О РАСПОЗНАВАНИИ АВТОМАТОВ С ИСПОЛЬЗОВАНИЕМ ГЕНЕРАТОРОВ ЗАВИСИМЫХ СЛУЧАЙНЫХ СИГНАЛОВ

Барашко А.С.
Кафедра ПМИ ДонНТУ

Abstract

Barashko A.S. On automata recognation with use the generators of dependent random signals. *Concept of generalized statistical equivalence of automata is entered and it is shown that it is lower than usual equivalence, but it is stronger than statistical equivalence of automata.*

ВВЕДЕНИЕ

Статья посвящена изучению статистических свойств детерминированных автоматов и продолжает цикл работ [1 - 3]. В работе [1] исследуется статистический метод распознавания сильно связных автоматов, основанный на подсчете частоты появления фиксированной подпоследовательности (фрагмента) в выходной последовательности автомата, на входе которого действует стационарный генератор независимых случайных сигналов. Основной результат этой работы заключается в доказательстве того, что распределение частоты появления фрагмента аппроксимируется нормальным законом. В [3] показано, что если на входе распознаваемого автомата использовать стационарный генератор зависимых случайных сигналов, то частота появления фрагмента также будет аппроксимироваться нормальным распределением. С целью оценки метода статистического распознавания автоматов, предложенного в [1], и его сравнения с детерминированным методом в 7-й главе монографии [2] было введено понятие статистической эквивалентности автоматов и выполнено ее сравнение с обычной эквивалентностью автоматов. Оказалось, что статистическая эквивалентность слабее обычной и поэтому статистический метод из [1] слабее детерминированного. Возникает вопрос об оценке статистического метода, использующего генератор зависимых случайных сигналов. Два сильно связных автомата назовем статистически k-эквивалентными, если они нераспознаваемы методом, использующим генератор зависимых случайных сигналов, при любом выборе фрагмента и генератора, глубина зависимости которого не превышает число k. При $k = 0$ имеем статистическую эквивалентность, изученную в [2]. Если сильно связные автоматы нераспознаваемы при любом выборе фрагмента и генератора случайных сигналов с произвольной глубиной зависимости, то они называются обобщенно статистически эквивалентными.

В данной работе для любого k найден эффективно проверяемый критерий статистической k-эквивалентности двух автоматов и показано, что обобщенная статистическая эквивалентность остается слабее обычной эквивалентности, но сильнее статистической эквивалентности автоматов. Формулируются нерешенные проблемы, касающиеся обобщенной статистической эквивалентности.

1. СТАТИСТИЧЕСКАЯ k-ЭКВИВАЛЕНТНОСТЬ

Рассмотрим сильно связный автомат Мили $A=(S,X,Y,\delta,\lambda)$, где $S =\{1,2,...,r\}$ - множество состояний, X - входной алфавит, Y - выходной алфавит, δ - функция переходов и λ - функция выходов. Для целого числа $k \geq 0$ положим $X^k = \{u \in X^* \mid |u| = k\}$, где X^*- множество всех последовательностей конечной длины, включая пустую последовательность e длины 0, а $|u|$ - длина последовательности u. Заметим, что $X^0 = \{e\}$. Будем считать, что источник зависимых случайных сигналов вначале генерирует с некоторой положительной вероятностью последовательность длины k, а затем вероятность появления любого входного символа отлична от нуля и зависит от суффикса длины k уже сгенерированной последоваельности. Пусть $p(u)$ - начальная вероятность появления последовательности $u \in X^k$ и $p(x|u)$ - условная вероятность появления символа $x \in X$, если суффикс длины k уже сгенерированной последовательности есть u. Тогда стационарный генератор случайных сигналов с глубиной зависимости k задается соотношениями

$$p(u) > 0, \quad \sum_{u \in X^k} p(u) = 1, \; p(x|u) > 0, \quad \sum_{x \in X} p(x|u) = 1, \tag{1}$$

справедливыми для всех $x \in X$ и $u \in X^k$.

Зафиксируем фрагмент $v \in X^*$, длина которого $|v|$ больше нуля. Для $n \geq |v| = t$ рассмотрим случайную величину $\zeta_A^n(v)$, определяемую соотношением $\zeta_A^n(v)=1$, если в выходной последовательности автомата A длины n последние t символов составляют фрагмент v, и $\zeta_A^n(v)=0$ в противном случае. В качестве статистической характеристики, по которой в [2] проводится статистическое распознавание автомата A, принимается частота $\nu_A^n(v)$ появления фрагмента v в выходной последовательности длины n. Формально $\nu_A^n(v) = \dfrac{1}{n-t+1} \sum\limits_{k=t}^{n} \zeta_A^k(v)$.

