3D Visual Method of Variant Logic Construction for Random Sequence

Huan Wang and Jeffrey Zheng

Abstract As Internet security threats continue to evolve, in order to ensure information transmission security, various encrypts and decrypts have been used in channel coding and decoding of data communication. While cryptography requires a very high degree of apparent randomness, random sequences play an important role in cryptography. Both Cellular Automata (CA) and RC4 contain pseudorandom number generators and may have intrinsic properties, respectively. In this chapter, a 3D visualization model 3DVM is proposed to display spatial characteristics of the random sequences from CA or RC4 keystream. Key components of this model and core mechanism are described. Every module and their I/O parameters are discussed, respectively. A serial of logic function of CA is selected as examples to compare with some RC4 keystreams to show their intrinsic properties in three-dimensional space. Visual results are briefly analyzed to explore their intrinsic properties including similarity and difference. The results provide support to explore the RC4 algorithm by using 3D dimensional visualization tools to organize its interactive properties as visual maps.

Keywords Pseudorandom sequence · CA · Stream cipher · RC4 keystream
3D maps

1 Introduction

Wireless Sensor Networks WSN and Wireless Networks WN are most popular and widely used types of network of this era. Because of the openness these types of

Project supported by NSF of China (613620214), the Key R&D project of Yunnan Higher Education Bureau (K1059178) and Yunnan Advanced Overseas Scholar Project (W8110305).

H. Wang
Yunnan University, Kunming, China
e-mail: lights127@gmail.com

J. Zheng (✉)
Key Laboratory of Yunnan Software Engineering, Yunnan University, Kunming, China
e-mail: conjugatelogic@yahoo.com

J. Zheng (ed.), *Variant Construction from Theoretical Foundation to Applications*,
https://doi.org/10.1007/978-981-13-2282-2_15

networks are not very much secure. To provide the security over the WSN and WN, algorithm used must be fast enough which can encrypt and decrypt data comparatively in less amount of time to require less resource too. In this concern, Wi-Fi Protected Access WPA and Wired Equivalent Privacy WEP protocols are used as standard. These standards have adopted the RC4 stream cipher algorithm to secure the data over the WN environment. These standard adopted RC4 algorithms because RC4 algorithm gives speedy encryption and decryption of data, utilize less hardware resource during processing, and easy to implement [1, 2]. Presently, RC4 algorithm is not secure in many aspects. Lots of weaknesses and attacks have been detected by the cryptanalysis [3, 4].

1.1 The Weakness of RC4

RC4 algorithm is a stream cipher under the symmetric ciphers algorithms. Typically, in a stream cipher, the keystream is the sequence which is combined digit by digit to the plaintext sequence for obtaining the ciphertext sequence. However, the data encryption is equivalent to a simple XOR with keystream. The keystream is generated by a finite state automaton called the keystream generator [5, 6]. The encryption can be broken if the plaintexts are encrypted using the same keystream. RC4 keystream generated by RC4 keystream generator is completely compromising the security of RC4.

Because it is very hard to trace the characteristics of keystream generators, random characteristics of keystream can be investigated on spatial characteristics of the keystream generator to test pseudorandom sequences. This chapter is the expansion work of [7] by Qingping Li from 2D to 3D. In this chapter, random sequences from given keystreams are collected in comparison with random sequences generated by sample logical function of 1D Cellular Automata to show their intrinsic properties in three-dimensional space of relationships.

