WSEAS TRANSACTIONS on COMMUNICATIONS

Issue 2, Volume 3, April 2004

ISSN 1109-2742 http://www.wseas.org

Stream Ciphers created by a Discrete Dynamic System for application in the Internet.

V.SOULIOTI [1], Y.BAKOPOULOS [1], S.KOUREMENOS [1,2], Y.VRETTAROS [1], S.NIKOLOPOULOS [2], A.DRIGAS [1]

1. National Center for Scientific Research "DEMOKRITOS"
Department of Technological Applications
P.O. Box 15310 Gr. Ag. Paraskevi Attikis, GREECE

2. National Technical University of Athens
School of Electrical and Computer Engineering
P.O. Box 15780 Gr. Zographou, Attikis, GREECE

jvr@imm.demokritos.gr http://imm.demokritos.gr

Abstract: A discrete dynamic system is utilized for the creation of random number series. It contains discontinuities based on the modulo and signum functions. The binary number series created show almost total randomness, as indicated by block entropy tests. The concept of a virtual cryptographic device is defined and analyzed. A new method, based on the above is proposed, for secure and easy application in the Internet and all digital networks in general.

Key-words: symbolic dynamics, stream ciphers, pseudorandom, incompressible, encryption, entropy, signum, modulo, security, internet.

1 Introduction

The Symbolic Dynamics of both continuous and discontinuous Discrete Dynamic Systems have been studied extensively [5], [6], [11], [18]. One of the most important applications is the generation of pseudorandom number series for use in encryption of messages in large area networks as the Internet. The applications make use of the chaotic pseudorandom behavior such systems may exhibit in their phase space trajectories [17], [2], [3], [4], [19].

The most common method is to use the symbolic series of the systems' evolution in time. The symbolic dynamics must fulfill certain demands so as to be suitable for the specific application. The created series must appear to be random to a third party and to be almost totally incompressible. It must be reproducible, in the sense that the same initial conditions must always reproduce exactly the same series every time. It must also be easy and fast to create, starting from a relatively small set of real valued parameters. Finally, the set of all different series that can be created by this method should be as large as possible, so that frontal attacks by brute force would be useless [16].

The dynamic systems used are mostly chosen because they show chaotic behavior. The most well known are the standard map, the logistic map, the tent map and some others based on discontinuous functions like the step function or the modulo function [17]. Discontinuous dynamic systems of higher dimension exist in abundance, such as the Sigma – Delta Modulation systems mentioned by many authors [5], [7], [8], [9], [13], [14], [15], [2], [3], [4], [19].

The authors of this work believe that such systems, suitably modified, can be applied with considerable success to random number generation and stream cipher creation.

The system examined here belongs to this class. It is a two dimensional variant of the above mentioned systems.

In its simplest form, with zero input and a signum discontinuity, it is described by Eqns. (I)

(I): $\chi_{n+1} = A \cdot \chi_n + B \cdot S(x_n) + U_n$

or:

$$\begin{pmatrix} x_1(n+1) \\ . \\ . \\ x_k(n+1) \end{pmatrix} = \begin{pmatrix} a_{11} & . & . & a_{1k} \\ . & . & . & . \\ . & . & . & . \\ a_{k1} & . & . & a_{kk} \end{pmatrix} \cdot \begin{pmatrix} x_1(n) \\ . \\ . \\ x_k(n) \end{pmatrix} +$$

$$+ \begin{pmatrix} b_{11} & . & . & b_{1k} \\ . & . & . & . \\ . & . & . & . \\ b_{k1} & . & . & b_{kk} \end{pmatrix} \cdot \begin{pmatrix} \mathrm{sgn}(x_1(n)) \\ . \\ . \\ \mathrm{sgn}(x_k(n)) \end{pmatrix} + \begin{pmatrix} w_1 \\ . \\ . \\ w_n \end{pmatrix}$$

where:

$$A = \begin{pmatrix} a_{11} & . & . & a_{1k} \\ . & . & . & . \\ . & . & . & . \\ a_{k1} & . & . & a_{kk} \end{pmatrix}, B = \begin{pmatrix} b_{11} & . & . & b_{1k} \\ . & . & . & . \\ . & . & . & . \\ b_{k1} & . & . & b_{kk} \end{pmatrix},$$