Пусть $A \in \{A_1, A_2\}$, где A_1, A_2 - сильно связные автоматы. Если законы распределения случайных величин $\nu_1^n(v)$ и $\nu_2^n(v)$ для автоматов A_1 и A_2 известны, то задача статистического распознавания автомата A сводится к задаче статистической проверки гипотез [4]. В [3] показано, что для фиксированного фрагмента v и генератора случайных сигналов с глубиной зависимости k распределение случайной величины $\nu_A^n(v)$ аппроксимируется нормальным законом. Математическое ожидание величины $\nu_A^n(v)$, соответствующее аппроксимирующему распределению, обозначим $q_k^A(v)$. Если $q_k^{A_1}(v) = q_k^{A_2}(v)$, то автоматы A_1 и A_2 нельзя распознать статистическим методом при фиксированном фрагменте v и фиксированном генераторе случайных сигналов с глубиной зависимости k.

Определение 1. Сильно связные автоматы A_1 и A_2 с соответственно одинаковыми входными и выходными алфавитами назовем статистически k-эквивалентными (обозначение $\underset{k}{\cong}$), если $q_k^{A_1}(v) = q_k^{A_2}(v)$ при любом выборе фрагмента v и генератора случайных сигналов с глубиной зависимости k, определяемого соотношениями (1). Сильно связные автоматы называются статистически эквивалентными, если они статистически 0-эквивалентны.

Определение 2. Сильно связные автоматы A_1 и A_2 с соответственно одинаковыми входными и выходными алфавитами назовем обобщенно статистически эквивалентными (обозначение $\cong$), если они статистически k-эквивалентны при любом натуральном k.

Рассмотрим вопрос о вычислении математического ожидания $q_k^A(v)$ для произвольных сильно связного автомата $A=(S,X,Y,\delta,\lambda)$, натурального числа k и $v\in Y^*$. При этом не будем фиксировать вероятности из (1), задающие генератор случайных сигналов. Поскольку $q_k^A(v)$ - математическое ожидание аппроксимирующего распределения, то оно не будет зависеть от вероятностей типа $p(u)$ и будет являться функцией переменных типа $p(x|u)$ для $x\in X$ и $u\in X^k$. Если $|X|= m$ ($|X|$ - мощность алфавита X), то $q_k^A(v)$ есть функция m^{k+1} переменных.

Пусть $n \geq 0$, $0 \leq t \leq n$ и $u\in X^n$. Через $\mathrm{sf}_t(u)$ обозначим суффикс длины t последовательности u, а через $\mathrm{pf}_t(u)$ - префикс длины t этой последовательности. Если $t= n$, то $\mathrm{sf}_t(u) = \mathrm{pf}_t(u) = u$, а если $t = 0$, то $\mathrm{sf}_t(u) = \mathrm{pf}_t(u) = e$. При $n \geq 1$ $\mathrm{sf}_1(u)$ ($\mathrm{pf}_1(u)$) - последний (первый) символ последовательности u. Аналогичные обозначения будем использовать для последовательностей в алфавите Y.

Положим $S_k=\{(i,u) \in S\times X^k \mid \exists_{j\in S}\ (i=\delta(j,u))\}$ и для произвольных $(i,u) \in S_k$, $x\in X$ определим функцию $\delta_k((i,u),x)= (\delta(i,x),\mathrm{sf}_k(ux))$. Легко видеть, что δ_k есть отображение $S_k\times X\rightarrow S_k$. Поэтому можно определить автомат без выходов $A_k =(S_k , X, \delta_k)$ и показать, что если A - сильно связный автомат, то A_k - также сильно связный. Автомату A_k и генератору случайных сигналов поставим в соответствие цепь Маркова [5] M^{A_k} , множество состояний которой совпадает с S_k. Будем предполагать, что состояния в S_k пронумерованы таким образом, что вначале идут состояния с первой компонентой 1, затем 2 и т.д. до r. Матрицу переходов цепи M^{A_k} обозначим $M^{A_k} =[m^{A_k}_{(i_1,u_1),(i_2,u_2)}]$ и элемент $m^{A_k}_{(i_1,u_1),(i_2,u_2)}$ определим равенством

$$m^{A_k}_{(i_1,u_1),(i_2,u_2)} = \sum_{x\in X^A(i_2,u_2|i_1,u_1)} p(x|u_1)\,,$$

где $X^A(i_2, u_2| i_1 , u_1) = \{x\in X| \ i_2=\delta(i_1, x)\ \&\ u_2=\mathrm{sf}_k(u_1x)\}$. Поскольку A_k - сильно связный автомат, то цепь M^{A_k} эргодическая и, значит, существует неподвижный положительный (без нулевых компонент) стохастический вектор-строка α^{A_k} с компонентами $\alpha^{A_k}_{(i,u)}$, где $(i ,u) \in S_k$, удовлетворяющий матричному равенству $\alpha^{A_k}\ M^{A_k} =\alpha^{A_k}$. Компонента $\alpha^{A_k}_{(i,u)}$ представляет собой предельную частоту попадания цепи M^{A_k} в состояние (i ,u).

