Identification of automata Markov models using a modified "forward-backward" algorithm

A R Nurutdinova
The Kazan branch of the Federal State Budget-Funded Educational Institution of Higher Education «The Russian State University of Justice», Kazan 420088, Russia

E-mail: nurutdinovaar@mail.ru

Abstract. In 1989, Rabiner L.R proposed the solution of the problem of identifying hidden Markov models by induction using the modified "forward-backward" algorithm. In this paper, we propose the application of a modification of this method for solving the recognition of automaton Markov models following sequences of finite simple homogeneous Markov chains generated on the basis of stochastic ergodic matrices, which in turn determine the class of the automata Markov model.

1. Introduction
Markov models are used to modeling complex systems and processes in various fields [1-4]. A separate direction of research in the theory of Markov processes comprises the problem of analysis (recognition, control and diagnostics) [5-8]. In [9-17] it is possible to find the results of studies of the analysis of entropy and asymptotic properties of discrete Markov processes. In [8, 9, 17] methods of classification and identification of Markov chains (MC), which arise when solving problems of recognition of automaton Markov models (AMM), are proposed. In particular, they are based on the calculation of functionals following sequences of finite length, taken from the output of AMM, with some error with respect to the ergodic stochastic matrix (ESM) that defines it. This reduces the accuracy of the analysis on the basis of these functionals with limitations on the length of the observed sequence, especially for CM of length of the order of $10^2 - 10^3$.

The problems of identification of automata Markov models were considered in [18, 19], and identification models based on the calculation of functionals directly from the CM realization taking into account the structure of ESM defining AMM were proposed to solve the above-described problem. According to the results, the methods obtained made it possible to increase the informativeness of the solution of the AMM recognition problem, to identify with a higher confidence probability for a smaller number of generated MC elements.

In this article we propose an approach to the problem of identifying automata Markov models defined on the basis of ergodic stochastic matrices on the basis of the Markov chain realizations generated by them, which is a modification of the "forward-backward" algorithm proposed for solving the speech recognition problem for the hidden Markov model in [3]. The identification model is considered for two classes of automata Markov models, determined on the basis of regular and cyclic ESM.
2. Modification of the «forward-backward» algorithm

The automata Markov models (AMM) without an exit is given as \( (S, \varphi(s'/s)) \), where \( S = \{s_i\} \), \( i = 0, n - 1 \) is a set of MC states, \( s, s' \in S \), \( \varphi(s', s) \) - transition function, defined by the stochastic matrix \( P_i = \{p_{ij}\} \) dimension \( n \times n \), \( i, j = 0, n - 1 \) [10]. If you set different functions \( \varphi(s', s) \) for the AMM, you can receive different AMM classes defined by ESM \( P \), \( P \in Q \), and solve the AMM identification problem for the sequence of MC that are derived from the specified AMM subclasses.

Let for AMM(\( P \in Q \)) we get a set \( S \) sequence of MC of length \( N \) as \( \hat{S}(N) = s_1, s_2, ..., s_N \), where \( s_i \) is state AMM at time \( t \), \( t = 1, N \). Let there be given a set of subclasses ESM \( P - Q \), AMM (\( P \in Q \)), which are defined depending on the structure ESM \( P \).

The AMM(\( P \)) identification task is defined - to determine the value of \( P(\hat{S}(N) | AMM(P)) \) - the probability that \( \hat{S}(N) \) is generated on the basis of AMM(\( P \)), where ESM \( P \) belongs to a given subclass \( Q \).

Suppose that \( \hat{S}_k(N) \) is MC of length \( N \), type similar to \( \hat{S}(N) \), for which a \( k \) moments of time, \( k < N \) which station \( s(t) \) are hidden from observation. Within proposed solution method the tasks are solved both for \( \hat{S}(N) \), and for \( \hat{S}_k(N) \).

In order to identify the sequence \( \hat{S}(N) = s(1)s(2)...s(N) \) according to the "forward-backward" algorithm the following array of variables are input [20]:

\[
\alpha_s(i) = P(s(1)s(2)...s(t), s(t) = s_i | AMM(P)), t = 1, N, i = 1, m.
\]

The algorithm includes three steps [3].

Step 1. Initialization of variables: \( \alpha_s(i) = \pi_y(i) \cdot z_i, i = 1, n, z_i = \begin{cases} 1: & s(t + 1) = s_j; \\ 0: & \text{otherwise} \end{cases} \).