1.2 CA

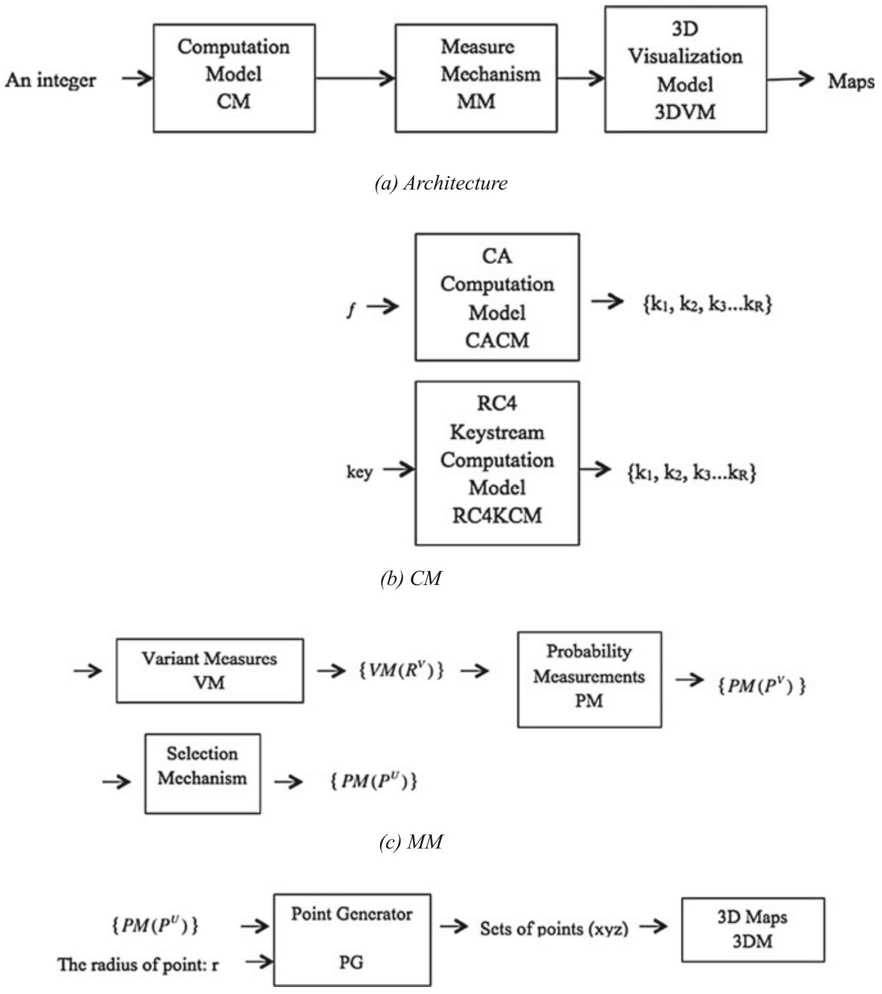
Cellular Automata is a great discovery in the twentieth century, and it forms a time series according to a given function in an iterations process by introducing logic function and related calculation methods in the natural pattern [8]. In 1985, S. Wolfram formed the sequential cipher from pseudorandom sequence generated from logic calculation using cellular automata. Because of the implicated expression of the logic function, the spatial characteristic cannot be directly observed from the function formula [9].

2 Architecture

2.1 Architecture

The architecture is shown in Fig. 1a. The three main components and their modules are shown in Fig. 2b–d, respectively.

In the first part of this system, two types of data sets are generated by CACM and RC4KCM, respectively. The data sets on either CACM or RC4KCM get into



(a) Architecture

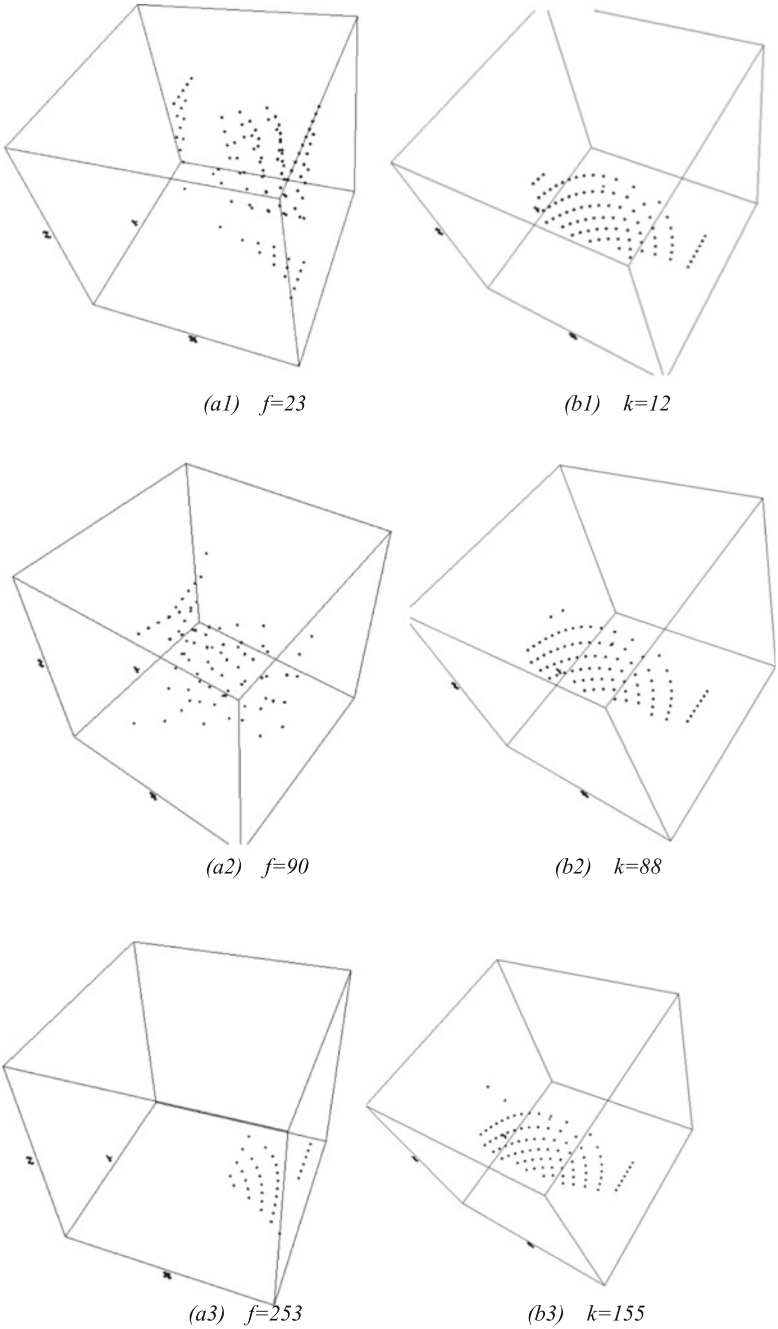
f
CA Computation Model CACM
{k1, k2, k3...kR}
key
RC4 Keystream Computation Model RC4KCM
{k1, k2, k3...kR}

