UDC 519.216.1

# SEQUENCE INTERNAL STRUCTURE FORMATION DURING PSEUDORANDOM GENERATION

**G. Vostrov, A. Khrinenko**
*Odessa National Polytechnic University*

***Abstract.*** *This work analyzes and examines problems of number sequence generation processes for pseudorandomness. Properties and internal structure of obtained sequences are investigated according to its influence of randomness measure of there sequences.*

***Key words:*** *chaos, pseudorandomness, nonlinear maps, inner structure.*

## Introduction

Different branched of science include so-called randomized algorithms, i.e. that require some source of true randomness for number sequence generation. Those algorithms increase the range of approaches that can be used to efficiently solve various problems. Nevertheless, a question of randomness is still open. In the axiomatization of Kolmogorov the random sequences were left outside the theory and it was proposed only general approaches to definition of randomness, such as von Mises approach. Even if some method for sequence generation is used a problem of uncertainty of approximation to true randomness appears since true random sequence is infinite and cannot be generated for real-world tasks. To solve this problem it was proposed to construct pseudorandom processes which would generate sequences that satisfy some certain requirements, such as degree of approximation to a given distribution law, successful results for some set of statistical tests and others. It turned out to be relatively easy to present explicit functions of such pseudorandom generations, which gives necessary sequences in cases where random-like objects should be used. One of such tools is so-called deterministic dynamic systems that are based on the study of the dynamics of iterative fixed points via recursive functions. Iterative cycles or orbits are considered as deterministic chaos [1], since they depend on initial conditions and demonstrate no regularity [2] and respectively could be used for pseudorandom sequence generation. One more advantage of dynamic systems is that to obtain required level of approximation to randomness dynamic systems allow to combine algebraic equations.

## 1. Computer analysis of inner structure in nonlinear maps

In this paper we continue to investigate processes occurring in maps that represent simple nonlinear dynamic systems [3] for analyzing internal structure of sequence, obtained by these maps. As process we consider a function $F$, that maps finite sequence (word) into sequences so that if for word $x$ value of $F(x)$ is determined and $y \subset x$, then $F(y)$ is also determined and $F(y) \subset F(x)$. Let's $\omega$ - some sequence. Process $F$ will be applied as long as it is possible. As a result we obtain parts of some new sequence S – result of application $F$ to $\omega$, that is $S = F(\omega)$. It is important to remind that maps (1, 2, 3) are continuous functions and map (4) is determined only on set of integers. Analyzed in this paper maps allow to make some conclusions of general consistent patterns in chaotic processes that emerge in complex dynamic systems. The maps are represented as follows:

$$x_{n+1} = \begin{cases} 2x_n, & 4x_n < p \\ p - 2x_n, & 4x_n \geq p \end{cases} \quad (1)$$

$$x_{n+1} = \begin{cases} 2x_n, & 2x_n < p \\ p - x_n, & 2x_n \geq p \end{cases} \quad (2)$$

$$x_{n+1} = \begin{cases} 2x_n, & 2x_n < p \\ 2x_n - p, & 2x_n \geq p \end{cases} \quad (3)$$

$$x_{n+1} = 4x_n (\mathrm{mod}\, p), \quad (4)$$

Presented nonlinear maps allowed to divide set of prime numbers $p$ into the class system, which is based on the length of the iteration cycles as a function of given prime numbers. It should be noted there is an infinite set of prime numbers for which the length of the period is significantly smaller than the dimension of the number and therefore form simple sequence structure. Table 1 shows the behavior of presented maps for different prime numbers. It is worth to mention, that map (1) algebraically congruent to map (4) on set of integers.

Table 1

Length of the period for generated sequences

| prime number | *m(1)* | *m(2)* | *m(3)* | *m(4)* |
|---|---|---|---|---|
| 148587941 | 74293970 | 111440955 | 148587940 | 74293970 |
| 148587949 | 142 | 193 | 284 | 142 |
| 148587953 | 37146988 | 55717993 | 74293976 | 37146988 |
| 164511349 | 82255674 | 123383511 | 164511348 | 82255674 |
| 164511371 | 82255685 | 123383528 | 164511370 | 82255685 |
| 168410987 | 84205493 | 126308240 | 168410986 | 84205493 |

For given prime numbers Fig. 1 and 2 represents the internal structure of the iterative processes in the maps, where dotted line shows obtained sequences and solid line shows internal parts inside sequences that gives maximum correlation value. This value is used to identify similar subsequences. As it is seen on these figures sequences that were obtain with map (1) and (2) for some subsequences give correlation value close to 1 that indicate influence of fixed points on sequence inner structure.

