On recognition of the shift register output function with a random input for a Markov model of an input sequence

Sergey Yu. Melnikov, Konstantin E. Samouylov

Peoples' Friendship University of Russia (RUDN University), 6, Miklukho-Maklaya St., Moscow, 117198, Russia

Abstract

The problem of recognizing the output function of a binary non-autonomous shift register is considered, the input of which receives a sequence of random variables connected in a simple homogeneous stationary Markov chain. The statistics of symbol frequencies in the output register sequence is used. The recognition problem is reduced to the form of integer minimization of the linear form module under linear constraints. The results are compared with the case when the input sequence is Bernoulli.

Keywords

shift register, random input machine, recognition of output functions

CEUR-WS.org/Vol-2639/paper-09.pdf

1. Introduction and background

In [1], the problem of recognizing the output function of a Moore automaton by the output sequence symbol statistics when the input is a sequence of independent identically distributed random variables was considered. Such a problem is a special case of the general problem of recognition of machines with random input and arises in a number of theoretical and applied problems. We list some of them. The problem of choosing the minimum set of multigrams in the output sequence, the probabilities of which characterize the automaton, leads to the problem of analyzing the algebraic properties of the probability distributions family of such multigrams as functions of the probability distribution at the input [2, 3].

In [4], a review of papers on a similar problem is given: the probabilities of multigrams of some stationary discrete random process are known, and it is required to determine whether this process is a function on a Markov chain. In [5], for a non-autonomous shift register with a random input, an estimate was obtained of the output n-grams size, the set of probabilities of which determines the equivalence class of such an automaton, and estimates of the number of equivalence classes.

We can represent any probabilistic automaton [6, 7] in the form of a sequential connection of a "source of probability" that randomly and independently generates the symbols of an

Workshop on information technology and scientific computing in the framework of the X International Conference Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems (ITTMM-2020), Moscow, Russian, April 13-17, 2020

melnikov@linfotech.ru (S. Yu. Melnikov); samuylov-ke@rudn.ru (K. E. Samouylov)

0000-0002-6368-9680 (K. E. Samouylov)

© 2020 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

alphabet with given probabilities and a deterministic finite automaton. In a number of works (see the review [8]), the problem of synthesizing such a "source of probability" was considered, in particular, developing criteria to determine whether the desired value is generated by a combination of Bernoulli random variables with probabilities from a given set.

The search for new calculation methods leads to the problem of analyzing a class of functions that describe the probabilities of multigrams in the output sequence as functions of real variables [9].

A large number of works are devoted to the problem of discrete devices testing by applying random sequences (tests) to them and analyzing the statistics at the outputs [10, 11], etc.). The problem of determining the malfunction of a device in such a formulation is the task of recognizing automata [12]. This direction is called random or pseudorandom testing. As noted in [13], the relevance of this direction is associated with the high volume of computations necessary to construct deterministic tests for complex discrete devices. A special place is taken by probabilistic compact testing methods, in which malfunctions are detected by comparing some statistical characteristics of the tested and reference device. Such a characteristic may be the frequency of occurrence in the output sequence of a fixed segment of the output symbols. The ability to change the analyzed segment and the parameters of the random sequence generator at the input of the device is a powerful tool in the hands of the researcher. In [Hamlet, 2006], situations are described where the use of deterministic tests is impractical and random tests should be used. The questions of estimating the length of random tests and the probability of detection of malfunctions are considered in [14] and [15]. We note the possibility of using probabilistic testing to detect covert channels in the analyzed equipment [16, 17].

In [18] and [19], the problem of diagnosing a discrete device with random input under conditions when the output sequence of the device is known with distortions is considered.

In [20], a binary shift register with a random Bernoulli input is considered, the output function of which is known up to a certain transformation. The output sequence of the register is specified, and the input sequence is known in the variants. A statistical criterion is built to verify that the output sequence could be obtained by this input variant. The criterion is based on counting the frequencies of certain bigrams in the output sequence.