(b) CM

(c) MM

(d) 3DVM 3D Visualization Model

Fig. 1 Variant 3D visualization system and key components



(a1) f=23

(b1) k=12

(a2) f=90

(b2) k=88

(a3) f=253

(b3) k=155

Fig. 2 Two sets of six 3D maps based on unified model in different conditions; **a**1–**a**3 for the file CA; **b**1–**b**3 for the file RC4

the MM module as input data. The main function of the VM is to output the four vectors of variant measurements. Using unified or non-unified method, six probability measurements are created by PM module. In order to establish 3D maps, three vectors of probability measurements are selected from the six probability measurements by the SM module. Three vectors determine a 3D spatial position. All vectors generate a 3D map using 3DVM.

There are six parameters in an input group, three sets of parameters in the intermediate group, and one set of parameters in the output group.

Input Group:

An integer indicates the serial number of logic function or the value of the key selected
An integer indicates which model is selected
An integer indicates the number of elements in the binary sequence
An integer indicates the number of elements in a segment
An integer indicates the method of selection mechanism
An integer indicates the control parameter for mapping

Intermediate Group:

A 0-1 vector generated by CA logic function or RC4 keystream generator
A set of four variant measures
A set of six probability vectors

Output Group:

3D maps

2.2 Computation Model of CA (CMCA)

CMCA module is used to measure the features of a logic function based on Cellular Automata (CA). Consider a logic function f: $Y = f(X)$ as a function of CA, the output sequence Y can be generated by the given initial input sequence X with 2 states. For N bits initial input sequence, a total of 2^n states are generated under the logic function f: $X \rightarrow Y$. A pair of vectors (X, Y) could be collected for their correspondences on the pair of input–output relationships. There are 2^n groups of this corresponding relationship.

Input Group:

X A 0-1 vector with N elements, $X \in B_2^n$
n An integer indicating a 0-1 vector with n elements,
f A function with 2 variables

Intermediate Group:

Y A 0-1 vector with N elements, $Y \in B_2^n$

Output Group:

$\forall Y$ Exhaustive set of all states of N bit vectors with 2^n elements

2.3 Computation Model of RC4 Keystream (RC4KCM)

For an L bits input keystream K, divided into G segments and $W = L/G$ bits of each segment with $G < L$. The value of parameter G determines the amount of points and W determines the spatial distribution for the output keystream in the phase space.
Input Group:

A 0-1 vector with L elements generated by RC4 keystream generator

L An integer indicates the number of elements in an input sequence,
G An integer indicates the number of segments divided,
W An integer indicates the number of elements in a segment.

Output Group:

G sets of W bits 0-1 vectors

The CMRC4 component uses an input vector as input, under different segment strategies to divide into several segments. The output of this component is G sets of W bits 0-1 vectors.

2.4 Measure Mechanism (MM)

The MM component shown in Fig. 1c is composed of three modules: Variant Measure (VM), Probability Measurement (PM), and Selection Mechanism (SM). Three parameters are listed as input signals; four variant measures are outputted from VM module, six probability measurements are created from variant measures by Probability Measurement (PM), under the Selection Mechanism (SM) module, and a set of triples interactive projections is selected.
Input Group:

V A symbol is selected from four types of transformations {⊥, +, −, T},
N An integer indicates the number of elements in an input vector

A 0-1 data vector
Intermediate Group:

$VM(R^V)$ A set of four variant measures
$PM(P^V)$ A set of four probability vectors

Output Group:

$U \subset V$ A set of three interactive projections under the SM condition, $U \subset V$
$PM\left(P^U\right)$ A set of three probability vectors

2.5 Variant Measure (VM)

Considering the transformation of every bit between input sequence $\{X_i\}_{i=0}^{N-1}$ and output sequence $\{Y_i\}_{i=0}^{N-1}$, there are a total of four types of transformations: 0 → 0, 0 → 1, 1 → 0, and 1 → 1 [10, 11].