$$\chi_n = \begin{pmatrix} x_1(n) \\ . \\ . \\ x_k(n) \end{pmatrix}, \chi_{n+1} = \begin{pmatrix} x_1(n+1) \\ . \\ . \\ x_k(n+1) \end{pmatrix},$$

$$S(\chi_n) = \begin{pmatrix} \mathrm{sgn}(x_1(n)) \\ . \\ . \\ \mathrm{sgn}(x_k(n)) \end{pmatrix}, U_n = \begin{pmatrix} w_1 \\ . \\ . \\ w_k \end{pmatrix},$$

Where A, is a rotation matrix in k dimensions, while B may be the identity matrix or any matrix with $|\det(B)| \leq 1$

The signum function is defined in this work as sign(x) = -1 if x < 0 and sign(x) = 1 otherwise.

A more complicated system is created by the introduction of appropriate input functions. One form of input consists of a perturbation of the rotation matrix in Eqn (I) [19], [1], [2], [3], [4]. To each term a_{ij} of the matrix a perturbation ε_{ij} is added. For example, if in two dimensions the rotation matrix has terms: $a_{11} = \cos(f) = a_{12}$, $a_{21} = \sin(f) = -a_{12}$, then a perturbation parameter can be added ε to a_{11} and a_{22}, so that $a_{11} = a_{22} = \cos(f) + \varepsilon$, leaving the other terms unchanged.

$$\begin{pmatrix} x_1(n+1) \\ x_2(n+1) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix} +$$

(II):

$$+ \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} \mathrm{sgn}(x_1(n)) \\ \mathrm{sgn}(x_2(n)) \end{pmatrix} + \begin{pmatrix} w_1(n) \\ w_2(n) \end{pmatrix}$$

A further step is to make use of the modulo function. As the authors of the present study have defined it, the function MOD[x;p] is equal to the value of the real variable x minus the product of p by the integral part of the quotient of the absolute value of x divided by p and by the signum of x: MOD(x;p) = x – sign(x) (p) INT(|x|/p). Here p is defined to be positive.

$$\begin{pmatrix} x_1(n+1) \\ x_2(n+1) \end{pmatrix} = MOD\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} x_1(n) \\ x_2(n) \end{pmatrix} +$$

(III):

$$+ \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} \mathrm{sgn}(x_1(n)) \\ \mathrm{sgn}(x_2(n)) \end{pmatrix} + \begin{pmatrix} w_1(n) \\ w_2(n) \end{pmatrix}; p \}$$

or, by coordinates:

(IIIa): $x_i(n+1) = \mathrm{MOD}\{\alpha_{i1}\chi_1(n) + \alpha_{i2}\chi_2(n) + b_{i1}\mathrm{sgn}(\chi_1(n)) + b_{i2}\mathrm{sgn}(\chi_2(n)) + w_i(n); p\}$

i = 1,2 (with obvious generalization in higher dimensions).

This is the form to be studied in this manuscript. It is one of the best fitted for random number generation with application to stream ciphers and Cryptographic Key Creation and Distribution.

There are several interesting ways to define a symbolic dynamics based on the class of systems described by Equation (IIIa). The best studied by the authors of this work is the most obvious. It is based on the properties of the signum function, as defined above.

Let the problem be restricted to its two – dimensional form. This may be done without any real loss of generality, since all the relevant concepts can be readily generalized to any dimensions.

The symbols used for the description of the system are four in number and are defined as follows: If $\mathrm{sign}(x_1(n)) = 1$ and $\mathrm{sgn}(x_2(n)) = 1$, then the value of the symbol is defined as s(n) = 0. If $\mathrm{sgn}(x_1(n)) = -1$ and $\mathrm{sgn}(x_2(n)) = 1$, then s(n) = 1. If $\mathrm{sgn}(x_1(n)) = -1$ and $\mathrm{sgn}(x_2(n)) = -1$, then s(n) = 2.

Finally, if $sgn(x_1(n)) = 1$ and $sgn(x_2(n)) = -1$, then $s(n) = 3$.