Step 2. Induction: \( \alpha_s(j) = \sum_{i=1}^{m} \alpha_s(i) \cdot p_{ij} \cdot z_j, t = 1, N - 1, j = n \).

Step 3. Calculate \( P(\hat{S}(N) | AMM(P)) = \alpha_y(s(N)) \).

For the case when not all states of the sequence MC are observable, when implementing step 2 of the identification algorithm, there is a logical expression:

\[
\alpha_s(j) = \sum_{i=1}^{m} \alpha_s(i) \cdot p_{ij} \cdot z'_j, z'_j = \begin{cases} 1: & s(t + 1) \text{ is hidden} \\ z'_j: & \text{otherwise} \end{cases}.
\]

If there are \( k \) hidden elements in the observed sequence \( s(N) \), then the probability is

\[
P(\hat{S}_k(N) | AMM(P)) = \sum_{i=1}^{m} \alpha_y(i).
\]

If all elements of \( \hat{S}(N) \) are observable, the computational complexity of the proposed method for solving the problem of identifying AMM is of the order \( O(N \cdot n) \). The presence of hidden elements in an amount comparable to the length of the sequence \( N \), increases the order of the computational complexity of the algorithm to \( O(N \cdot n^2) \).

In the studies [18-20] we estimations of the complexity of the algorithms for identifying AMM. In particular, it was shown that using functionals based on l-gram, \( l = 2,3 \), the order of the computational complexity of the identification algorithm is \( O(N \cdot n^2) \) for \( O(N \cdot n^2) \), therefore, it is advisable to use these methods for small values of \( n \). For large values of \( n \), it is more efficient to use an algorithm
based on frequency characteristics, whose computational complexity has order $O(N \cdot n)$ or the proposed "forward-backward" algorithm. The proposed model allows us to quantify the probability of identifying the sequence of the Markov chain for the possibility of generating a given AMM.

We consider the problem of identifying AMM based on the ESM of the cyclic matrix [16], where the class $SC_r$ of AMM studied is determined by the period $r$ of the cyclic MC.

Let’s consider the set of the maximum entropy CSM, \([21]: P \in P_n (CSMr)\), where \(P_n (CSMr)\) is family CSM of dimension \(n \times n\), obtained by partition of a set \(n\) states of AMM into \(r\) nonempty subsets, \([P_n (1 | CM_r)] = S(n, r)\).

The probability of identifying such classes of AMM by the "forward-backward" method is determined in accordance with the following algorithm [22]:

Step 1. Initialization of variables:

$$\alpha'_1 (i) = \begin{cases} 1: \ (\pi_0 (i) > 0) \ & (z_i = 1), \ i = 1, m, \ z_i = 1: \ s(t + 1) = s_j, \ t = 0. \\ 0: \ otherwise \end{cases}$$

Step 2. Induction:

$$\alpha'_1 (j) = \begin{cases} 1: \ (\exists \alpha'_1 (i) \cdot p_j > 0) \ & (z_j = 1), \ t = 1, N - 1, \ j = 1, m. \\ 0: \ otherwise \end{cases}$$

Step 3. If $\alpha'_1 (s(N)) > 0$, then $s(N) \in SC_r$ and probability of identifying the $S(N)$ membership to a given subclass $SC_r$ is determined according to the formula:

$$P(S(N)|AMM(P)) = 1 - \left(\frac{n-1}{n}D\right)^N.$$  

If the step 2 of algorithm results in $\sum_{j=1}^{m} \alpha'_1 (j) = 0$, then $s(N) \notin SC_r$.

In order to identify the sequence, which elements are partially hidden from observation, when implementing step 2 of the identification algorithm, there is a logical expression:

$$\alpha'_{i+1} (j) = \begin{cases} 1: \ (\exists \alpha_{i+1} (j) \cdot p_j > 0) \ & (z'_j = 1), \ s(t + 1) - hidden \\ 0: \ otherwise \end{cases}$$

Besides, if $s(N)$ is hidden from observation, then:

$$\alpha'_N (\hat{S}_N (N)) = \begin{cases} 1: \ \exists \alpha_N (i) > 0 \\ 0: \ otherwise \end{cases}.$$

In papers [22-23] methods were proposed with the help of functionals formed on the basis of the characteristic feature of a cyclic stochastic matrix. For this methods computational complexity of algorithms for obtaining the probability of identification of cyclic MC is equal to $O(r \cdot n \cdot N \cdot S(n, r))$.