Define the variant representation as follows:

$$V = \begin{cases} \perp, X_i = 0, Y_i = 0; & \\ +, X_i = 0, Y_i = 1; & 0 \le i \le N, \quad X_i, Y_i \in B_2 \\ -, X_i = 1, Y_i = 0; & \\ \mathrm{T}, X_i = 1, Y_i = 1; & \end{cases}$$

For any N bit 0-1 vector $X, X = X_0X_1 \ldots X_i \ldots X_{N-1}X_N, 0 \le i \le N, X_i \in B_2, X_i \in B_2^N$ under 2-variable function f, N bit 0-1 output vector $Y, Y = Y_0Y_1 \ldots Y_i \ldots Y_{N-1}Y_N, 0 \le i \le N, Y_i \in B_2, Y_i \in B_2^N$. Let Δ be the variant measure function.

$$\Delta(X \to Y) = \sum_{i=0}^{N-1} \Delta(X_i \to Y_i) = \langle R_\perp, R_+, R_-, R_\mathrm{T} \rangle,\ N = R_\perp + R_+ + R_- + R_\mathrm{T}, R_0$$
$$= R_\perp + R_+, R_1 = R_- + R_\mathrm{T}$$

Example
$N = 13$, $Y = f$ (X).

$$\mathrm{X} = 1001011100101$$
$$\mathrm{Y} = 0010110101100$$
$$\Delta(X \to Y) = -\perp + - + \mathrm{T} - \mathrm{T}\perp + \mathrm{T}-$$
$$\langle R_\perp + R_+ + R_-, R_\mathrm{T} \rangle = \langle 3, 3, 4, 3 \rangle, R_0 = 6, R_1 = 7, N = 13$$

Input and output pairs are 0-1 variables for only four combinations. For any given function, the quantitative relationship of $\{\perp, +, -, \top\}$ is directly derived from the input/output sequences. Four meta measures are determined [12].

Input Group:

V A symbol is selected from four types of transformations $\{\perp, +, -, \mathrm{T}\}$,
N An integer indicates the number of elements in an input vector

A 0-1 data vector

Output Group:

$VM\left(R^V\right)$	A set of four variant measures
R_0	An integer indicates the number of 0 in an input vector
R_1	An integer indicates the number of 1 in an input vector

2.6 Probability Measurement (PM)

Variant measure parameters and the other three parameters are listed as input signals; the output of probability signals is calculated as eight measurements in two groups by following the given equations.

The first group of probability signal vectors ρ is called a non-unified model and defined as follows:

$$\begin{cases} \rho = \frac{R^V}{N} = R_{\perp}, R_{+}, R_{-}, R_{\mathsf{T}} \\ \rho_{\alpha} = \frac{R_{\alpha}}{N}, \alpha \in \{\perp, +, \_, \mathsf{T}\} \end{cases} \quad \& \quad \begin{cases} \rho_0 = \frac{R_0}{N} \\ \rho_1 = \frac{R_1}{N} \end{cases}$$

The second group of probability signal vectors $\tilde{\rho}$ is called a unified model and defined as follows:

$$\begin{cases} \tilde{\rho} = \frac{R^V}{R_0 | R_1} = R_{\perp}, R_{+}, R_{-}, R_{\mathsf{T}} \\ \rho_{\alpha} = \frac{R_{\alpha}}{R_0}, \alpha \in \{\perp, +\} \\ \rho_{\beta} = \frac{R_{\beta}}{R_1}, \beta \in \{\_, \mathsf{T}\} \end{cases} \quad \& \quad \begin{cases} \rho_0 = \frac{R_0}{N} \\ \rho_1 = \frac{R_1}{N} \end{cases}$$

Under such condition, the output signals of the PM module can be expressed as a pair of probability vectors in quaternion forms $PM\left(P^V\right) = \{\rho, \tilde{\rho}\}$.

