion and cluster; impact factor; Bayesian model

I. INTRODUCTION

Nonlinear transformation (NT) [1][2][3][4][5], which exists at any kinds of secure communication systems, is an important part of cryptographic study. The cryptography includes private key and public key cryptography. The private key cryptography is a secure process which uses one secret key to confuse and recover the information, and the private key process system is depend on a pseudorandom number generator, such as HASH [6][7][8], the block cipher[9][10][11][12], the stream cipher[13][14][15]. The public key cryptography is a process which is used two keys to ensure the information, and it is based on a mathematic problem, such as integer factorization and Elliptic curve [16][17][18][19]. Whether private or public key, the NT is evolved [20][21][22] from value tables supported by development of the information technology[23]. Since the Future Internet [24][25][26][27][28], wireless sensor[29][30], RFID[31][32] and quantum communication [33][34][35] are developed, the application of block cipher, stream cipher and HASH[36] is prevalence to ensure the communication's secure. The RSA [37], ECC[38] are used at mini-information secure condition, such as digital signature, key exchange schemes. The Lattice-based cryptosystem is simply introduced at the paper. The NT Impact factor (IF) of cryptography is a simple and popular way to value the NT influence at Bayesian model [39][40][41].

This paper focuses on the nonlinear transformation impact factor of cryptography, which is clustering with the different time, different scales of nonlinear transformation and different application environment. The rest of the paper is organized as follows. Section 2 contains the two inversed study directions which are confusion process and cluster process at cryptography. The nonlinear indexes and NT IF according to Bayesian model are described in section 3. Section 4 includes conclusion about the effect of NT IF.

II. CONFUSION AND CLUSTER PROCESS AT CRYPTO

The confusion and cluster process are the certain parts of encipher and decipher of information at cryptography. The whole cryptographic system is a nonlinear process because of the sharing of NT or nonlinear mathematic problem. The confusion process is the nonlinear part besides the key

 ISBN: 978-1-61208-226-4

addition and linear diffusion process at secret key condition, and it is the whole operation process at the public key condition. Meanwhile the cluster process is the first step which is a simple differential sort process at whenever the secret key cryptographic system's condition of white box, gray box or black box, and is the exponent of the public key cryptographic. It shows the sort of online cryptography according to the change of time at figure 1. The published online cryptography amount has a jump as figure 2.