Введем в рассмотрение систему матриц $M^{A_k}(y)= [m^{A_k}_{(i_1,u_1),(i_2,u_2)}(y)]$ для $y\in Y$. Элемент $m^{A_k}_{(i_1,u_1),(i_2,u_2)}(y)$ равен $p(\mathrm{sf}_1(u_2)|\ u_1)$, если $(i_2=\delta(i_1, \mathrm{sf}_1(u_2))\ \&\ y=\lambda(i_1, \mathrm{sf}_1(u_2))\ \&\ \mathrm{pf}_{k-1}(u_2)= \mathrm{sf}_{k-1}(u_1))$, и 0 в противном случае. Если $v' = y_1 \ldots y_n$, где $y_i \in Y$, то матрица $M^{A_k}(v')=[m^{A_k}_{(i_1,u_1),(i_2,u_2)}(v')]$ определяется через произведение матриц $M^{A_k}(v')=$ $M^{A_k}(y_1) \ldots M^{A_k}(y_n)$.

Аналогично [3] можно показать, что при $v\in Y^*$ и $|v| \geq k$ математическое ожидание вычисляется по формуле

$$q_k^A(v)= \alpha^{A_k}\, M^{A_k}(v)\xi^{A_k}\,, \tag{2}$$

где ξ^{A_k} - единичный вектор-столбец размерности $|S_k|$. Формула (2) оказывается справедливой для любых фрагментов $v\in Y^*$. При нефиксированных вероятностях типа $p(x|u)$ элементами матриц M^{A_k} и $M^{A_k}(v)$ будут являться полиномы от m^{k+1} переменных, а компонентами вектора α^{A_k} - рациональные функции тех же переменных, поскольку α^{A_k} находится из матричного равенства $\alpha^{A_k} M^{A_k} = \alpha^{A_k}$. Таким образом, q_k^A есть отображение вида $Y^*\to\Re_k$, где $\Re_k$ - множество рациональных функций m^{k+1} переменных. В [2] безотносительно к автомату аксиоматически определен класс статистических отображений. К этому классу принадлежит и отображение q_k^A.

Напомним некоторые понятия из [2], относящиеся к статистическим отображениям. На Y^* зафиксируем лексикографический порядок $v_0, v_1, \ldots,$ удовлетворяющий условиям $v_0=e$ и $|v_i| \le |v_{i+1}|$ для всех $i \ge 0$, и рассмотрим бесконечномерную матрицу $Q_k^A = [q_k^A(v_i v_j)]$. Ранг этой матрицы rank Q_k^A объявляется рангом статистического отображения q_k^A (обозначение rank q_k^A). Его величина не превышает мощности множества S_k, т.е. rank $q_k^A \le |S_k|$. Пусть rank $q_k^A = r_k^A$ и $(\overline{w},\overline{w'}) = ((w_1,\ldots,w_{r_k^A}),(w_1',\ldots,w'_{r_k^A}))$ - такая пара систем последовательностей в Y^*, что $w_1 = w_1' = e$ и rank $[q_k^A(w_i w_j')] = r_k^A$. Пара $(\overline{w},\overline{w'})$ называется ранговой парой систем последовательностей в Y^* для отображения q_k^A. Для обозначения соответствующих $(r_k^A \times r_k^A)$-матриц введем следующую символику: $F_k^A(\overline{w},\overline{w'}) = [q_k^A(w_i w_j')]$, $F_k^A(\overline{w},v,\overline{w'}) = [q_k^A(w_i v w_j')]$, где $v\in Y^*$. Аналогично доказательству теоремы 7.13 из [2] можно получить эффективно проверяемый критерий статистической k-эквивалентности сильно связных автоматов для любого натурального k.

Теорема 1. *Пусть A,B - сильно связные автоматы с соответственно одинаковыми входными и выходными алфавитами, k - произвольное натуральное число и* $(\overline{w},\overline{w'})$ *- ранговая пара систем последовательностей для отображения* q_k^A. *Автоматы A и B статистически k-эквивалентны тогда и только тогда, когда выполняются следующие условия:*

1. $r_k^A = r_k^B$.
2. $F_k^A(\overline{w},\overline{w'}) = F_k^B(\overline{w},\overline{w'})$.
3. $F_k^A(\overline{w},y,\overline{w'}) = F_k^B(\overline{w},y,\overline{w'})$ *для всех* $y\in Y$.

Пусть $k \ge 1$ и $0 \le t < k$. Матрицы M^{A_k} и $M^{A_k}(v)$, $v\in Y^*$ определяются автоматом A и набором вероятностей $p(x|u)$, $x\in X$, $u\in X^k$. Если для всех $x\in X$ и $u\in X^k$ положить $p(x|u) = \dfrac{1}{m^{k-t}} p(x|\,\mathrm{sf}_t(u))$, то в этом частном случае матрицы M^{A_k} и $M^{A_k}(v)$, $v\in Y^*$ будут определяться автоматом A и набором вероятностей $p(x|w)$, $x\in X$, $w\in X^t$. Поскольку теперь матрицы M^{A_k}, $M^{A_k}(v)$ и M^{A_t}, $M^{A_t}(v)$ $(v\in Y^*)$ определяют одно и то же аппроксимирующее распределение $v_A^n(v)$, то имеет место равенство $q_k^A = q_t^A$. Поэтому справедливо следующее утверждение.