Input Group:

V	A symbol is selected from four types of transformations $\{\perp, +, -, \mathrm{T}\}$,
N	An integer indicates the number of elements in an input vector
$VM\left(R^V\right)$	A set of four variant measures
R_0	An integer indicates the number of 0 in an input vector
R_1	An integer indicates the number of 1 in an input vector

Output Group:

$PM\left(P^V\right)$	A set of four probability vectors

2.7 Selection Mechanism Module

The SM Module is composed of two models: Non-unified Model and Unified Model. Under different constructions, two models are established respectively as follows.

Non-unified Model

Selecting two measurements from four combinations $\{\tilde{\rho}_{\perp}, \tilde{\rho}_{+}, \tilde{\rho}_{-}, \tilde{\rho}_{\mathrm{T}}\}$, there will be C_4^2 choices. And then selecting one measurement from two combinations $\{\rho_0, \rho_1\}$, there will be C_2^1 choices. A 3-tuple S is defined as follows:

$$\begin{cases} S = (\rho_\alpha, \rho_\beta, \rho_\gamma) \\ S' = (\rho_\beta, \rho_\alpha, \rho_\gamma), \quad \alpha,\ \beta \in V,\ \gamma \in \{0, 1\},\ \alpha \neq \beta \\ S = S' \end{cases}$$

Unified Model

Selecting two measurements from four combinations $\{\tilde{\rho}_{\perp}, \tilde{\rho}_{+}, \tilde{\rho}_{-}, \tilde{\rho}_{\mathrm{T}}\}$, there will be C_4^2 choices. And then selecting one measurement from two combinations $\{\rho_0, \rho_1\}$, there will be C_4^2 choices. A 3-tuple $\tilde{S}$ is defined as follows:

$$\begin{cases} \tilde{S} = (\tilde{\rho}_\alpha, \tilde{\rho}_\beta, \tilde{\rho}_\gamma) \\ \tilde{S}' = (\tilde{\rho}_\beta, \tilde{\rho}_\alpha, \tilde{\rho}_\gamma), \quad \alpha,\ \beta \in V,\ \gamma \in \{0, 1\},\ \alpha \neq \beta \\ \tilde{S} = \tilde{S}' \end{cases}$$

Under such condition, the output signals of the SM module can be expressed as a 3D visual model in 3-tuples forms S or $\tilde{S}$. Specifically ρ_α or $\tilde{\rho}_\alpha$ determines the value of X-axis, ρ_β or $\tilde{\rho}_\beta$ determines the value of Y-axis, and ρ_γ or $\tilde{\rho}_\gamma$ determines the value of Z-axis.

Input Group:

$PM(P^V)$	A set of four probability vectors

Output Group:

$U \subset V$	A set of three interactive projections under the SM condition, $U \subset V$
$PM(P^U)$	A set of three probability vectors

2.8 Visualization Model

Using a visual model, *all possible measurements are calculated exhaustively on all G-1 vectors. Each 3-tuple* can be drawn as a point in three-dimensional space (*xyz*-space). All G-1 points are constructed in the phase space for the selected keys.

3 Sample Results on 3D Maps

In this section, two types of data sets are selected to illustrate their differences on 3D maps for comparison. The first type of data sets is generated by CA. The second type of data sets is generated by RC4.

3.1 Visualization Results of Unified Model

See Fig. 2.

3.2 Visualization Results of Non-unified Model

See Fig. 3.

3.3 Visualization Results of CA with Different Length of Initial Sequence

See Fig. 4.