The problem considered in [1] can be formulated as follows. What information about the output function of the automaton whose input is a sequence of independent random variables can we obtain if the parameters of the "input" distribution and the relative frequency of occurrence of the symbol in the output sequence are known? How will this information change if the parameters of the "input" distribution are changed? What output functions are indistinguishable?

Similar problems can be considered in a more general form, assuming a simple or complex [21] Markov dependence in the input sequence. The main ideas of the developed approach, in principle, can be transferred from the "independent" to the "Markov" case. In fact, a sequence of random variables connected into a finite Markov chain can be considered as a sequence of states of a suitable finite automaton (the transition graph of which is the transition graph of the Markov chain under consideration) that processes a sequence of independent random variables with a properly selected distribution [22]. Let the output function of such an automaton be identical (output = state). Then the machine that processes the Markov chain can be represented as a serial connection of two machines, the input of the first of which receives a sequence of independent random variables. Thus, the problem reduces to the one considered above, only for a more complex automaton (Figure 1).

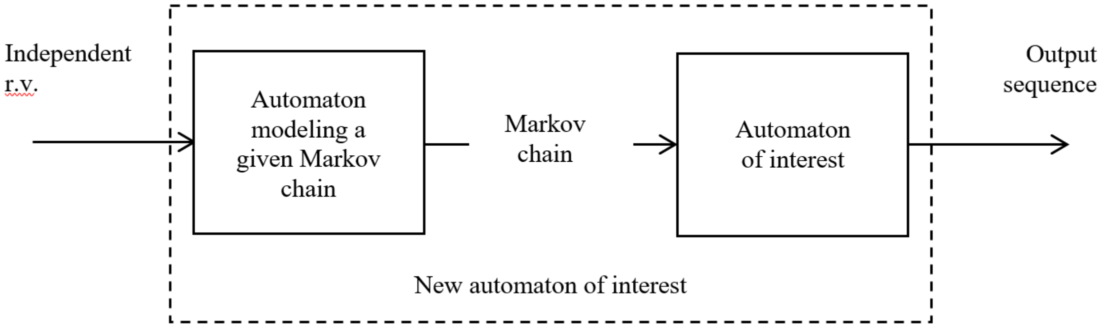
Independent r.v.
Automaton modeling a given Markov chain
Markov chain
Automaton of interest
Output sequence
New automaton of interest

Figure 1: New automaton design

The same can be done if the input sequence is a more general random process — a function on a finite Markov chain. In this case, the output function of the first automaton may not be identical. Certain difficulties in this direction are associated both with an increase in the number of states of a new combined automaton, and therefore with an increase in the dimension of computational problems, and with the requirement of strong connectivity of the serial connection of two automata, the first of which should provide all possible models of randomness in the input sequence.

For the case of a binary shift register, we will use the direct representation of probability functions as functions of the Markov chain parameters. A Markov chain with two states is determined by the transition probability matrix and the initial probability distribution, which we assume to be stationary. The transition probability matrix has a size of 2×2 , and due tostochasticity it is determined by two real parameters λ and ξ, $0 < \lambda, \xi < 1$. These parameters can be interpreted as the probabilities of transitions "from state 0 to state 1" and "from state 1 tostate 0", respectively. In [23], for the case when the symbols of the input sequence are connected into a simple homogeneous stationary Markov chain, an expression is obtained for the probability of the symbol "1" in the output sequence of the register. The same work describes the equivalence relation on a set of Boolean functions that occurs when registers with these output functions have the same probability functions. Here we consider the problem of determining the equivalence class of the output function from symbol statistics, reduce it to the form of the discrete optimization problem, and compare the Bernoulli and Markov cases in terms of obtaining information for recognizing the output function and dimensions of discrete optimization problems.