Теорема 2. *Пусть A,B - сильно связные автоматы с соответственно одинаковыми входными и выходными алфавитами, $k \geq 1$ и $0 \leq t < k$. Тогда, если $A \underset{k}{\cong} B$, то $A \underset{t}{\cong} B$.*

2. СПЕКТР ГЕНЕРАТОРА

К сожалению, эффективно проверяемый критерий обобщенной статистической эквивалентности автоматов получить не удалось. Однако, для конкретных неэквивалентных автоматов удалось доказать их обобщенную статистическую эквивалентность, в связи с чем можно сделать вывод, что она слабее обычной эквивалентности. В этом доказательстве используются свойства последовательностей, порождаемых генератором зависимых случайных сигналов [6]. Рассмотрим эти свойства.

Любому генератору случайных сигналов с глубиной зависимости $k \geq 1$, заданному соотношениями (1), соответствует семейство $\{\beta^t(k)\}$ $(t \geq 0)$ положительных (без нулевых компонент) стохастических векторов, которое называется спектром данного генератора. Для произвольного натурального t положительный стохастический вектор-строка $\beta^t(k)= [\beta_u^t(k)]$ $(u \in X^t)$ имеет размерность m^t $(m=|X|)$, а его компонента $\beta_u^t(k)$ представляет собой предельную частоту появления последовательности u в бесконечной случайной последовательности, порожденной генератором. Покажем как вычисляются вектора спектра. Рассмотрим цепь Маркова M_k, состояниями которой являются последовательности из X^k. Матрицу переходов этой цепи также обозначим $M_k = [m_{u_1,u_2}]$, где $u_1, u_2 \in X^k$, считая множество X^k упорядоченным. Элемент m_{u_1,u_2} $(m^k \times m^k)$-матрицы M_k положим равным $p(\mathrm{sf}_1(u_2)|\ u_1)$, если $\mathrm{sf}_{k-1}(u_1)= \mathrm{pf}_{k-1}(u_2)$, и 0 в противном случае. Цепь M_k - регулярна [5] и для нее существует неподвижный положительный стохастический вектор $\beta^k(k)$, однозначно определяемый матричным равенством $\beta^k(k)\ M_k= \beta^k(k)$. Остальные векторы спектра вычисляются следующим образом. Если $t< k$, то $\beta_u^t(k)= \sum\limits_{v \in X^{k-t}} \beta_{vu}^k(k) = \sum\limits_{v \in X^{k-t}} \beta_{uv}^k(k)$. Для $t \geq k$, $u \in X^t$ и $x \in X$ справедлива формула $\beta_{ux}^{t+1}(k) = \beta_u^t(k)\ p(x|\ \mathrm{sf}_k(u))$, индуктивное применение которой позволяет вычислить компоненты любого вектора $\beta^t(k)$ из спектра при $t > k$. Пусть $p(w|u)$ - вероятность появления на выходе генератора последовательности w при условии, что суффикс длины k уже появившейся последовательности равен u, где $u \in X^k$. Если $w=x_1 \ldots x_n$ $(x_i \in X)$, то

$$p(w|u)= p(x_1|\ u)p(x_2 \mid \mathrm{sf}_k(ux_1)) \ldots p(x_n \mid \mathrm{sf}_k(ux_1 \ldots x_{n-1})).$$

В [6] показано, что

$$\beta_u^t(k)p(w|\ \mathrm{sf}_k(u))= \beta_{uw}^{t+\tau}(k) \text{ для всех } t \geq k,\ \tau > 0 \text{ и } w \in X^\tau; \qquad (3)$$

$$\beta_u^t(k)= \sum_{w \in X^\tau} \beta_{uw}^{t+\tau}(k) = \sum_{w \in X^\tau} \beta_{wu}^{t+\tau}(k) \text{ для всех } t, \tau > 0 \text{ и } u \in X^t. \qquad (4)$$

Итак, генератор случайных сигналов с произвольной глубиной зависимости порождает бесконечные последовательности в алфавите X, которые обладают тем свойством, что для любой конечной последовательности из X^* существует положительная предельная частота ее появления в бесконечной последовательности и эту предельную частоту можно вычислить.

Класс спектров генераторов зависимых случайных сигналов можно описать не обращаясь к свойствам генераторов. Пусть X - конечный алфавит, $|X|= m \geq 2$ и для любого натурального $t \geq 0$ X^t упорядочено. Допустим для каждого t=0,1,2,... и $u \in X^t$ $\beta^t = [\beta_u^t]$ - произвольный стохастический вектор размерности m^t, компоненты которого упорядочены в соответствии с порядком на X^t. Семейство векторов $\{\beta^t\}$ ($t \geq 0$) назовем спектром относительно X.

Определение 3. Спектр $\{\beta^t(k)\}$ ($t \geq 0$) относительно алфавита X назовем k-спектром ($k \geq 1$ - целое число), если

1. $\beta_u^k = \sum\limits_{y \in X} \beta_{yu}^{k+1}(k)$ для всех $u \in X^k$;

2. при $t \leq k$ $\beta_u^t(k) = \sum\limits_{x \in X} \beta_{ux}^{t+1}(k)$ при $u \in X^t$;

3. при $t > k$ $\beta_{ux}^t(k) = \beta_u^{t-1}(k)\beta_{\mathrm{sf}_k(u)x}^{k+1}(k) / \beta_{\mathrm{sf}_k(u)}^k(k)$ для всех $x \in X$ и $u \in X^{t-1}$.