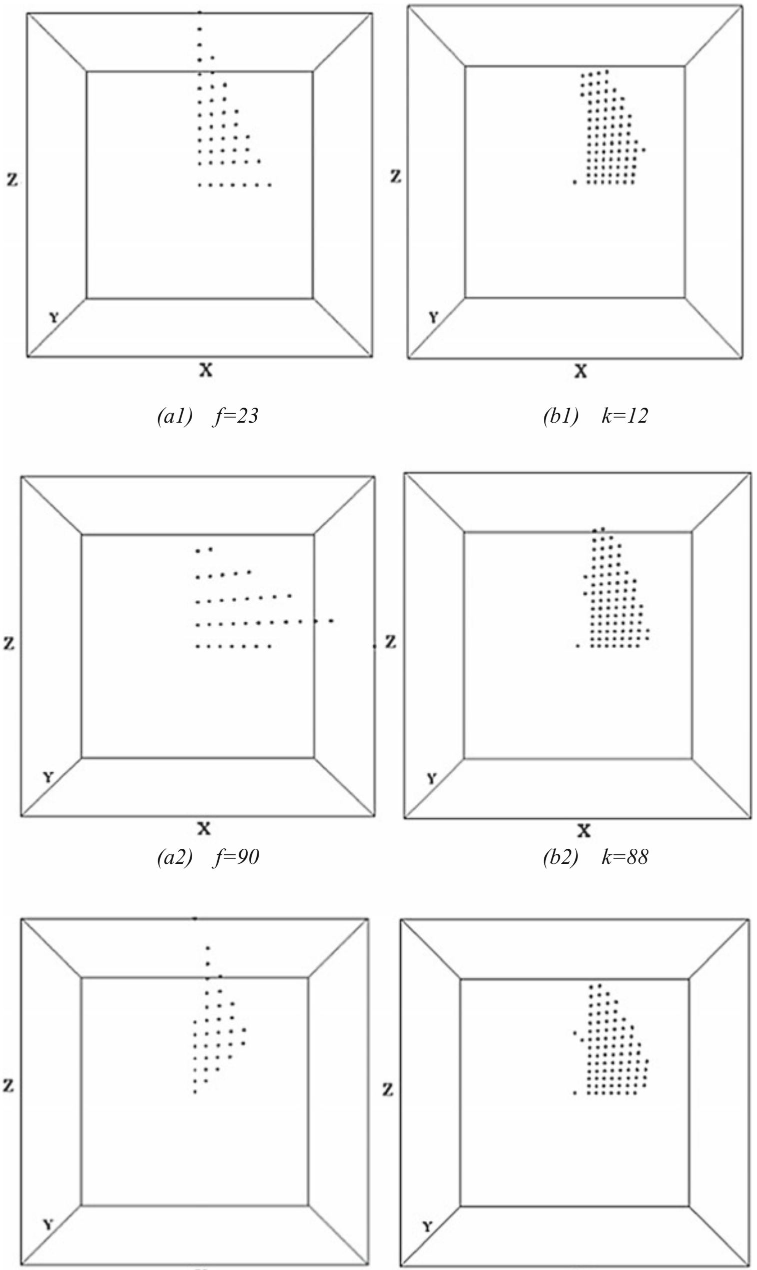
3.4 Visualization Results of RC4 Keystream with Different Segment Strategies

See Fig. 5.

4 Analysis of Results

The above 27 3D maps contain different information. Some important conclusions will be discussed in detail in this section.

The first group of results shown in Fig. 2 presents two sets of six 3D maps constructed by the unified model from two data files: CA and RC4 to illustrate their 3D spatial characteristics. Three 3D maps of each group in Fig. 2a1–a3 show 3D spatial characteristics of CA with different logic functions. In this group, No. 23, 90, 253 functions are selected as examples to compare each other. And three 3D maps of each group in Fig. 2b1–b3 show 3D spatial characteristics of RC4 with 20



(a1) f=23

(b1) k=12

(a2) f=90

(b2) k=88

(a3) f=253

(b3) k=155

Fig. 3 Two sets of six 3D maps based on non-unified model in different conditions; **a**1–**a**3 for the file CA; **b**1–**b**3 for the file RC4

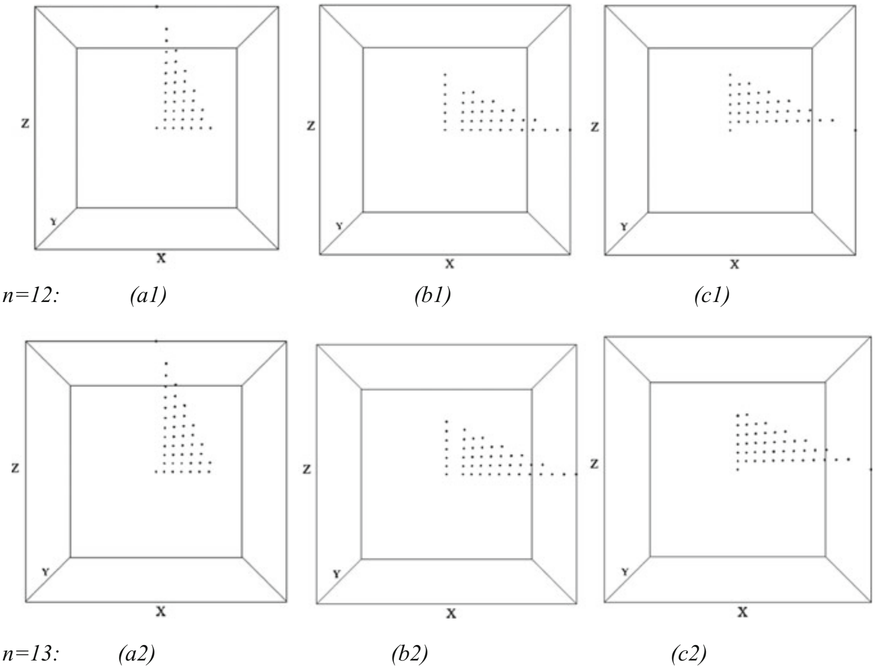


Fig. 4 Three sets of nine 3D maps under different conditions; **a**1–**a**2 for the logic function $f = 15$ and non-unified model; **b**1–**b**2 for the logic function $f = 100$ and non-unified model; **c**1–**c**2 for the logic function $f = 170$ and non-unified model

bits of every segment and different given keys. In this group, keys: 12, 88, and 155 are selected as examples to compare each other. From a distribution viewpoint, different logic function can be distinguished by their three-dimensional spatial characteristics from CA files, e.g., (a1–a3). Different from CA, for RC4 keystream, all spatial distributions are always in a plane, e.g., (b1–b3).