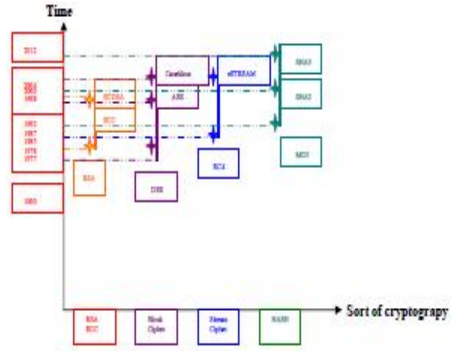

Figure 1. The sort of cryptography online according to time

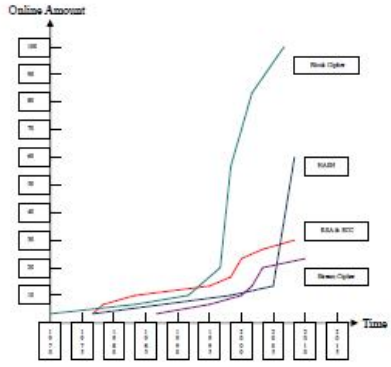

Figure 2. Online Cryptography Amount by Years

A. *Confusion and Cluster Process Precisely Depicted*

The confusion processes and cluster processes of cryptography sort out to 5 categories, which are confusion of block cipher, of stream cipher, of HASH, of RSA and of ECC in this paper. The confusion process of block cipher has a NT, which is from N bits to M bits. At the beginning of the published block cipher, the confusion process is constructed by 6 bites to 4 bits NT, which is activated 4 to 8 times and repeated 16 rounds, such as DES in 1976. IDEA is also an old block cipher which is 16 bits to 16 bites NT repeated 8.5 rounds in 1991. Updated block cipher is constructed by 8 bits to 8 bits NT activated 8 to 16 times and repeated 10, 12 or 14 rounds, such as AES in 1998. Camellia is a new block cipher which is the same nonlinear part similar as AES repeated 18 rounds in 2000. The stream cipher has none any nonlinear part at first, such as A5 used in the GSM cellular telephone standard. From November 2004 to April 2008, the project founded by EU ECRYPT network founded, eSTREAM, identify new stream ciphers suitable for widespread adoption. Salsa20, which is based on software, is one of the winners. The NT of Salsa20 is 128 bits to 128 bits repeated 20 rounds. Trivium based on hardware is one of the winners. Its NT is 228 bits to 228 bits at 4th rounds. Those are some nonlinear emblematical transformations.

The old HASH SHA1 is published at 1995, which is 96 bits to 96 bits NT repeated 80 rounds. SHA2 has a NT which is 96 bits to 96 bits activated 2 or 4 times repeated 64 or 80 rounds. A new SHA3 competition is an open competition from October 31, 2008 to the end of 2012, and the final round candidates have occurred on December 10, 2010. Keccak of finalists has a NT, which is 192 bits of 64 bit memory to 64 bits repeated 24 rounds. Skein of finalist has a NT, the designers call it MIX, which is 128 bits to 128 bits repeated 72 rounds or 256 bits to 256 bits repeated 80 rounds.

RSA and ECC are another typical type of cryptography comparing to block cipher, stream cipher and HASH, being called the public key cryptography. While the RSA patent expired in 2000, the NT of RSA is the naïve bit amount of prime number at ANSI X9.31 in 1998, which is 1024 bits. Meanwhile, the NT of ECC is the 192, 224, 256, 384, and 521 bits according to FIPS 186-3 in 2009. The Lattice based cryptography is a new mathematic problem appears with the post-quantum cryptography. The "ideal lattice" designed by Craig Gentry, which is announced by IBM at 6.2009, NL is depicted by the Lattice problem in n-dimensional Euclidean space R^n with a strong periodicity property.

B. *Nonlinear Transformations at a Confusion process*

The confusion process is to confuse the information to pseudo-random data which cannot be reversed at the useful-life time of the information. The different cryptography uses different nonlinear transformation to obtain the random. The block ciphers, being the popularity from public game AES, have active or non-active model at NT. If the box or the value table implement at NT part, the active number of box affects the nonlinear complexity and speed directly. More active boxes, more complexity the NT is. The mathematic problem NT depends on the exponent of the nonlinear function. There are some distinctions, which are the speed and secure influences, between all boxes activated and part of boxes activated at confusion process. There are some useful nonlinear factors at sub-key rolling in process. The information block and nonlinear complexity of sub-key are covered up by NT because they are far lower exponent than NT's. And then, the effect of NT is diffused with a linear array. By the way, a block cipher is constituted by iterative rounds which include information block, sub-key addition, NT and diffusion. The stream ciphers have a high speed character, which is usually suitable to the secure of wireless communication and a high-capacity bandwidth environment. The stream ciphers usually constitute by linear feedback shift register (LFSR) or nonlinear feedback shift register (NFSR) fusing by NT. If the NFSR is the part of a stream cipher, the resilience function, which has both better balance and better anti-differential analysis, is a better choice than bent function. The NT part of stream cipher can be just similar to a whole

 ISBN: 978-1-61208-226-4

rounds or mini rounds of block cipher. And more and more simple NT fused directly to LFSR or NFSR is appearing, such as Trivium, Grain, or NUSH. Information is confused by a simple NT of stream cipher has lower secure but higher speed. HASH is a stream cipher or a block cipher adding kinds of digest function. The digest function is considered as a linear part so that the NT of HASH process refers to the stream or block cipher. So the NT of HASH is the same as the stream or block cipher.