Для задания k-спектра $\{\beta^t(k)\}$ ($t \geq 0$) достаточно найти положительный стохастический вектор $\beta^{k+1}(k)$, удовлетворяющий соотношению

$$\sum_{y \in X} \beta_{yu}^{k+1}(k) = \sum_{x \in X} \beta_{ux}^{k+1}(k) \text{ для всех } u \in X^k. \tag{5}$$

Остальные векторы спектра определяются однозначно по вектору $\beta^{k+1}(k)$. Вектор $\beta^t(k)$ вычисляется по формуле из условия 2 при $t \leq k$ и по формуле из условия 3 при $t > k$.

Оказывается, что класс спектров генераторов с глубиной зависимости k совпадает с классом k-спектров. Используя это утверждение, можно показать, что

$$\beta_u^t(k) = \sum_{v \in X^\tau} \beta_{uv}^{t+\tau}(k) = \sum_{v \in X^\tau} \beta_{vu}^{t+\tau}(k) \text{ для всех } t, \tau, k > 0 \text{ и } u \in X^t. \tag{6}$$

3. СРАВНЕНИЕ ЭКВИВАЛЕНТНОСТЕЙ АВТОМАТОВ

Пусть на входе сильно связного автомата A действует генератор случайных сигналов с глубиной зависимости $k \geq 1$. Для любого фрагмента $v \in Y^*$ математическое ожидание $q_k^A(v)$ вычисляется по формуле (2). Можно ли $q_k^A(v)$ выразить через компоненты стохастических векторов спектра генератора? Покажем, что для некоторых автоматов это возможно.

Определение 4. Автомат A назовем полным, если $S_k = S \times X^k$ для всех $k \geq 1$.

Для автомата A и $x \in X$ рассмотрим отображение $\delta_x : S \to S$, определив его равенством $\delta_x(s) = \delta(s,x)$ для всех $s \in S$, и через Val δ_x обозначим область значений этого отображения.

Определение 5. Будем говорить, что A - автомат с перестановочной функцией переходов, если Val $\delta_x = S$ для всех $x \in X$.

Справедливы следующие утверждения.

Лемма 1. *Автомат A - полный тогда и только тогда, когда он является автоматом с перестановочной функцией переходов.*

Лемма 2. *Пусть $A=(S,X,Y,\delta,\lambda)$- сильно связный автомат с перестановочной функцией переходов, у которого $|S|= r$ и $X=\{x_1,...,x_m\}$. Если на входе A действует генератор случайных сигналов с глубиной зависимости $k \geq 1$, то справедливо следующее соотношение:*

$$\alpha^{A_k}_{(i,u)} = \frac{1}{r}\beta^k_u(k) \text{ для всех } i\in S \text{ и } u\in X^k.$$

Для автомата A и произвольных $i\in S$ и $v\in Y^*$ положим $X^A(i,v)=\{w\in X^{|v|} \mid \lambda(i,w)=v\}$.

Теорема 3. *Пусть A - сильно связный автомат с r состояниями и перестановочной функцией переходов, на входе которого действует генератор случайных сигналов с глубиной зависимости $k \geq 1$. Тогда для любого $v\in Y^*$ справедливо равенство*

$$q^A_k(v) = \frac{1}{r}\sum_{i\in S}\sum_{w\in X^A(i,v)}\beta^{|v|}_w(k). \tag{7}$$

Доказательство. Для $v\in Y^*$ рассмотрим вектор-столбец $\eta^{A_k}(v) = M^{A_k}(v)\xi^{A_k}$, компонентами которого являются элементы $\eta^{A_k}_{(i,u)}(v)$, где $(i,u)\in S_k= S\times X^k$. Тогда из (2) находим $q^A_k(v)=\sum_{i\in S}\sum_{u\in X^k}\alpha^{A_k}_{(i,u)}\eta^{A_k}_{(i,u)}(v)$. Учитывая, что $\eta^{A_k}_{(i,u)}(v)=\sum_{w\in X^A(i,v)}p(w|u)$, а также лемму 2, получаем $q^A_k(v) = \frac{1}{r}\sum_{u\in X^k}\beta^k_u(k)\sum_{i\in S}\sum_{w\in X^A(i,v)}p(w|u)$. Используя (3) и (4), приходим к (7). Теорема доказана.

Рассмотрим сильно связные неэквивалентные автоматы A и B, которые заданы таблицами переходов-выходов 1 и 2.