The second group of results shown in Fig. 3 presents two sets of six 3D maps constructed by non-unified model. It is interesting to observe that all maps (no mater CA data files or RC4 keystream data files) have planar distribution, e.g., (a1–a3) and (b1–b3).

The third group of results shown in Fig. 4 presents three sets of six 3D maps constructed by non-unified model from CA data files with different lengths of the initial sequence and given logic functions. Figure 4a1–a2 shows 3D maps for the No. 15 function, (b1–b2) shows 3D maps for the No. 100 function, and (c1–c2) shows 3D maps for the No. 170 function. The overall relationship of multiple-variable logic functions for spatial characteristics can be shown clearly. For example, under the non-unified model, no matter what logic functions are, all spatial distributions are always in a plane, e.g., (a1–a2), (b1–b2), and (c1–c2). Different lengths of initial

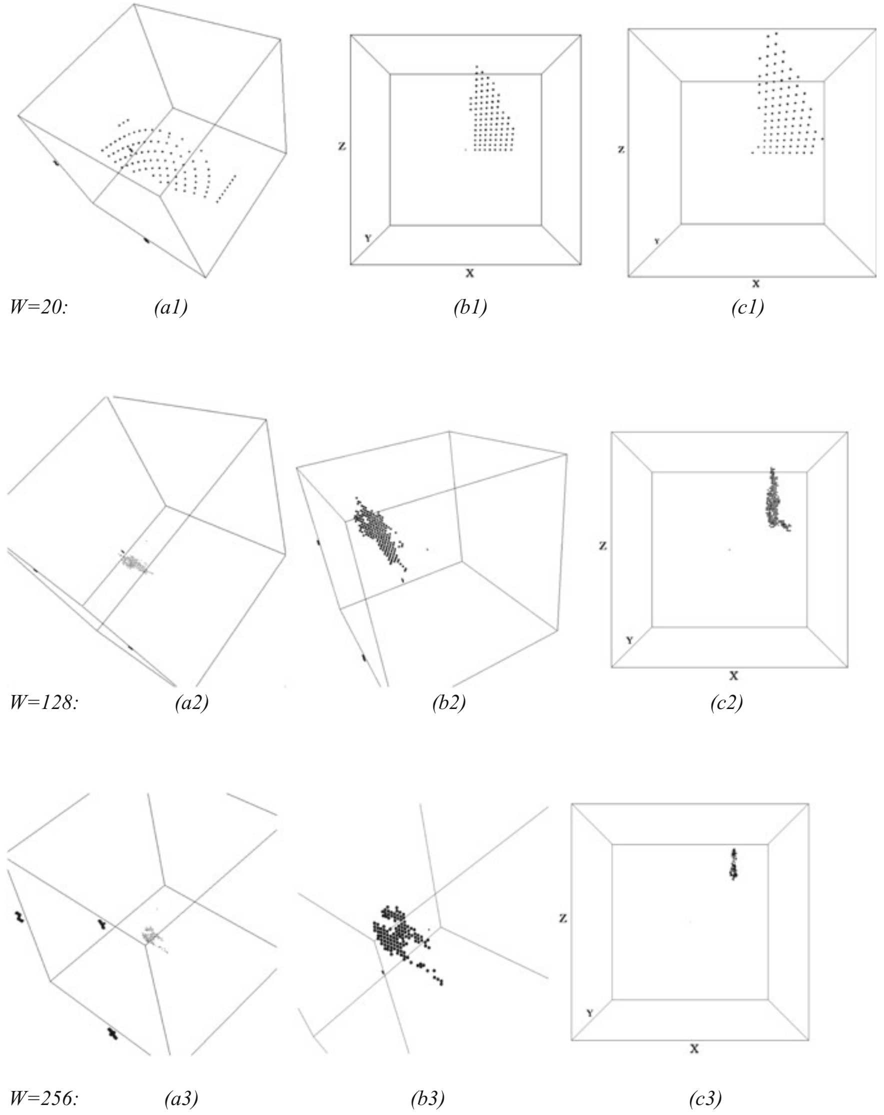


Fig. 5 Three sets of nine 3D maps under different conditions; **a**1–**a**3 for the key = 90 and unified model; **b**1–**b**3 for the key = 90 and non-unified model; **c**1–**c**3 for the key = 123 and non-unified model

sequence ($n = 12, 13$) have different spatial characteristics distribution with the same given logic function, e.g., (a1–a2), (b1–b2) and (c1–c2).