So, for every vector x(n) in configuration space, with coordinates $x_1(n)$ and $x_2(n)$, will correspond a symbol s(n) taking values from the set: {0, 1, 2, 3}. The symbolic series created depends on the following parameters: The initial values of the coordinates, $x_1(0)$ and $x_2(n)$. The rotation angle f. The perturbation parameter ε. The modulo parameter p. And finally the number of iterations n, defining the length of the symbol string.

2 Study of the Symbolic Dynamics

The dynamic system described by Eqn. (I), in its two – dimensional version, contains only one source of nonlinearity, the signum function defined in the introduction. The discontinuity of this function may lead the system into chaotic behavior. As a result the trajectories of the system in configuration space will be parts of a fractal set and the symbolic series will be aperiodic [13], [14], [19], [1], [2]. In such a case, the symbolic series will appear random to a third party. Although in theory there are initial conditions that would lead to fractal trajectories for every angle f in the interval: $f \in (0, \pi/2)$, in fact there have been observed fractal trajectories only in occasional values of f, specifically for $f = \pi/6$ (30^0) and $f = 4\pi/9$ (80^0) [Fig 1 a,b]. The rest of the orbits are periodic, with the symbolic series following strict rules. There are no stable points and the trajectories are stable, with well defined basins of attraction.

From the point of view of random number generation, if block entropy methods are applied [12], it can be demonstrated that most aperiodic series obtained this way contains significant structure and is compressible to the degree that it is not suitable for cryptographic key applications. It is possible that the same holds for systems of higher dimension, although this is still an open question, pending investigation. Various other methods of evaluation also indicate clearly that the systems of equation (I) are not suited for cryptographic key creation.

Equation (II) leads to an entirely different situation, as far as the behavior of the trajectories in phase space and the aperiodic symbolic series are concerned. As it will be shown elsewhere, there are infinite isolated unstable point orbits, while all other orbits are aperiodic (Fig 2). This is due to the fact that the introduced perturbation ε destroys the periodicity of the orbits by making the determinant of the linear matrix larger than one. This leads to a richness of trajectories in phase space much more complex than that of the previous case (Eqn.I), but still, from the point of view of random number generation, there is not enough randomness and incompressibility for stream cipher creation and distribution of cryptographic keys [12].

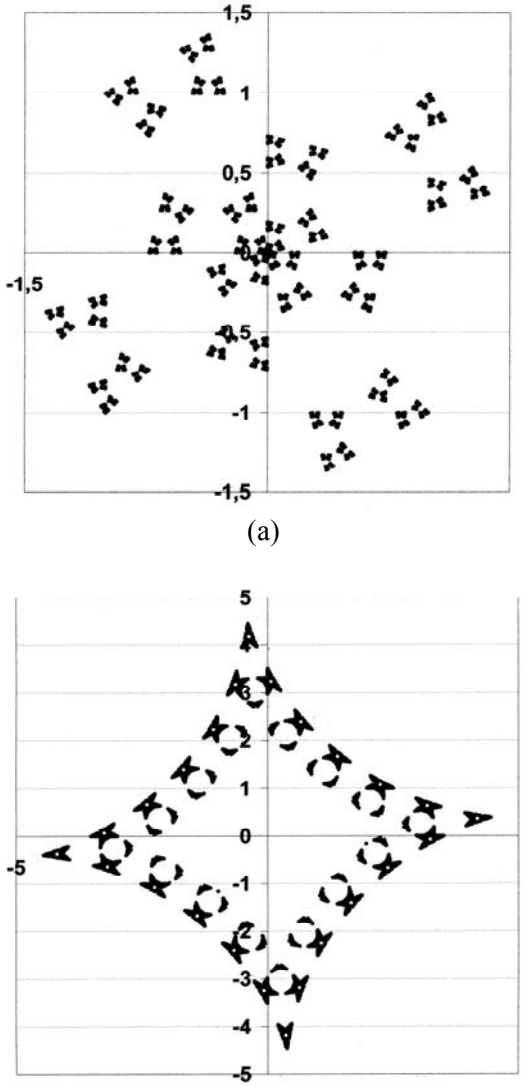

Fig.1 (a): f=30^o
(b): f=80^o

The creation of symbolic series which pass successfully the tests of randomness and incompressibility is achieved by the utilization of the systems derived from Equation (III). This family of systems contains two discontinuous functions, the signum function, included in all definitions so far, and the modulo function which is only used in Eqn (III). These two factors of