Таблица 1.

i	x: 0	x: 1
1	1/ 0	2/ 0
2	2/ 1	1/ 1

Таблица 2.

i	x: 0	x: 1
1	1/ 0	2/ 1
2	2/ 1	1/ 0

Докажем, что $A \cong B$. Поскольку A и B - автоматы с перестановочными функциями переходов, то для вычисления математических ожиданий $q^A_k(v)$ и $q^B_k(v)$ ($k \geq 1$, $v\in\{0,1\}^*$) можно воспользоваться формулой (7). Поэтому для произвольных $k\geq 1$, $t\geq 0$ и $v\in Y^t$ имеют место равенства

$$q^A_k(v) = \frac{1}{2}\Big(\sum_{w\in X^A(1,v)}\beta^t_w(k) + \sum_{w\in X^A(2,v)}\beta^t_w(k)\Big),\ q^B_k(v) = \frac{1}{2}\Big(\sum_{w\in X^B(1,v)}\beta^t_w(k) + \sum_{w\in X^B(2,v)}\beta^t_w(k)\Big). \tag{8}$$

Доказательство равенства $q^A_k(v) = q^B_k(v)$ базируется на следующем утверждении, справедливость которого можно установить индукцией по длине последовательностей v, w и u.

Лемма 3. *Для автоматов A, B и произвольных $t\geq 0$, $v\in Y^t$ существуют такие w, $u\in X^t$, что*

$$X^A(2, 0v)=\varnothing,\ X^A(1, 0v)=\{w0, w1\},\ X^B(1, 0v)=\{0w\},\ X^B(2, 0v)=\{1w\}, \tag{9}$$

$$X^A(1, 1v) = \varnothing,\ X^A(2, 1v) = \{u0, u1\},\ X^B(1, 1v) = \{1u\},\ X^B(2, 1v) = \{0u\}.$$

Теперь все подготовлено для доказательства эквивалентности $A \cong B$. Если $v = e$. то $q_k^A(v) = q_k^B(v) = 1$. Пусть $v' \in Y^t$ и $t \geq 1$. Рассмотрим два случая.

Случай 1: $v' = 0v$. В соответствии с леммой 3 в этом случае существует $w \in X^{t-1}$, удовлетворяющее (9). Используя (8) и (4), находим

$$q_k^A(v') = \frac{1}{2}(\beta_{w0}^t(k) + \beta_{w1}^t(k)) = \frac{1}{2}\beta_w^{t-1}(k)\,,\ q_k^B(v') = \frac{1}{2}(\beta_{0w}^t(k) + \beta_{1w}^t(k)) = \frac{1}{2}\beta_w^{t-1}(k)$$

и делаем вывод, что $q_k^A(v') = q_k^B(v')$.

Случай 2: $v' = 1v$. В этом случае лемма 3 гарантирует существование $u \in X^{t-1}$, удовлетворяющего (9). На основании (8) и (4) получаем

$$q_k^A(v') = \frac{1}{2}(\beta_{u0}^t(k) + \beta_{u1}^t(k)) = \frac{1}{2}\beta_u^{t-1}(k)\,,\ q_k^B(v') = \frac{1}{2}(\beta_{0u}^t(k) + \beta_{1u}^t(k)) = \frac{1}{2}\beta_u^{t-1}(k)$$

и убеждаемся в справедливости равенства $q_k^A(v') = q_k^B(v')$.

Поскольку глубина зависимости k генератора была выбрана произвольной, то можно сделать вывод, что автоматы A и B обобщенно статистически эквивалентны. Учитывая обычную неэквивалентность A и B, можно считать доказанным следующее утверждение.

Теорема 4. *Существуют неэквивалентные сильно связные автоматы, которые являются обобщенно статистически эквивалентными, т.е. обобщенная статистическая эквивалентность слабее обычной эквивалентности сильно связных автоматов.*

Теперь установим, что обобщенная статистическая эквивалентность сильнее статистической эквивалентности автоматов. Рассмотрим сильно связные неэквивалентные автоматы C и D, которые заданы таблицами переходов-выходов 3 и 4.

Таблица 3.

i	x = 0	x = 1
1	2/ 1	1/ 0
2	1/ 0	3/ 0
3	3/ 0	2/ 0

Таблица 4.

i	x = 0	x = 1
1	2/ 0	1/ 0
2	1/ 1	3/ 0
3	3/ 0	2/ 0

Используя теорему 1 при $k = 0$, покажем статистическую эквивалентность автоматов C и D. Пусть p - вероятность появления на выходе генератора независимых случайных сигналов символа 0. Можно показать, что ранги статистических отображений q_0^C и q_0^D одинаковы и равны 3. Ранговая пара $(\overline{w}, \overline{w}')$ систем последовательностей для отображения q_0^C определяется равенством $\overline{w} = \overline{w}' = (e, 0, 00)$. Соответствующие вычисления позволяют найти

$$F_0^C(\overline{w},\overline{w}) = F_0^D(\overline{w},\overline{w}) = \frac{1}{3}\begin{bmatrix} 3 & 3-p & 3-2p \\ 3-p & 3-2p & 3-3p+p^3 \\ 3-2p & 3-3p+p^3 & 3-4p+3p^3-p^4 \end{bmatrix},$$

$$F_0^C(\overline{w},0,\overline{w}) = F_0^D(\overline{w},0,\overline{w}) = \frac{1}{3}\begin{bmatrix} 3-p & 3-2p & 3-3p+p^3 \\ 3-2p & 3-3p+p^3 & 3-4p+3p^3-p^4 \\ 3-3p+p^3 & 3-4p+3p^3-p^4 & 3-5p+7p^3-6p^4+2p^5 \end{bmatrix},$$

$$F_0^C(\overline{w},1,\overline{w}) = F_0^D(\overline{w},1,\overline{w}) = \frac{1}{3}\begin{bmatrix} p & p & p(1-p^2) \\ p & p & p(1-p^2) \\ p(1-p^2) & p(1-p^2) & p(1-p^2)^2 \end{bmatrix}$$

и на основании теоремы 1 сделать вывод о статистической эквивалентности автоматов C и D.