The confusion process of public key cryptography is to ensure the mini-scale communication key's security. The mathematic confusion process is nonlinear because exponent is more than 1. The NT of RSA is up to 210 in 2010. Then the successor ECC has a popular confusion process at the equation $y^2 = x^3+ax+b$. So the NT of ECC is considered as 3^2. Until now, the idea model of the Lattice based confusion process is a no known bounds iteration of vectors at least 2 equivalence dimensions. The NT of lattice based confusion process is $(\text{number at least } 2)^n$, the n is unknown.

C. Nonlinear Transformations at a Recovery Process

A recovery process is to unveil the cipher step by step at the condition of White-box, grey-box and black-box. If the confusion process is totally published, the recovery process is at a white-box situation. If there only are confusion date or some parts of confusion process, the recovery process is at a black-box or a grey-box situation. The NT at a recovery process based on a white-box has to operate carefully rounds by rounds, or just operates the reversed process of the mathematic problem. At this situation, the NT is strictly equal to the reverse of NT at confusion process. According the grey-box and black-box conditions, the cluster process is necessary to congregate the confusion date. By the NT of cluster process, the data is sorted by some factors, such as the linear index, the nonlinear index, characters of Pseudo-Randomness or avalanche, including by other mathematic ways and means. Then those sorts are the results of first glimpse of the recovery process, in which the grey-box is covered to the white-box and the black-box to grey-box. The black-box is equivalent to several grey-boxes which also are considered as the scale of white-box based on the results. When the cluster data has a certain sameness with a certain white-box, the grey-box can be equal to the white-box at the scale. In conclusion, the black boxes mix with white boxes according to results of NF cluster at recovery process can be converted to grey boxes, which are prepared to be white.

So, the cluster process of NT can recovery white-box completely and converse the deep color box to undertone box in basis of confusion date clustering.

III. NONLINEAR INDEXES BASED ON BAYESIAN MODEL

Bayesian inference is a rational engine for solving such problems within a probabilistic framework, and consequently is the heart of most probabilistic models of nonlinear indexes the cryptography. The nonlinear indexes, which include the private-key cryptography and public-key cryptography are sorted at naïve Bayesian model, whenever it's confusion or cluster process. The simple NT index reflect the nonlinear character itself meanwhile the NT IF combine the networks or cryptography protocols into NT index.

A. Nonlinear transformations index at cryptography

The nonlinear transformations index at cryptography is a kind of rough description about the strength of cryptography. The NT index of block cipher is the product of the active boxes number of a block and exponent of nonlinear function. For example, the NT index of standard AES which has 16 actives s-boxes at each of 10 rounds is that product of 8 and 160. The NT index of stream cipher is the sum of each part of nonlinear functions' exponent. The stream cipher Grain has a NT index which is the product of 3 and 6. The NT index of HASH evolved from the certain cipher, which is block cipher or stream cipher, is the same count method as the certain cipher. The NT index of public-key cryptography, such as RSA or ECC, is that 1024 or 9 referring to the standard. The ECC has the homothetic nonlinear character as private-key cryptography.

If only consider the NT index, the block cipher is equal to many times of stream cipher, such as NT index of AES is about 72 times of Grain. Meanwhile, RSA is almost equal to 120 times of ECC. Thus, standard AES can be secure the 72 times bandwidth data the same as Grain. The signature with RSA is equal to 72 signatures of ECC. The figure3 shows the NT IF by years at different networks environment.

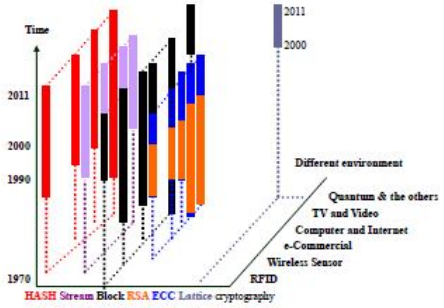

Figure 3. NT IF at different environments according to time

B. The Nonlinear Transformations Impact Factor Based on Naïve Bayesian Model

The NT impact factor is related to NT index according to the application of cryptography at the naïve Bayesian model. The NT index reflects the nonlinear character directly and NT IF combines with the networks environment fused in protocol. So T_0 is the NT index which is the cryptography's nonlinear measure and the T_1 is that is the different networks. Naive Bayesian networks are identified with theories at the lowest, most concrete level of the abstraction hierarchy, level T_0. The categorization grammars are typically identified with the T_1-level theories that define hypothesis spaces of T_0-level structures and assign prior probabilities to those hypotheses, thereby guiding inferences about the naive network structure T_0 mostly likely to have given rise to some observed dataset d. A Bayesian learner evaluates a naive network hypothesis T_0 based on its posterior probability:

 ISBN: 978-1-61208-226-4

$$P(T_0|d,T_1) = (P(d|T_0)P(T_0|T_1)) / P(d|T_1) \quad (1)$$

where the denominator is

$$P(d|T_1) = \sum P(d|T_0)P(T_0|T_1) \quad (2)$$

The naive grammar T_1 specifies a probabilistic process for generating naive-network hypotheses. With such a Bayesian model, the NT IF of cryptography is discussed in a causal way which is considered as naive condition. NT IF is the iteration of Bayesian model result and the amount of the cryptography. There is one kind of NT based on ciphers' application among different networks which occurs between private-key cryptography and public-key cryptography. So the NT IF is identical as the cryptography amount online. The table1 is the NT IF result considered the impact by the frequency of same networks environment.

TABLE I. NT IF OF CRYPTOGRAPHY AT DIFFERENT ENVIRONMENT IN 2011 ACCORDING TO BAYESIAN MODEL

NT IF	HASH	Stream	Block	RSA	ECC	Lattice
RFID	X	X	1.172	X	X	X
Wireless Sensor	0.844	0.348	1.192	0.439	0.738	X
e-Commercial	0.839	0.333	1.187	0.444	0.743	X
Com & Inter	0.849	0.338	1.182	0.449	0.748	X
TV & Video	0.834	0.343	1.177	0.434	0.733	X
Qum & other	X	X	1.167	X	X	0.167

IV. CONCLUSION

The nonlinear character of both private-key cryptography and public-key cryptography is expressed by the nonlinear transformation's impact factor simply. At confusion process, the NT IF is a complexity of nonlinear character. At cluster process, the NT IF is the appearance color of nonlinear box or the known degree of a cryptography algorithm. At any networks, higher the NT IF is, the more popularity the cryptography is used. The reason is that more amounts of users require stronger secure level. Furthermore, the NT's IF of cryptography at confusion and cluster process based on Bayesian model, which is an advanced description of the nonlinear character according to the different application environments, demonstrates that higher NT IF is a suitable wider width-band data communication. The NT IF can be a label of a cryptography system indicating the suitable crowd and suitable application environment. Contrarily, the extent of cryptography's secure can be measure up the NT impact factor's precision deeply.

ACKNOWLEDGMENT

This research is partly financially supported by the zygx2010J068. We would like to deeply thank the various people who provided us with very useful and helpful suggestions. The authors sincerely thanks for YuQi He, the academician of Academy of Sciences of the United States and Department of Automation Tsinghua University and Harvard University, who is our tutor of the Bayesian model at Sciencenet.CN.

REFERENCES

[1] Svetla Nikova, Vincent Rijmen and Martin Schläffer, Nonlinear transformations S-box Noekeon the other design of hardware Secure Hardware Implementation of Nonlinear Transformations in the Presence of Glitches, Journal of Cryptology, Vol. 24, Number 2, April 2011, Pages 292-321

[2] D. R. Stinson and J. L. Massey, An infinite class of counterexamples to a conjecture concerning nonlinear resilient transformations. Journal of Cryptology, 1995, Vol. 8, Number 3, Pages 167-173

[3] Stanislav V. Smyshlyaev, Perfectly Balanced Boolean Transformations and Golić Conjecture, Journal of Cryptology, 3 July 2012, Pages 464-483

[4] Carlisle Adams and Stafford Tavares, The structured design of cryptographically good s-boxes, Journal of Cryptology, 1990, Volume 3, Number 1, Pages 27-41