2. Statement of the problem

Let V_n be the space of n-dimensional binary vectors, F_n be the set of Boolean functions of n arguments, $n = 1, 2, \dots$. For $f(x_1, x_2, \dots, x_n) \in F_n$ by $A_f = (X = \{0,1\}, V_n, Y = \{0,1\}, h, f)$ we denote the Moore automaton with set of states V_n, the transition function h, determined by the rule $h((a_1, \dots, a_n), x) = (a_2, \dots, a_n, x)$, where $x, a_i \in \{0, 1\}$, $i = 1, 2, \dots, n$, and output function $f(x_1, x_2, \dots, x_n)$. The automaton A_f is a shift register with a size of n.

Suppose that the automaton A_f, $f \in F_n$, $n = 1, 2, \dots$ receives a sequence of binary random variables $x^{(i)}$, $i = 1, 2, \dots$, connected in a simple homogeneous stationary Markov chain with the transition probability matrix

$$\begin{pmatrix} 1-\lambda & \lambda \\ \xi & 1-\xi \end{pmatrix}, \quad 0 < \lambda, \xi < 1. \tag{1}$$

Our task is to determine the accuracy with which the output function $f(x_1, x_2, \dots, x_n)$ can be recognized by the relative frequency of the symbol "1" in the output sequence of automaton A_f and propose a recognition method.

3. Recognition of the equivalence class by the value of the probability function

In [23], the definition of the Markov weight structure of a Boolean function $f \in F_n$ was introduced as an integer vector $\overline{m}(f)$ consisting of sums of values $f(x_1, x_2, \dots, x_n)$ on special subsets V_n formed by vectors with the same bigram markings.

The vector $\overline{m}(f)$ of the Markov weight structure defines the class of statistical equivalence of the function f. It is easy to see that the dimension $d(n)$ of the vector $\overline{m}(f)$ grows quadratically with n. Direct calculation of the number of different bigram markings of vectors from V_n leads to the following result.

Theorem 1. If $n \geqslant 2$:

$$d(n) = \begin{cases} \frac{3}{4}n^2 - n + 2, & \text{if } n \text{ is even,} \\ \frac{3}{4}(n-1)^2 - n + 2, & \text{if } n \text{ is odd.} \end{cases} \tag{2}$$

Consider the set $A = \{\overline{m}(f), f \in F_n\}$ of vectors formed by the vectors of all possible Markov weight structures as the set of integer points in the $d(n)$-dimensional Euclidean space $R^{d(n)}$.

Using the result of the Lemma of [23], it is easy to see that $A = A_1 \times A_2 \times A_3$, where

$$\begin{aligned} A_1 &= \left\{ \left(a_{ij}^{(1)}, (i,j) \in I_1 \right) \middle| 0 \le a_{ij}^{(1)} \le \binom{n-j-1}{i} \binom{j-1}{i-1}, \quad a_{ij}^{(1)} - \text{integer} \right\}, \\ A_2 &= \left\{ \left(a_{ij}^{(2)}, (i,j) \in I_2 \right) \middle| 0 \le a_{ij}^{(2)} \le 2 \binom{n-j-1}{i-1} \binom{j-1}{i-1}, \quad a_{ij}^{(2)} - \text{integer} \right\}, \\ A_3 &= \left\{ \left(a_{ij}^{(3)}, (i,j) \in I_3 \right) \middle| 0 \le a_{ij}^{(3)} \le \binom{n-j-1}{i-1} \binom{j-1}{i}, \quad a_{ij}^{(3)} - \text{integer} \right\}. \end{aligned} \tag{3}$$

Thus, A is the set of integer points of a $d(n)$-dimensional parallelepiped. Let M_n be a family of real functions defined on a square $0 < \lambda$, $\xi < 1$ of the form $\overline{R}(\lambda, \xi) \left(\overline{a} - \overline{b} \right)^T$, where $\overline{a}, \overline{b} \in A$, $\overline{a} \neq \overline{b}$, the vector $\overline{R}(\lambda, \xi)$ is defined in [23], the coordinates of this vector are fractional rational functions of λ and ξ. Let Ω_n be the set of all those points of the square at which at least one function in the family M_n is equal to zero. It is easy to see that Ω_n is the union of a finite number of smooth curves in a unit square. In particular, it can be shown that both diagonals of the square, $\{(\lambda, \xi) | \lambda + \xi = 1\}$ and $\{(\lambda, \xi) | \lambda = \xi\}$ belong to the set Ω_n, if $n \geqslant 3$. The set Ω_n has Lebesgue measure zero.