Теперь покажем, что автоматы C и D не являются обобщенно статистически эквивалентными. Заметим, что C и D - автоматы с перестановочными функциями переходов. Поэтому для вычисления математических ожиданий можно пользоваться формулой (7). Пусть $A=(S,X,Y,\delta,\lambda)$ - автомат с перестановочной функцией переходов, $k \geq 1$ и $v\in Y^*$. При вычислении $q_k^A(v)$ по формуле (7) необходимо знать множества входных последовательностей $X^A(i,v)$ для всех $i\in S$. Опишем метод вычисления этих множеств.

Для автомата A, $r=|S|$ и $y\in Y$ рассмотрим $r\times r$-матрицу $P^A(y)=[\ p_{ij}^A(y)\ ]$, в которой элемент $p_{ij}^A(y)$ представляет собой формальную сумму входных сигналов из $X^A(j,y|\ i)= \{x\in X|\ \ \delta(i,x)=j\ \&\ \ \lambda(i,x)=y\}$. Эти матрицы можно умножать, если под операцией умножения понимать конкатенацию входных последовательностей, а под операцией сложения - их формальную сумму (коммутативную). Будем считать,что если $X^A(j,y|\ i)=\varnothing$, то $p_{ij}^A(y)=\varnothing$, и $\varnothing+u=u+\varnothing=u$, $\varnothing u=u\varnothing=\varnothing$ для всех $u\in X^*$. Если $v=y_1 \dots y_n$ $(y_i\in Y)$, то положим $P^A(v)=P^A(y_1)\dots P^A(y_n)$, где справа от знака равенства стоит определенное выше произведение матриц. Пусть $\gamma^A(v)=[\gamma_i^A(v)]$ - вектор-столбец, полученный из матрицы $P^A(v)$ формальным суммированием ее строк. Компонента $\gamma_i^A(v)$ $(i\in S)$ этого вектора будет представлять собой формальную сумму последовательностей из $X^A(i,v)$. Для вектора $\gamma^A(v)$ через $\|\gamma^A(v)\|$ обозначим формальную сумму его компонент, при этом выражение $u+\dots+u$ (c слагаемых) будем записывать в виде $c\cdot u$, где $u\in X^*$. Предположим, что для целого $k\geq 1$ и $|v|=t$ $\tau_k(\|\gamma^A(v)\|)$ является суммой $\beta_u^t(k)$ для всех последовательностей u, входящих в формальную сумму $\|\gamma^A(v)\|$ (если в формальной сумме перед u стоит натуральный коэффициент c, то в полученной сумме будет слагаемое $c\,\beta_u^t(k)$). Тогда из (7) получаем

$$q_k^A(v) = \frac{1}{r}\,\tau_k(\|\gamma^A(v)\|). \tag{10}$$

Используем формулу (10) при вычислении математических ожиданий для автоматов C и D. Находим

$$P^{C}(0)=\begin{bmatrix}1 & \varnothing & \varnothing\\ 0 & \varnothing & 1\\ \varnothing & 1 & 0\end{bmatrix},\ \gamma^{C}(0)=\begin{bmatrix}1\\ 0+1\\ 0+1\end{bmatrix},\ P^{D}(0)=\begin{bmatrix}1 & 0 & \varnothing\\ \varnothing & \varnothing & 1\\ \varnothing & 1 & 0\end{bmatrix},\ \gamma^{D}(0)=\begin{bmatrix}0+1\\ 1\\ 0+1\end{bmatrix}.$$

Далее, вычисляя $\gamma^{C}(0^4)=(P^{C}(0))^3\gamma^{C}(0)$ и $\gamma^{D}(0^4)=(P^{D}(0))^3\gamma^{D}(0)$, получаем

$$\gamma^{C}(0^4)=\begin{bmatrix}1111\\ 0111+1101+1111+1110+1010+1011+1001+1000\\ 1011+1110+1111+1101+1100+0010+0101+0111+0110+0011+0001+0000\end{bmatrix},$$

$$\gamma^{D}(0^4)=\begin{bmatrix}1111+1110+1101+1011+1010+0111+0101+0100\\ 1111+1110+1011+1001+1000\\ 1111+1101+1100+0111+0110+0011+0001+0000\end{bmatrix}.$$