The fourth group of results shown in Fig. 5 presents three sets of nine 3D maps for the different conditions including segments strategies and keys. In this group, three types of segment strategies ($W = 20$, 128, 256) are proposed to compare.

Combinations of three set use the same key e.g., (a1–a3), (b1–b3), and (c1–c3) to observe them conveniently. The dispersity of points increased with reducing the bit length of each segment. Obviously, the spatial distribution of points with 256 bits of each segment is more concentrated than the distribution of points with 20 bits, as shown in (a1–a2), (b1–b2), and (c1–c2). 3D map shows some commonalities of the spatial distribution of different keys and different segment strategies. First, under this construction, different keys can be distinguished by their three-dimensional spatial characteristics in the model, e.g., (b1–c1), (b2–c2), and (b3–c3). Second, no matter what keys or segment strategies are, all spatial distributions are always in a plane. Third, the distribution features are varying from key to key and segment strategy to segment strategy.

5 Conclusions

Both the similarities and the differences may indicate those maps with comparable mechanism to express keystream with different given keys and in their high levels of relationships applying to the stream cipher mechanism. The spatial property of random sequence can be detected from the distribution of cluster point in the 3D maps discussed in details. Different spatial distributions are illustrated to show various distributions on each phase space for relevant logic function or keystream. For example, no matter what keys or segment strategies are, all spatial distributions are always in a pane. And all maps (no mater CA data files or RC4 keystream data files) are planar distribution under non-unified model. Spatial distribution properties like this provide useful information for further exploring the RC4 stream cipher. This construction could provide remarkable insights to spatial information on stream cipher construction via 3D maps. Further explorations are required on this scheme.

Acknowledgements Thanks to the school of software Yunnan University, to the key laboratory of Yunnan software engineering for excellent working environment, to the Yunnan Advanced Overseas Scholar Project (W8110305), to the Key R&D project of Yunnan Higher Education Bureau (K1059178), and to National Science Foundation of China (61362014) for the financial support to this project.

References

1. H. Brandon, A.J. Patricia, Information Warfare. Inf. Syst. Educ. J. **4**(49) (2006). http://isedj.org/4/49/. ISSBN: 1545-679X
2. O.S. Suhaila, S.P. Mansoor, Performance analysis of Stream Cipher algorithms, in *The 3rd International Conference on Advanced Computer Theory and Engineering (ICATE)*, Chengdu, China (2010)
3. J.K. Fahime, V.M. Mohammad, R.N. Hamid, H. Payman, A new symmetric cryptographic algorithm to secure E-commerce transactions, in *The International Conference on Financial Theory and Engineering*, Dubai, United Arab Emirates (2010)

4. C.S. Lamba, Design and analysis of Stream Cipher for Network security, in *The 2nd International Conference on Communication Software and Networks*, Singapore (2010)
5. M.J.B. Robshaw, *Stream Cipher.* RSA Laboratories Technical Report TR-701. Retrieved from http://citeseerx.ist.psu.edu/ (1995)
6. S. Bruce, *Applied cryptography* (Wiley, CRC Press, 1997)
7. Q. Li, J. Zheng, 2D spatial distributions for measures of random sequences using conjugate maps, in *The Proceedings of the 11th Australian Information Warfare and Security Conference*, Perth 1–9, 2010. http://ro.ecu.edu.au/isw/34
8. S. Wolfram, *Theory and Applications of Cellular Automata* (World Scientific Press)
9. L. Shiyong, T. Xinhua, *Nonlinear Study and Complexity Study* (Harbin Institute of Technology Press)
10. J. Zheng, C. Zheng, T.L. Kunii, A framework of Variant Logic Construction for Cellular Automata, in *Cellular Automata—Innovative Modelling for Science & Engineering* (InTech Press, 2011)
11. S. Alejandro, *Cellular Automata-Innovative Modelling for Science and Engineering* (InTech Press, 2011)
12. J. Zheng, C. Zheng, T.L. Kunii, Interactive maps on variant phase spaces—from measurements—micro ensembles to ensemble matrices on statistical mechanics of particle models, in *Emerging Applications of Cellular Automata* (InTech Press, 2013)