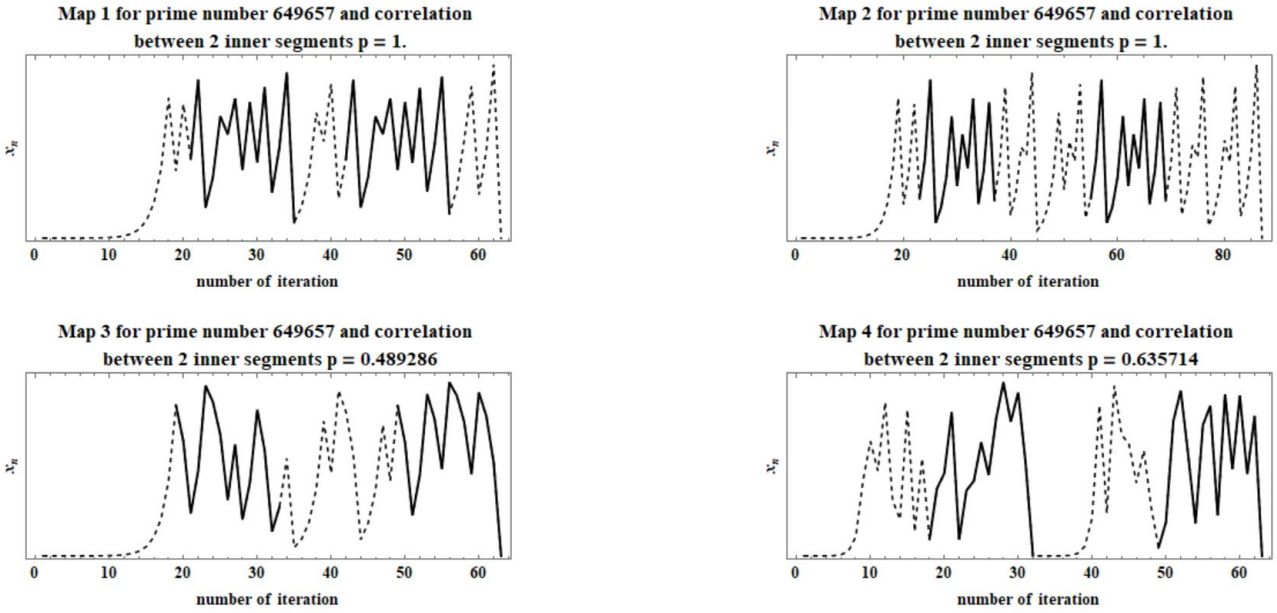


Fig. 1 – Sequence structure for the maps and prime number 649657

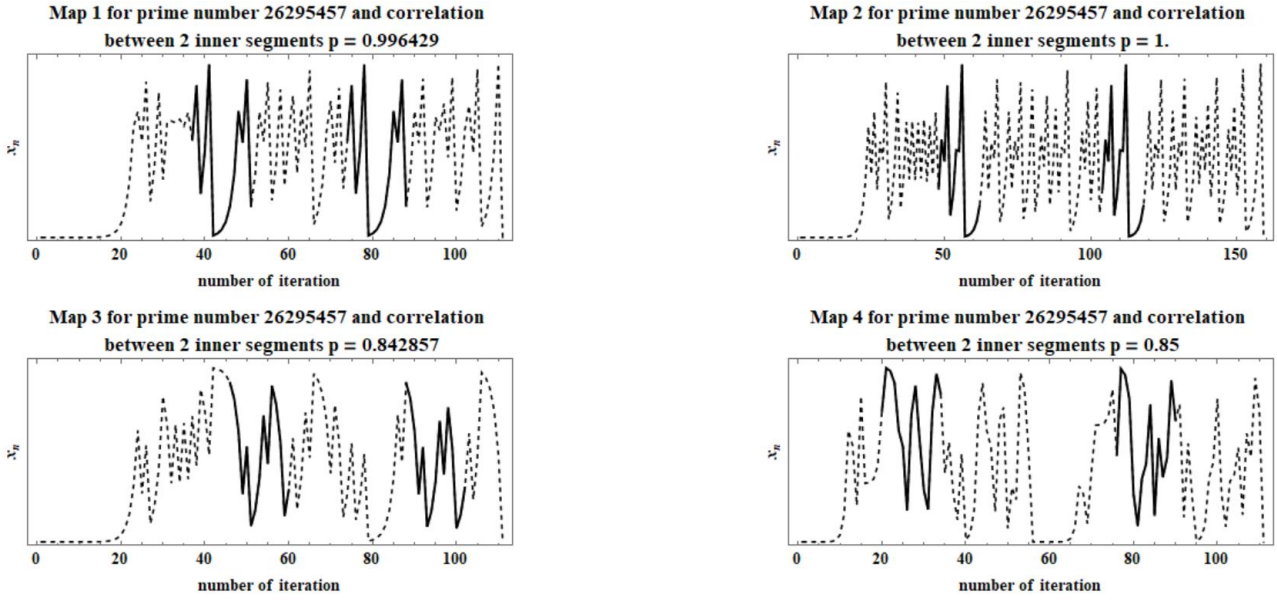


Fig. 2 – Sequence structure for the maps and prime number 26295457

**2. Approaches to definition of randomness and its measures**

There are several approaches to determining randomness and, accordingly, methods for evaluating the measure of randomness of a particular sequence. In work [4], four algorithmic properties are distinguished for describing randomness: frequency stability, chaotic character, typicality, unpredictability. Each of them represents its own algorithmic face of randomness, and each of them, with more or less grounds, can claim to be a strict mathematical definition of the concept of randomness.

This paper analyzes chaotic character and unpredictability of the internal structure of the generated sequences to form the concept of a truly random sequence. When considering the concept of randomness, the Kolmogorov complexity theory is used, where core idea is based on the fact that the complexity of an object is determined by the length of its description. If an object is indescribable, then there is no description for it and, accordingly its complexity tends to infinity. When the internal structure of numerical sequences is considered, the presence of internal similar subsequences means that these internal structures could be grouped into separate classes and for each class can be assigned a description, which will reduce the size of the description of the entire sequence. Thus, when considering the internal structure of the generated sequences, it is necessary to build such generators that will produce sequences where the subsequences will have the least degree of similarity.

Proceeding to the consideration of unpredictability, we understand the sequence as unpredictable if in whatever order we choose its elements, knowledge of the values of already selected members does not allow to predict the value of the next element that we intend to choose. Since this work investigates processes in nonlinear dynamic systems, unpredictability is the result of the sensitive dependence on initial conditions of the systems. A sequence is called predictable if there is a computable map for it, which allows to obtain an element of the sequence based on previous values. Thereby, periodic similar subsequences allow to calculate sequence elements for these subsequences with known level of similarity. It is known that any chaotic sequence is unpredictable. However, the question of the coincidence of classes of chaotic and unpredictable sequences remains open.