complexity and apparent pseudorandom behavior, lead to chaotic symbolic series and possibly, to some degree, may be used for cryptographic key creation and distribution [17].Yet it is only in combination with the perturbation factor ε that the method reaches its full potential. The virtual encryption machines defined by the authors [1], [2], [3], [4], [19], and described by Eqn (III) create cryptographic keys that pass, not only the standard tests known so far in the relevant literature [16], but also the new tests based on the use of the block entropy concept [12].

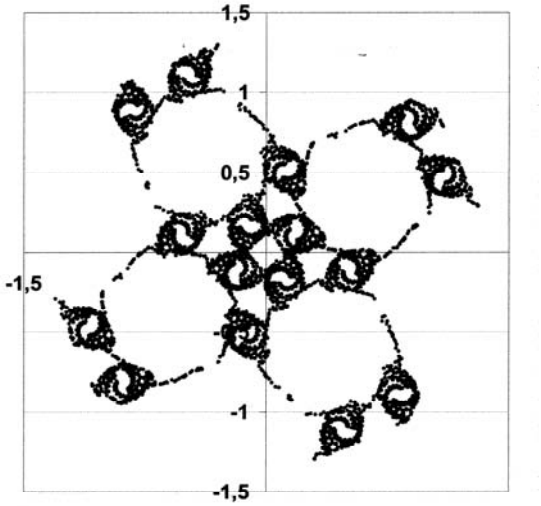

Fig. 2 f=24°
Perturbation parameter ε=0.01

3 Phase space pictures and symbolic series of Eqn (III) in two dimensions.

The two – dimensional system (Eqn (IIIa)) contains at least five parameters which control its behavior: The angle of rotation f, the perturbation parameter ε, the parameter p of the modulo function and the values of the initial conditions $x_1(0)$ and $x_2(0)$. (A sixth parameter, the number of iterations which defines the length N of the symbolic series, does not immediately influence the randomness of the series, once N is large enough to yield satisfactory statistics. Usually the length of the binary stream cipher should be of the order of 1,000000 digits).

As seen in Fig.3, the phase space profile does not present any structure. This is an indication of the total apparent randomness which is the first attractive feature of the system as a random number generator. Another factor is the absolute repeatability of the cipher created by the system. This is guaranteed by the equations of the system. The last property that a random number generator must have is the very large set of ciphers it can produce. Theoretically, the set may be infinite, but in practice its size depends on the properties of the random number generator in use.

The two – dimensional system described by Eqn (IIIa) is capable of creating a very large set of different stream ciphers. This is due to the extreme sensitivity of the trajectories in phase space to the initial conditions. The dependence of the system to the initial conditions is partly controlled by the accuracy of the calculations mainly with regard to the signum discontinuity. The important point is how to define the proposition '$x_i(n) = 0$', where $x_i(n)$ is a coordinate of a trajectory point at time n (i = 1,2). In the discrete mathematics of computer calculations there is no such thing as '0'. So it is postulated that '$x_i(n) = 0$' if $|x_i(n)| \leq \zeta$, where ζ is defined as a very small positive number, for example $\zeta = 10^{-20}$. This parameter controls the behavior of the system and the form of the phase space trajectories. Therefore it is critical for the structure of the symbolic series and its sensitivity to initial conditions. A variation of the order 10^{-20} near an axis may change the properties of the series from that point on [1],[2],[3],[4], [19]. So the uncertainty of the initial conditions is controlled by the architecture of the computer system and the platform on which the application will be materialized. There is no theoretical limit to the sensitivity of the dynamic system.