Theorem 2. For $(\lambda, \xi) \in \{(0,1) \times (0,1)\} \setminus \Omega_n$, by the value of $P_f(\lambda, \xi)$, we can unambiguously indicate the class of statistical equivalence of the function f. This equivalence class corresponds to the vector of the Markov weight structure $\overline{\mu_0}$, which is the only solution of the equation

$$P_f(\lambda, \xi) = \overline{R}(\lambda, \xi)\,(\overline{\mu_0})^T \text{ under restriction } \overline{\mu_0} \in A. \tag{4}$$

If $(\lambda, \xi) \in \Omega_n$, then this equation has more than one solution, and the vector of the Markov weight structure of f is contained among these solutions.

4. Recognition of the equivalence class using the symbols statistics

Let $y^{(i)} = f\left(x^{(i)}, x^{(i+1)}, \dots, x^{(i+n-1)}\right)$, $i = 1, 2, \dots$ be the output sequence of the automaton A_f in the scheme described above,

$$Y_N = \frac{1}{N} \sum_{i=1}^{N} y^{(i)}, \qquad N = 1, 2, \dots. \tag{5}$$

The binary sequence $y^{(i)}$ is stationary, therefore, the law of large numbers is valid [24]:

$$Y_N \to P_f(\lambda, \xi) \quad \text{(probability convergence).} \tag{6}$$

Theorem 3. Let $(\lambda, \xi) \in \{(0,1) \times (0,1)\} \setminus \Omega_n$. The probability that the minimization problem

$$\begin{cases} \left| Y_N - \overline{R}(\lambda, \xi) \overline{\mu}^T \right| \to \min \\ \overline{\mu} \in A \end{cases} \tag{7}$$

has a unique solution, tends to 1 with $N \to \infty$.

This unique solution corresponds the class of statistical equivalence of output function f. For the case $(\lambda, \xi) \in \Omega_n$, we can only say that the vector of the weighted Markov structure of the true output function f is contained among the solutions of problem (7).

5. Conclusion

Comparing the results obtained for the case when the input sequence is a Markov chain with the results obtained for the case of a Bernoulli sequence, the following can be noted.

1. Probabilistic functions in both cases are determined by the weights (sums of values) of the Boolean function of the outputs on special subsets of space V_n. These subsets consist of vectors that have the same probability in a particular model (in the independent case, vectors of the same weight, in Markov, vectors with the same frequency of occurrence of

bigrams). In the first case, the probability function is a polynomial in the parameter p of the Bernoulli scheme; in the second, it is a fractional rational function in the parameters λ and ξ, that determine the Markov chain.