For the modified «forward-backward» algorithm computational complexity with the search of values $P(\hat{S}(N) | AMM(P))$, $P \in P_n (CSMr)$, is equal to $O(n \cdot N \cdot S(n, r))$ and is less dependent on the index of CSM. If $s(N)$ is hidden from observation, then computational complexity increases by $n$ comparison operations.

Thus, the proposed modified "forward-backward" algorithm is quite effective in estimating computational complexity. In addition, the method allows solving the recognition problem for sequences with hidden elements.

**Acknowledgments**

This work was supported by RFBR Grant 18-01-00120a «Specialized devices for generating and processing data sets in the architecture of programmable logic devices class FPGA». 


References
[1] Levin B R 1985 Probabilistic models and methods in communication and management systems. (Moscow, Radio i sviaz’) 312 [In Russian]
[2] Friedman W F 1985 Military cryptanalisis Part I. Vol. Z. (Aegean Park Press. Laguna Hills CA) 342
[3] Lawrence R and Rabiner A 1989 Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition Proc. IEEE. 1989 77 (2) 257–286
[4] Teptin G M and Ivanov K V 2008 Markov models of Defense Means for Automated Special Purpose Systems Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-matematicheskie Nauki 150 (4) 41–53
[5] Raskin L G 1975 Analysis of stochastic systems and elements of the optimal control theory (Moscow, Sov. Radio) 344 [In Russian]
[6] Pospelov D A 1970 Probabilistic automata (Moscow, Energy) 88 [In Russian]
[7] Bukharaev R G 1967 On the estimation of the minimal number of input values of an automaton that generates a homogeneous Markov chain Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-matematicheskie Nauki 129 (4) 3–11 [In Russian]
[8] Bukharaev R G 1967 The representability of events in probabilistic automata Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-matematicheskie Nauki 127 (3) 7–20 [In Russian]
[9] Alpin Yu A and Alpina V S 2008 The Perron-Flobenius Theorem – a proof with the use of Markov chain. Uchenye Zapiski Kazanskogo Universiteta. Seriya Fiziko-matematicheskie Nauki 150 (4) 41–53 [In Russian]
[10] Bukharaev R G 1970 Probabilistic automata (Kazan, Izd. KGU)188 [In Russian]
[11] Zakharov V M, Nurmeev N N, Salimov F I and Shalagin S V 2001 Analysis of stochastic matrices using multivariate classification methods. Diskretnaja matematika i ee prilozhenija. Materialy Mezdunar. nauch. sem. 2 (3) 156–159 [In Russian]
[12] Zakharov V M, Nurmeev N N, Salimov F I and Shalagin S V 2000 Classification of ergodic stochastic matrices using methods of cluster and discriminant analysis Issledovanija po informatike 2 91–106 [In Russian]
[13] Lorenc A A 1976 Reliability and speed of probabilistic automata (Riga, Zinatne) 112
[14] Romanovskiy V I 1949 Discrete Markov chains (Moscow, Gostekhizdat) 436 [In Russian]
[15] Fedotov N G 1990 Methods of stochastic geometry in pattern recognition (Moscow, Radio i sviaz’) 144 [In Russian]
[16] Kemeny J G and Snell J L 1976 Finite Markov Chains (NY, Springer Publishing Company) 272
[17] Li T, Jadh D and Zelner A 1977 Evaluation of parameters of Markov property models by aggregated temporary rows (Moscow, Statistics) 221 [In Russian]
[18] Nurutdinova A R and Shalagin S V 2010 The method of identification of automaton Markov models on the basis of sequences generated by them Vestnik KGTU im. A.N. Tupoleva 1 94–99 [In Russian]
[19] Nurutdinova A R and Shalagin S V 2010 Multiparameter classification of automation Markov models based on their generated sequence of states Applied discrete mathematics 4 41–54 [In Russian]
[20] Shalagin S V and Nurutdinova A R 2016 Identification of Markovian Automata Sub-classes International Journal of Pharmacy and Technology 8 (3) 15327–15337
[21] Khinchin A Ya 1953 Concept of entropy in the theory of probability Success of Mathematical Sciences 3 3–20 [In Russian]
[22] Shalagin S V and Nurutdinova A R 2016 Identification Algorithms of Simple Homogeneous Markov Chains of Cyclic Class and Their Complexity Analysis International Journal of Pharmacy and Technology 8 (3) 14926–14935
[23] Shalagin S V and Nurutdinova A R 2017 Identification of a sequence of measurements of economic parameters based on a hidden Markov model Problems of analysis and modeling of regional socio-economic processes 159–162 [In Russian]