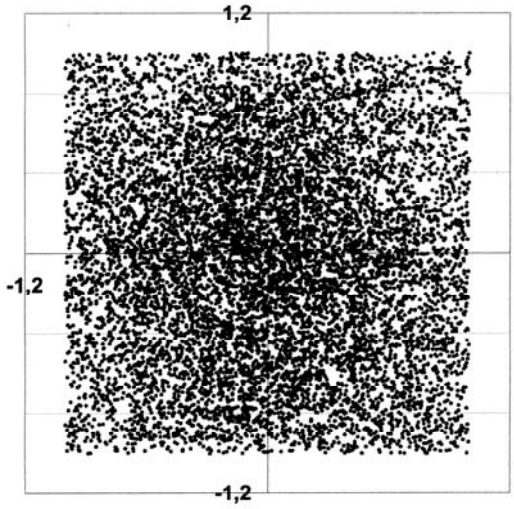

Fig.3 mod1
f=24
ε=0,8

So it may be concluded that there are two possible ways to increase the set of keys that may be created: one, to increase the dimension of the

system and to add input functions and parameters. The other, to increase the precision of the calculations and make use of a ζ as small as possible. Even though there are many open questions in both methods, it is obvious that the versatility of this system is a definite advantage over other random number generators.

4 The communication protocol.

The proposed communications protocol for the Internet consists of three parts:

4.1 Preliminaries

The first of the users, A, (by tradition called Alice), decides on a specific virtual cryptographic machine [3], [4]. Then a series of N different sets of parameter values for the regulation of the machine is chosen. Each set of parameter values may be used for the creation of a specific cryptographic key. The codes created by these sets of parameters are checked by the methods of lumping and gliding entropy [12]. The keys that are created this way are checked and verified for incompressibility and apparent randomness by the method of K. Karamanos, (or any other method or combination of methods are preferred by the users) [16]. The created keys are classified as emergency keys and filed under the generic name E.K. Each key is labeled by a number, EK1, EK2, … EKN. Then another key is prepared in the same way and given the label K1.

All the above information, the description of the virtual encryption machine, the procedure of its regulation and use, the set of emergency keys EK1 … EKN and the key K1 are the contents of the first message M0 to be sent from A (Alice) to the second user B (Bob).

4.2 Initiation of Communication

User A will have to choose a secure method of sending the initial message M0 to the user B. This method may be an established protocol of Quantum Key Distribution. Since The QKD method will be used only once to initiate the communication process and not for everyday communications, the users will not have to bear the expense in money and time usually associated with these protocols. But the security of communications will be the same as that of a QKD protocol, thanks to the properties of the virtual encryption machine.

Or the user may prefer any secure method such as personal contact or the use of trusted messengers. It should be stressed again that this method will be used only once, to initiate the communication.

By sending the message M0, Alice will have established a communications line with Bob which will offer them the security of Vernam type (or one time pad) methods with an ease and speed of used comparable to that of sending a simple e-mail. So, the message M1 would be encrypted and subsequently decrypted by the use of K1. The message M1 should contain a set of parameter values for the key K2, to be used for message M2, and so forth and so on, theoretically for as long as desirable.

4.3 Security Countermeasures

In theory, the communications method is secure as far as the protocol steps are followed faithfully and there is no inside leak of essential information. This security is based mainly on the apparent randomness and the extremely large set of keys that the method has the potential to create. In real life there are many ways that vital bits of information may come into the hands of illegitimate third parties and then be used to compromise the security of the whole protocol.

There are many methods of attack by eavesdroppers against both classical and quantum key distribution protocols [1] (and references therein). There are also various methods of defense and a method of communication should have some degree of adaptability in its tactics. So against an eavesdropper, (traditionally called Eve), who is trying a split universe attack, there are some emergency keys, EK1 to EKN, to resume communications and (so to speak) ‘sort’ the split created by Eve. Against somebody trying to ‘listen in’ on the communication lines so as to learn something of the hardware utilized in the protocol, (Trojan horse attacks), the encryption machine gives Alice and Bob the capability to ‘spam’ the opposition by a deluge of ‘dummy’ messages consisting of pseudorandom series of digits. Such dummies would automatically be rejected by the legitimate users as unintelligible, but would cost the prospective Eves a disproportionate amount of time and resources [1], [2], [3], [4], [19]. The subject is under study but obviously the concept of a protocol capable to adapt to specific forms of

especially dangerous attacks is very desirable in an Internet application.