Considering the internal structure of the sequences obtained on the basis of the above-described maps, the problem of finding and evaluating such structures arises. Presence or absence of which reflects a particular measure of approaching a given sequence to a random one. Accordingly, the map that generates sequences with fewer similar internal subsequences and a smaller length of these subsequences can be considered for further analysis on the possibility of using it as a pseudorandom sequence generator. The problem of choosing a similarity measure for the evaluation of subsequences appears. In this paper, the correlation is chosen as the measure of similarity, as a measure of the strength of association between two subsequences. Subsequences will be called similar if the measure takes values greater than 0.5. Among the various correlation coefficients, we will use the Spearman correlation, since Spearman rank correlation test does not carry any assumptions about the distribution of the data and is the appropriate correlation analysis when the variables are measured on a scale that is at least ordinal. Spearman correlation does not assume linear type of relationship between compared variables. And also it helps to reduce effect of extreme variations in subsequence values. The Spearman ranks correlation coefficient can be calculated using the following equations:

$$r_s = 1 - \frac{6\sum d_i^2}{N(N^2 - 1)}, \qquad (5)$$

where $d_i$ is the difference between ranks for each data pair, and value $N$ is the number of data pairs. Spearman correlation performs analysis based on the ranks of data, thus it can represent the similarity of the shape of two distributions. Spearman correlation calculates the *p*-value the same way as linear regression and correlation, except that it does it on ranks, not measurements. To evaluate the internal structure of the sequences obtained on the basis of the maps, the following method is used and involves following steps:

1. The position of the first peak is calculated in order to remove initial exponential component from consideration;
2. Determine the size of the initial subsequence for evaluation with the following elements of the sequence;
3. Using the single shift, the value of the Spearman correlation of the subsequence from step 2 with proportional subsequences is calculated;
4. The obtained correlation values are filtered according to a given level;
5. The size of the initial subsequence is reduced by 1 as long as it exceeds 10 elements and steps 1-4 are repeated.

Thus, this method allows us to obtain a hierarchy of internal cycles according to the cycle length, as well as the degree of similarity of the found structures. This hierarchy can be used for

further sequence evaluation. According to step-by-step approach to similar subsequence search we obtain set of values that could identify subsequences that are longer than initial evaluation subsequence. Table 2 shows the number of similar overlapping subsequences for the presented maps based on this approach. In this table columns *m(n)* show the period length for a map *n* and columns *pn(n)* show number of overlapping inner patterns for map *n*, where *n* – number of the corresponding map.

Table 2

Number of overlapped similar inner subsequences

| **prime number** | ***Class*** | ***m(1)*** | **pn(1)** | ***m(2)*** | **pn(2)** | ***m(2)*** | **pn(3)** | ***m(2)*** | **pn(4)** |
|---|---|---|---|---|---|---|---|---|---|
| 293 | Low | *146* | 1522 | *219* | *4140* | 292 | *8347* | 146 | *491* |
| 521 | Low | *130* | 1378 | *187* | *3356* | 260 | *6665* | 130 | *488* |
| 26295457 | High | *111* | 504 | *159* | *1858* | 111 | *753* | 111 | *793* |
| 31675363 | High | *181* | 3038 | *262* | *7035* | 362 | *28942* | 181 | *2776* |

## Conclusion

The results obtained in the work show that, chaotic character and property of unpredictability for analyzed maps (1) and (2) demonstrate a very high degree of similarity for many internal cycles. Used correlation test enabled to identify those inner subsequences for further analysis. While the best results in the number of similar internal cycles are shown by the maps (4), which confirms the previously obtained results for this map when statistical tests were used to estimate the measure of randomness. At the same time, the greatest number of similar internal cycles is demonstrated by the map (3), even considering that the sequence length for this mapping is more than the others. For solid pseudorandomness generation methods that generate sequences with fewer similar internal subsequences and a smaller length of these subsequences should be considered.

## References

1. Schuster, H. (1995). Deterministic chaos. Weinheim: VCH. p. 320.
2. Hirsch, M., Smale, S. and Devaney, R. (2013). Differential equations, dynamical systems, and an introduction to chaos. Amsterdam: Academic Press. p. 423.
3. Vostrov, G., Khrinenko, A. (2018). Pseudorandom processes of the number sequence generation. ELTECS.
4. Uspenskiy, V.(2009). Four algorithmic faces of randomness [Chetyre algoritmicheskikh litsa sluchaynosti]. Moscow: MCNMO, p 40.
5. Sharkovsky, A. (1988). Attractors of trajectories and their pools [Attraktory traektorij i ih bassejny]. Kyiv: «Scientific book». p. 322.