2. The relations of statistical equivalence on F_n are introduced on the basis of the identity of probability functions; from the statistical equivalence of two Boolean functions with a Markov input dependence follows their statistical equivalence with a Bernoulli input. The equivalence of two functions is equivalent to the equality of their weights on the mentioned subsets V_n. In the independent case the number of such subsets is $n + 1$, in the Markov case the number of such subsets is $\frac{3}{4}n^2O(n)$. The number of equivalence classes for the Bernoulli case is equal to $\exp\left(\frac{n^2}{2} + O(n \ln n)\right)$, for the Markov case is equal to $\exp\left(\frac{5n^3}{4} \ln n + O(n^3)\right)$.
3. In the set of all kinds of distributions (the parameter p of the Bernoulli distribution belongs to the interval $(0, 1)$, the parameters (λ, ξ) of the transition matrix of the Markov chain belong to the square $(0, 1)\times(0, 1)$), a subset Ω of the “special” distributions is distinguished. For all distributions not from the set Ω, the value of the probability function allows us to uniquely indicate the equivalence class of the output function. If the distribution belongs Ω, then there are at least two nonequivalent output functions for which the probability values of “1” in the output sequence coincide. The Lebesgue measure of the set Ω of “special distributions” in the distribution space is zero: in the first case it turns out to be a finite set of points of the segment, in the second – the set of points of a finite number of smooth curves inside a square.
4. The problem of determining the equivalence class of a function from the value of a probability function reduces to the problem of solving the equation in integers under linear constraints; the problem of determining the equivalence class of a function from the statistic value of a sample average output sequence is reduced to the problem of minimizing a linear form module under linear constraints. The number of unknowns in the Bernoulli case is equal $n + 1$, in the case of Markov dependence it is approximately equal $\frac{3}{4}n^2 - n$.

Thus, analyzing the problem of recognizing an unknown function of outputs by symbol statistics, we can say that the proposed approach, while complicating the probabilistic nature of the input sequence (switching from the Bernoulli scheme to the Markov chain), leads, on the one hand, to the possibility of obtaining more detailed information about the recognizable function; on the other hand, the dimension of discrete optimization problems increases.

Acknowledgments

The publication has been prepared with the support of the “RUDN University Program 5-100”. The reported study was funded by RFBR, project numbers 18-00-01555 (18-00-01685), 19-07-00933.

References

[1] S. Y. Melnikov, K. E. Samouylov, The recognition of the output function of a finite automaton with random input, in: V. M. Vishnevskiy, D. V. Kozyrev (Eds.), Distributed Computer and Communication Networks, volume 919, Springer International Publishing, Cham, 2018, pp. 525–531. doi:10.1007/978-3-319-99447-5_45.

[2] A. S. Barashko, O range i statisticheskom otobrazhenii sil'nosvyaznogo avtomata, Kibernetika (1987) 56–60.

[3] H. Ito, S. Amari, K. Kobayashi, Identifiability of hidden markov information sources and their minimum degrees of freedom, IEEE Transactions on Information Theory 38 (1992) 324–333. doi:10.1109/18.119690.

[4] M. Vidyasagar, The complete realization problem for hidden markov models: a survey and some new results, Math. Control Signals Syst. (2011) 1–65. doi:10.1007/s00498-011-0066-7.

[5] M. I. Rozhkov, Algoritmicheskie voprosy identifikacii konechnyh avtomatov po raspredeleniyu vyhodnyh m-gramm, dis. ... d.t.n., 05.13.19 - metody i sistemy zashchity informacii. informacionnaya bezopasnost, Moskva, MGIEM, 2012.

[6] R. G. Buharaev, Osnovy teorii veroyatnostnyh avtomatov, Nauka, Moskva, 1985.

[7] A. Paz, Introduction to probabilistic automata, Academic Press, London, 1971.

[8] R. M. Kolpakov, Diskretnye preobrazovaniya veroyatnostnyh raspredelenij, in: Sovremennye problemy matematiki i mekhaniki, volume III. Matematika, Izd-vo Moskovskogo universiteta, Moskva, 2009, pp. 35–50.

[9] A. V. Ryabinin, Stohasticheskie funkcii konechnyh avtomatov, in: Algebra, logika i teoriya chisel, 7 temat. konf. mekh.-mat. fak. MGU, Moskva, 1986, pp. 77–80.

[10] D. Y. Golembiovskij, Diagnostika cifrovyh ustrojstv. Mikroprogrammnoe i veroyatnostnoe testirovanie, Politekhn. in-t, Saratov, 1993.