5 Incompressibility and apparent randomness of the series created. The lumping and gliding entropy method.

A fundamental part of the proposed method of communication is the evaluation and verification of the created keys for incompressibility, therefore for apparent randomness. For a series of pseudorandomly generated numbers in the binary system there are many standard methods of evaluation [16]. The proposed procedure includes an innovative method of studying number series [12]. It is based on the concept of lumping and gliding entropy.

If a series of binary digits, 0 or 1, is totally random, the possibilities of randomly choosing a digit and finding it to be either 0or 1 should be exactly p(0)= p(1) = ½. Then the Shannon Entropy E = - {p(0)ln(p(0)) + p(1)ln(p(1))} should lave a maximum value, E = ln(2). By a natural generalization, if a 'word' of n digits, chosen from a long series of N digits, n << N, in a specific way, has a possibility of having a certain content of digits. The cases to be considered are 2^n in number and the possibility p(i) of a word $A_i(n) = a_1a_2\ldots a_n$ will take the value of $1/2^n$ in the case of a totally random series. The basic methods of obtaining words of n digits are two: the lumping method and the gliding method.

In the lumping method, in a series of N digits, the first word consists of the elements 1 to n, the next one of the digits n + 1 to 2n etc. to N. In the gliding method, the first word is the digits 1 to n, the second the digits 2 to n + 1, etc. In the above mentioned paper [12], evidence is given that the lumping method is the most reliable. Either way, the normalized entropy $E(n) = \{\sum p_i(n)\ln(p_i(n))\}/n$, i = 1, 2,..$2^n$ should take the maximum value E(n) = ln(2) in the case of complete randomness and lesser values otherwise.

Obviously this method has the added advantage that it can be applied to series of numbers other than digital, such as decimal numbers or the four digits series of 0, 1, 2 and 3 created by the systems of Equations (I), (II), and (III). The application of the above method to samples of symbolic series created by the two dimensional system of Equation (III) was successful, indicating almost complete incompressibility. [Fig 4].

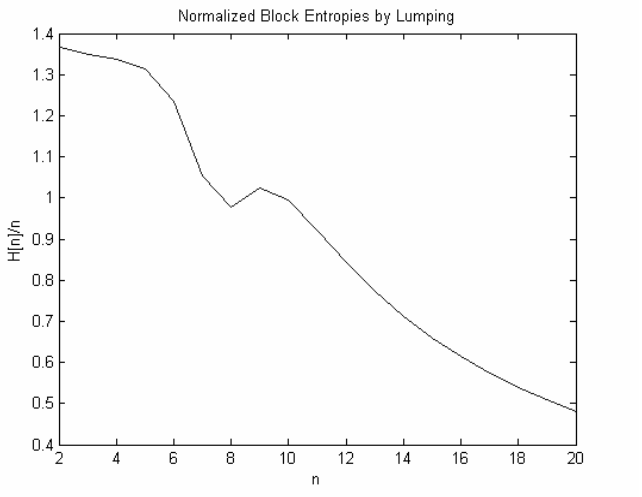

Fig.4 Normalized block entropies by Lumping

6. Conclusion

The virtual encryption machine in the form of a discrete dynamic system with discontinuity presented in this work has the properties required for application in random number generation and cryptographic key application in a digital network environment. The high level of security achieved by the incompressibility and apparent randomness of the created keys, the very large number of keys the virtual encryption machine has the ability to produce and the obvious repeatability of the process of key creation make the protocol proposed here especially attractive for Internet applications. The innovative methods of evaluation and verification of apparent randomness [12], increase the level of security and reliability of the protocol against some very dangerous forms of eavesdroppers' attacks. Finally, the concept of adaptation of the protocol to defend in real time against specific forms of attacks [2], although still under study, seems very promising for security and protection against even some conceived forms of attack that are not realistic at the present level of technology but may very well present a real threat at the near future. So it seems that further study and development of the protocol presented here may lead to substantial advances in the technology and methods of communication security.

7. Acknowledgements.

The authors would like to express their thanks to Prof. S. Kotsios, Prof. O. Feely, Prof. A. Bountis and Prof. N. Kalouptsidis for valuable discussions, suggestions and encouragement. They would also like to thank Drs K. Limniotis, P. Risomiliotis and Y. Kominis, as well as S. Domoxoudis, Y.