Воспользовавшись формулой (10), можно выразить соответствующие математические ожидания через компоненты вектора $\beta^4(k)$ спектра генератора. Имеем

$$q_k^C(0^4)=\frac{1}{3}(\beta^4_{0000}(k)+\beta^4_{0001}(k)+\beta^4_{0010}(k)+\beta^4_{0011}(k)+\beta^4_{0101}(k)+\beta^4_{0110}(k)+$$
$$+2\beta^4_{0111}(k)+\beta^4_{1000}(k)+\beta^4_{1001}(k)+\beta^4_{1010}(k)+2\beta^4_{1011}(k)+\beta^4_{1100}(k)+$$
$$+2\beta^4_{1101}(k)+2\beta^4_{1110}(k)+3\beta^4_{1111}(k)),$$

$$q_k^D(0^4)=\frac{1}{3}(\beta^4_{0000}(k)+\beta^4_{0001}(k)+\beta^4_{0011}(k)+\beta^4_{0100}(k)+\beta^4_{0101}(k)+\beta^4_{0110}(k)+$$
$$+2\beta^4_{0111}(k)+\beta^4_{1000}(k)+\beta^4_{1001}(k)+\beta^4_{1010}(k)+2\beta^4_{1011}(k)+\beta^4_{1100}(k)+$$
$$+2\beta^4_{1101}(k)+2\beta^4_{1110}(k)+3\beta^4_{1111}(k)).$$

Эти выражения можно упростить, применяя (6). В итоге получаем

$$q_k^C(0^4)=\frac{1}{3}(\beta^1_{1}(k)+2\beta^2_{11}(k)+\beta^3_{000}(k)+\beta^3_{101}(k)+\beta^4_{0010}(k)),$$

$$q_k^D(0^4)=\frac{1}{3}(\beta^1_{1}(k)+2\beta^2_{11}(k)+\beta^3_{000}(k)+\beta^3_{010}(k)+\beta^4_{1101}(k)).$$

Второе выражение можно преобразовать таким образом, чтобы оно отличалось от первого только одним слагаемым. Используя (6), находим

$$\beta^3_{010}(k)+\beta^4_{1101}(k)=\beta^4_{0100}(k)+\beta^4_{0101}(k)+\beta^4_{1101}(k)=\beta^3_{101}(k)+\beta^4_{0100}(k).$$

Поэтому

$$q_k^D(0^4)=\frac{1}{3}(\beta^1_{1}(k)+2\beta^2_{11}(k)+\beta^3_{000}(k)+\beta^3_{101}(k)+\beta^4_{0100}(k)).$$

Для доказательства того, что $q_k^C(0^4)\neq q_k^D(0^4)$ при некотором $k\geq 1$, достаточно найти такой k-спектр, в котором $\beta^4_{0010}(k)\neq\beta^4_{0100}(k)$. Положим k=3, $0<\varepsilon<1/2^4$ и определим положительный стохастический вектор $\beta^4(3)$ следующим образом: $\beta^4_u(3)$ равно $1/2^4+\varepsilon$, если (u= 0011 или u= 1010), $1/2^4-\varepsilon$, если (u= 0010 или u= 1011), и $1/2^4$ в остальных случаях. Непосредственной проверкой можно убедиться, что вектор $\beta^4(3)$ удовлетворяет соотношению (5) и, следовательно, задает k-спектр, который порождается некоторым генератором с глубиной зависимости 3. Из определения вектора $\beta^4(3)$ видно, что $\beta^4_{0010}(3)\neq\beta^4_{0100}(3)$ и, значит, $q_3^C(0^4)\neq q_3^D(0^4)$. Таким образом, автоматы C и D не являются обобщенно статистически эквивалентными. В результате доказано следующее утверждение.

Теорема 5. *Существуют сильно связные статистически эквивалентные автоматы, которые не являются обобщенно статистически эквивалентными, т. е. обобщенная статистическая эквивалентность сильнее статистической эквивалентности сильно связных автоматов.*

4. ЗАКЛЮЧЕНИЕ

На примере использования генераторов для статистического распознавания автоматов показано, что генераторы зависимых случайных сигналов мощнее независимых. Однако, не удалось выяснить, является ли проблема проверки обобщенной статистической эквивалентности произвольных сильно связных автоматов (в частности, автоматов с перестановочными функциями переходов) алгоритмически разрешимой.

ЛИТЕРАТУРА

1. Барашко А.С. Математические основы контроля дискретных устройств методом счета фрагментов // Кибернетика и системный анализ. - 1994. - № 2. - С. 44 - 51.
2. Барашко А.С., Скобцов Ю.А., Сперанский Д.В. Моделирование и тестирование дискретных устройств. - Киев: Наук. думка, 1992. - 285 с.
3. Барашко А.С., Павлив А.Н. Обобщенный подход к статистическому распознаванию автоматов // Кибернетика и системный анализ. - 1998. - № 1. - С. 46 - 56.
4. Фукунага К. Введение в статистическую теорию распознавания образов. - М.: Наука, 1979. - 367 с.
5. Кемени Д.Д., Снелл Д.Л. Конечные цепи Маркова. - М.: Наука, 1970. - 271 с.
6. Барашко А.С. Свойства последовательностей, порождаемых генератором зависимых случайных сигналов // Доповζдζ НАН Укра|ни. – 1998. - № 3. – С. 104 – 109.

Поступила в редакційну колегію 28.12.2002