## ФОРМУВАННЯ ВНУТРІШНЬОЇ СТРУКТУРИ ПРИ ГЕНЕРАЦІЇ ПСЕВДОВИПАДКОВИХ ПОСЛІДОВНОСТЕЙ

**Г. М. Востров, А. О. Хріненко**
*Одеський національний політехнічний університет*

***Анотація.*** *Дана робота продовжує дослідження і аналіз проблем, що виникають в нелінійних динамічних системах у процесі формування числових псевдовипадкових послідовностей. Моделювання нелінійних процесів відбувається за допомогою групи відображень, що дозволяють отримувати цілочисельні послідовності і проводити аналіз внутрішньої структури цих послідовностей. Розглянуто властивості і внутрішню структуру отриманих послідовностей відповідно до їх впливу на міру випадковості. В роботі проводиться пошук внутрішніх підпослідовностей і, відповідно, аналізуються використані відображення з погляду на кількість подібних внутрішніх циклів в послідовностях. Аналіз внутрішніх циклів відбувається з погляду алгоритмічних властивостей випадковості, таких як хаотичність та непрогнозованність. Хаотичний характер отриманих послідовностей аналізується з позиції концепції істинно випадкових послідовностей в теорії складності Колмогорова, де основна ідея полягає у тому, що складність і хаотичність послідовності визначається довжиною її описання. Відповідно, при аналізі внутрішньої структури наявність подібних підпослідовностей дозволяє скоротити описання всієї послідовності. Таким чином,висувається вимога щодо ступеня подібності будь-яких обраних підпослідовностей при генерації псевдовипадкових послідовностей. Аналізуючи послідовності з точки зору непрогнозованості, розглядається можливість отримання елементів послідовності на основі інформації про попередні елементи цієї послідовності. Відповідно, наявність періодичних подібних підпослідовностей порушує умову непрогнозованості і не дозволяє*

*розглядати таку послідовність як псевдовипадкову. Оскільки розглядається міра подібності внутрішніх підпослідовностей в якості міри подібності для аналізу послідовностей було обрано коефіцієнт кореляції Спірмена, оскільки дана кореляційна міра не містить припущень щодо закону розподілу чисел у послідовності, а також дозволяє зменшити ефект надмірних відхилень серед числових значень у послідовності. Розглянутий метод оцінки послідовностей дозволяє отримати ієрархію внутрішніх циклів відповідно до їх довжини, а також міри подібності між ними. Результати отримані в ході роботи дозволяють оцінити кожне з використаних відображень та зробити висновок, що для отримання надійних псевдовипадкових послідовностей необхідно будувати методи, що дозволять отримувати найменшу кількість внутрішніх подібних підпослідовностей і, відповідно, найменшою мірою їхньої подібності.*

***Ключові слова:*** *хаос, псевдовипадковість, нелінійні відображення, внутрішня структура.*

## ФОРМИРОВАНИЕ ВНУТРЕННЕЙ СТРУКТУРЫ ПРИ ГЕНЕРАЦИИ ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

**Г. Н. Востров, А. О. Хриненко**
*Одесский национальный политехнический университет*

***Аннотация.*** *Данная работа анализирует и рассматривает проблемы, которые возникают в процессах числовой генерации псевдослучайных последовательностей. Свойства и внутренняя структура полученных последовательностей рассматривается относительно влияния внутренней структуры на меру случайности данных последовательностей.*

***Ключевые слова:*** *хаос, псевдослучайность, нелинейные отображения, внутренняя структура.*

Received on 26.11.2018

**George Vostrov**, Ph. D. of Technical Sciences, Associate Professor of the Department of Applied Mathematics and Information Technologies, Odessa National Polytechnic University. Shevchenko ave., 1, Odessa, Ukraine. E-mail: vostrov@gmail.com, тел. +380503168776

**Востров Георгій Миколайович**, кандидат технічних наук, доцент кафедри прикладної математики та інформаційних технологій Одеського національного політехнічного університету. Проспект Шевченко, 1, Одеса, Україна.
E-mail: vostrov@gmail.com, тел. +380503168776

**ORCID ID**: 0000-0003-3856-5392

**Khrinenko Andrii**, master of the Department of Applied Mathematics and Information Technologies, Odessa National Polytechnic University. Shevchenko ave., 1, Odessa, Ukraine. E-mail: khrinenko.andrew@gmail.com, тел. +380637515228

**Хріненко Андрій Олегович**, магістр кафедри прикладної математики та інформаційних технологій Одеського національного політехнічного університету. Проспект Шевченко, 1, Одеса, Україна.
E-mail: khrinenko.andrew@gmail.com, тел. +380637515228

**ORCID ID**: 0000-0001-6000-2102