Loukidis and E. Koukianakis for valuable cooperation and assistance in the creation of this work.

References

[1] Yannis Bakopoulos, 'Application of Dynamic Systems for Cryptographic Key Distribution' *15th Congress on Nonlinear Dynamics, Chaos and Complexity* Patras Aug. 19 – 30, 2002 (A. Bountis).

[2] Yannis Bakopoulos, Yannis Vrettaros, Athanasios Drigas, 'An automatic process for the reliable and secure creation and distribution of quantum keys' *National Patent No 1003891, OBI*, 2002.

[3] Yannis Bakopoulos, Vassiliki Soulioti, 'A protocol for secure communication in digital networks' *National Patent No 1004308 OBI, 2003*.

[4] Yannis Bakopoulos, Vassiliki Soulioti, 'A protocol for secure communication in digital networks' PCT/GR 03/ 00035 2003.

[5] L. O Chua. and T.Lin, 'Fractal Pattern of second-order non-linear digital filters: A new symbolic analysis' (1988) *IEEE Trans. CAS 35*, pp. 648 – 658.

[6] Robert L Devaney. *Physica* 10D (1984) pp. 387 – 393.

[7] O. Feely and L. O. Chua 'Nonlinear Dynamics of a class of analog - to - digital converters', *Int. J. Bifurcation and Chaos*, Vol. 2, 1992, pp. 325 – 340.

[8] Orla Feely "Nonlinear Dynamics and Chaos in Sigma – Delta Modulation", Journal of the Franklin Institute Vol. 331B, No. 6, 1995 pp. 903 – 936.

[9] Orla Feely 'Nonlinear Dynamics of Chaotic Double-Loop Sigma Delta Modulation', ISCAS 1994: pp.101-104.

[10] T. Habutsu. et al. 'A secret key cryptosystem by iterating a chaotic map' *International Conference on the Theory and Application of Cryptographic Techniques,*Springer Verlag, DE pp 127 – 140, XP000607774

[11] Leo P. Kadanov, and Chao Tang, *Proc. Natl. Acad. Sci. USA* Vol. 81, pp. 1276 – 1279, February 1984, Physics.

[12] K. Karamanos "Entropy analysis of substitutive sequences revisited" J. *Phys. A, Math. Gen*. 34, (2001) 9231 – 9241.

[13] Stelios Kotsios and Orla Feely *NDES Congress Spain* '96.

[14] Stelios Kotsios and Orla Feely 'The model – matching problemfor a special class of discrete systems with discontinuity'*IMA Journal of Mathematical Control & Information* (1998) Vol. 15, pp 93 – 104.

[15] Stelios Kotsios 2000 'Symbolic Sequences Generated by a Special Class of Discrete Sysems with Discontinuity and Input' *Nonlinear Dynamics* 22 pp.175 – 191 (and refs therein).

[16] George Marsaglia "A Current View of Random Generators" Keynote Address, *Computer Sciense and Statistics: 16th Symposium on the Interface*, Atlanta, 1984 (It appeared in *"The Proceedings"* of the Conference, published by Elsevier Press).

[17] S. Papadimitriou, A. Bezerianos, T. Bountis, G. Pavlides, "Secure Communication protocols with discrete nonlinear chaotic maps", *Journal of Systems Architecture*, Vol. 47, No 1, 2001, pp. 61 – 72.

[18] James Rössler et al., *PHYSICAL REVIEW A*, VOLUME 39, NUMBER 11, JUNE 1 1989, pp.5954 – 5960.

[19] V. Soulioti 'A study on Discrete Dynamic Systems with a Linear Part and Discontinuity', *15th Congress on Nonlinear Dynamics, Chaos and Complexity* Patras Aug. 19 – 30, 2002 (A. Bountis).

[20] Richard J. Hughes et al 'Method and apparatus for free space quantum key distribution in daylight' *US 2001/055389*, December 27, 2001.

[21] Yuan et al 'Method and system for establishing a cryptographic key agreement using linear protocols', *US 5 966 444,* Oct. 12 1999

[22] Tohru Kohda et al 'Enciphering/Deciphering apparatus and method incorporating random variable and keystream generation' *USPatent 6 014 445* Jan 11, 2002.