[11] C. N. Hadjicostis, Probabilistic detection of fsm single state-transition faults based on state occupancy measurements, IEEE Transactions on Automatic Control 50 (2005) 2078–2083. doi:10.1109/TAC.2005.860270.

[12] N. G. Kushik, Metody sinteza ustanovochnyh i razlichayushchih eksperimentov s nedeterminirovannymi avtomatami, dis. ... k.f.-m.n., 05.13.01 - sistemnyj analiz, upravlenie, obrabotka informacii, Tomsk, 2013.

[13] A. S. Barashko, Y. A. Skobcov, D. V. Speranskij, Modelirovanie i testirovanie diskretnyh ustrojstv, Naukova dumka, Kiev, 1992.

[14] A. S. Barashko, N. V. Cherevko, Nekotorye sposoby ocenki dliny sluchajnogo testa, in: Teoriya i modelirovanie upravlyayushchih sistem, Naukova dumka, Kiev, 1989, pp. 10–15.

[15] G. V. Petruhnova, Ocenka dliny sluchajnoj posledovatel'nosti dlya operacii vosproizvedeniya informacii v kontrol'nom ispytanii, in: O. Y. Kravca (Ed.), Sovremennye problemy informatizacii v sistemah modelirovaniya, programmirovaniya i telekommunikaciyah, volume 9, Izd. "Nauchnaya kniga", Voronezh, 2004, pp. 303–305.

[16] E. E. Timonina, Analiz ugroz skrytyh kanalov i metody postroeniya garantirovanno zashchishchennyh raspredelennyh avtomatizirovannyh sistem, dis. ... d.t.n. 05.13.17 - teoreticheskie osnovy informatiki, Moskva, RGGU, 2004.

[17] A. A. Grusho, N. A. Grusho, E. E. Timonina, Vklyuchenie novyh zapretov v sluchajnye posledovatel'nosti, Inform. i ee primen. 8 (2014) 46–52.

[18] E. Athanasopoulou, C. N. Hadjicostis, Probabilistic approaches to fault detection in networked discrete event systems, IEEE Transactions on Neural Networks 16 (2005) 1042–1052. doi:10.1109/TNN.2005.853430.

[19] E. Athanasopoulou, L. Li, C. N. Hadjicostis, Probabilistic failure diagnosis in finite state machines under unreliable observations, in: in Proceedings of the 8th International Workshop on Discrete Event Systems - WODES'06, Ann Arbor, MI., 2006, pp. 301–306. doi:10.1109/WODES.2006.1678446.

[20] B. A. Sevastyanov, The conditional distribution of the output of an automaton without memory for given characteristics of the input, Discrete Mathematics and Applications 4 (1994) 1–6. doi:10.1515/dma.1994.4.1.1.

[21] A. S. Barashko, O raspoznavanii avtomatov s ispol'zovaniem generatorov zavisimyh sluchajnyh signalov, in: Naukovi praci Donec'kogo nacional'nogo tekhnichnogo universitetu, Problemi modelyuvannya ta avtomatizacii proettuvannya (MAP-2002), DonNTU, Donec'k, 2002, pp. 61–71.

[22] A. S. Davis, Markov chains as random input automata, The American Mathematical Monthly 68 (1961) 264–267.

[23] S. Y. Mel'nikov, K. E. Samujlov, Veroyatnostnye funkcii i statisticheskaya ekvivalentnost' dvoichnyh registrov sdviga so sluchajnym markovskim vhodom, in: Informacionno-telekommunikacionnye tekhnologii i matematicheskoe modelirovanie vysokotekhnologichnyh sistem : materialy Vserossijskoj konferencii s mezhdunarodnym uchastiem. Moskva, RUDN, 13–17 aprelya 2020 g., 2020, pp. 49–54.

[24] V. S. Korolyuk, N. I. Portenko, A. V. Skorohod, A. F. Turbin, Spravochnik po teorii veroyatnostej i matematicheskoj statistike, Nauka, Moskva, 1985.