[5] Nenad Dedić, Gene Itkis, Leonid Reyzin and Scott Russell, Upper and Lower Bounds on Black-Box Steganography, Journal of Cryptology, 2009, Vol. 22, Num 3, Pages 365-394

[6] Lars R. Knudsen, Xuejia Lai and Bart Preneel, Attacks on Fast Double Block Length Hash Transformations, Journal of Cryptology, 1998, Volume 11, Number 1, Pages 59-72

[7] ebastiaan Indesteege and Bart Preneel, Practical Collisions for EnRUPT. Journal of Cryptology, 2011, Volume 24, Number 1, Pages 1-23

[8] David Cash, Dennis Hofheinz, Eike Kiltz and Chris Peikert, Bonsai Trees, or How to Delegate a Lattice Basis. Eurocrypt'10 Proceedings of the 29th Annual international conference on Theory and Applications of Cryptographic Techniques Pages 523-552

[9] Debra L. Cook, Moti Yung and Angelos D. Keromytis, Elastic block ciphers: method, security and instantiations, International Journal of Information Security, 2009, Volume 8, Number 3, Pages 211-231.

[10] Joan Daemen and Vincent Rijmen AES proposal[R]: Rijndeal, NIST, FIPS PUB 197, 11. 2001

[11] Matsui, M., Nakajima, J., and S. Moriai, A Description of the Camellia Confuseion Algorithm, RFC 3713, April 2004.

[12] Serge Vaudenay, Decorrelation: A Theory for Block Cipher Security, Journal of Cryptology, 2003, Volume 16, Number 4, Pages 249-286

[13] T.W. Cusick, C. Ding, A. Renvall, Stream Ciphers and Number Theory, Published: APR-1998, ISBN 10: 0-444-82873-7, Imprint: NORTH-HOLLAND

[14] Matthew Robshaw and Olivier Billet, New Stream Cipher Designs The eSTREAM Finalists, Lecture Notes in Computer Science, Volume 4986, 2008

[15] Daniel J. Bernstein, The Salsa20 Family of Stream Ciphers, Lecture Notes in Computer Science, 2008, Volume 4986, Pages 84-97

[16] Steven D. Galbraith, Xibin Lin and Michael Scott, Endomorphisms for Faster Elliptic Curve Cryptography on a Large Class of Curves, Journal of Cryptology, Volume 24, Number 3 / July 2011, Pages 446-469

[17] Johan Håstad and Mats Näslund, Practical Construction and Analysis of Pseudo-Randomness Primitives, Journal of Cryptology, 2008, Volume 21, Number 1, Pages 1-26

[18] S. Micali and C. P. Schnorr, Efficient, perfect polynomial random number generators, Journal of Cryptology, 1990, Volume 3, Number 3, Pages 157-172

 ISBN: 978-1-61208-226-4

[19] Omer Barkol, Yuval Ishai and Enav Weinreb, On d-Multiplicative Secret Sharing, Journal of Cryptology, 2010, Volume 23, Number 4, Pages 580-593

[20] Eran Tromer, Dag Arne Osvik and Adi Shamir, Efficient Cache Attacks on AES, and Countermeasures, Journal of Cryptology, 2010, Volume 23, Number 1, Pages 37-71

[21] M. Bellare, A. Boldyreva, L. Knudsen and C. Namprempre, On-line Ciphers and the Hash-CBC Constructions, CRYPTO '01 Proceedings of the 21st Annual International Cryptology Conference on Advances in Cryptology, Pages 292 - 309

[22] David M. Goldschlag, Stuart G. Stubblebine and Paul F. Syverson, Temporarily hidden bit commitment and lottery applications, International Journal of Information Security, 2010, Volume 9, Number 1, Pages 33-50

[23] Dibyendu Chakrabarti, Subhamoy Maitra and Bimal Roy, A key pre-distribution scheme for wireless sensor networks: merging blocks in combinatorial design, International Journal of Information Security, 2006, Volume 5, Pages 105-114

[24] Renbin Xiao, Tinggui Chen and Chunhua Ju, Research on Product Development Iterations Based on Feedback Control Theory in a Dynamic Environment, International Journal of Innovative Computing, Information and Control, Volume 7, Number 5(B), May 2011, Pages. 2669-2688

[25] Serap Ataya, Marcelo Maserab, Challenges for the security analysis of Next Generation Networks, Information Security Technical Report, Vol. 16, Issue 1, February 2011, Pages 3-11

[26] William Walker, Mobile telephony security compromises, Information Security Technical Report, Volume 15, Issue 3, August 2010, Pages 134-136

[27] Allan Tomlinson Corresponding Author Contact Information, Po-Wah Yau, John A. MacDonald, Privacy threats in a mobile enterprise social network, Information Security Technical Report, Volume 15, Issue 2, May 2010, Pages 57-66

[28] Roman, R.E, Alcaraz, C., Lopez, J. The role of Wireless Sensor Networks in the area of Critical Information Infrastructure Protection, Information Security Technical Report, Volume 12, Issue 1, 2007, Pages 24-31

[29] Svendsen, N.K.a , Wolthusen, S.D. Connectivity models of interdependency in mixed-type critical infrastructure networks, Information Security Technical Report, Volume 12, Issue 1, 2007, Pages 44-55

[30] Igure, V.M., Laughter, S.A., Williams, R.D. Security issues in SCADA networks, Computers and Security, Volume 25, Issue 7, October 2006, Pages 498-506

[31] Arne Tauber, A survey of certified mail systems provided on the Internet, Computers & Security, Volume 30, Issues 6-7, September-October 2011, Pages 464-485

[32] Roberts, C.M, Radio frequency identification (RFID), Computers and Security, Volume 25, Issue 1, February 2006, Pages 18-26

[33] Alhazmi, O.H., Malaiya, Y.K., Ray, I. Measuring, analyzing and predicting security vulnerabilities in software systems, Computers and Security, Volume 26, Issue 3, May 2007, Pages 219-228

[34] Miron Abramovici, A solution for on-line trust validation, Anaheim, CA, USA, June 09-June 09, ISBN: 978-1-4244-2401-6, 2008 IEEE International Workshop on Hardware-Oriented Security and Trust

[35] Jens-Peter Kaps, Gunnar Gaubatz, Berk Sunar, Cryptography on a Speck of Dust, Computer, vol. 40, no. 2, Feb. 2007, doi:10.1109/MC.2007.52, Pages 38-44

[36] Philip O'Kane, Sakir Sezer, Kieran McLaughlin, "Obfuscation: The Hidden Malware," IEEE Security and Privacy, vol. 9, no. 5, Sep./Oct. 2011, Pages 41-47

[37] Romain Giot ,Mohamad El-Abed, Baptiste Hemery, Christophe Rosenberger, Unconstrained keystroke dynamics authentication with shared secret, Computers & Security, Vol. 30, Issues 6-7, September-October 2011, Pages 427-445

[38] Gary S.-W. Yeo and Raphael C.-W. Phan, On the security of the WinRAR confuseion feature, International Journal of Information Security, 2006, Volume 5, Pages 115-123C

[39] Yuanyuan Wang, Yunming Ye, Xutao Li, Michael K. Ng and Joshua Huang, Hierarchical Information-Theoretic Co-Clustering for High Dimensional Data, International Journal of Innovative Computing, Information and Control, Volume 7, Number 1, January 2011, ISSN 1349-418X, Pages 487-500

[40] Vladimir S. Udaltsov, Jean-Pierre Goedgebuer ,Laurent Larger, Jean-Baptiste Cuenot, Pascal Levy, William T. Rhodes , Cracking chaos-based confuseion systems ruled by nonlinear time delay differential equations, Physics Letters A 308 Pages 54–60

[41] Massimiliano Zanin1, Alexander N. Pisarchik, Boolean Networks for Cryptography and Secure Communication, Nonlinear Sci. Lett. B, Vol. 1, No.1, 2011, Pages 25-32

 ISBN: 978-1-61208-226-4