

УДК 519.2:621.391

Ю.С. Харин, А.Л. Костевич

**РАСПОЗНАВАНИЕ СТОХАСТИЧЕСКИХ СИГНАЛОВ
С ИСПОЛЬЗОВАНИЕМ ЦЕПЕЙ МАРКОВА s -ГО ПОРЯДКА***

Для решения задачи дискриминантного анализа стохастических сигналов предлагается переходить к дискретному сигналу, описываемому цепью Маркова s -го порядка; задача распознавания исходного сигнала решается в пространстве параметров цепей Маркова. Построены байесовское решающее правило, подстановочное байесовское решающее правило классификации цепей Маркова s -го порядка, и в асимптотике сближающихся классов найдено выражение для вероятности ошибочной классификации. Предложенный подход успешно применен к биомедицинским сигналам для диагностики коронарной ишемической болезни.

Введение

При создании и анализе компьютерных систем распознавания сигналов различной природы адекватной математической моделью является модель стохастического (случайного) сигнала $x(t) \in X$, где $X \subseteq \mathbb{R}^N$ – пространство значений сигнала ($N \geq 1$), $t \in \mathfrak{T}$ – время, $\mathfrak{T} \subseteq \mathbb{R}^1$ – временная область, на которой наблюдается подлежащий распознаванию сигнал. Если $N > 1$, то стохастический сигнал – векторный. В компьютерных системах обрабатываются сигналы с дискретным временем, поэтому \mathfrak{T} – дискретное множество, а сигнал $x(t)$ принято называть стохастической (случайной) последовательностью, либо временным рядом [1]. Приведем примеры задач распознавания стохастических сигналов [2 - 11]: распознавание электрокардиограмм, распознавание шумовых сигналов в технической диагностике, распознавание элементов речевых сигналов, распознавание радио- и гидролокационных сигналов; распознавание кардиозаболеваний по шумовым сигналам кровотока.

В настоящее время при распознавании стохастических сигналов можно выделить два основных подхода [2]. Первый подход состоит в отыскании «малого числа» так называемых «информативных» признаков распознаваемых сигналов, преобразовании наблюдаемых сигналов из исходного функционального пространства высокой размерности в пространство признаков и, наконец, построении разделяющих гиперповерхностей в пространстве признаков методами дискриминантного анализа [3,4]. Основная трудность этого подхода состоит в отсутствии эффективных методов формирования «информативных» признаков и в отсутствии либо сложности физической модели распознаваемых сигналов, используемой при построении признаков.

Второй подход использует вероятностную модель временного ряда в исходном функциональном пространстве сигналов, представляющую собой некоторое регулярное (параметрическое, семипараметрическое или непараметрическое) семейство вероятностных распределений [2,5,6]. Приведем примеры таких вероятностных моделей [1,2,12]: временной ряд с трендом, авторегрессионный ряд, временной ряд скользящего среднего, временной ряд ARIMA (p, d, q).

В данной статье предлагается использовать подход, объединяющий свойства первого и второго подходов, а именно: осуществляется дискретизация пространства значений X наблюдаемого сигнала и затем для полученного дискретного по времени и по значениям процесса строится вероятностная модель цепи Маркова s -го порядка [12].

1. Представление сигналов с помощью цепей Маркова s -го порядка

Пусть на основном вероятностном пространстве (Ω, F, P) определены $L \geq 2$ классов $\Omega_1, \dots, \Omega_L$ стохастических сигналов с дискретным временем $t \in \mathfrak{T} = \{1, 2, \dots, n\}$ (временных рядов):

*Данная работа выполнена при поддержке МНТЦ, проект В-705.

$$\{x_t^{(1)} : t \in \mathfrak{T}\}, \{x_t^{(2)} : t \in \mathfrak{T}\}, \dots, \{x_t^{(L)} : t \in \mathfrak{T}\} \quad (1)$$

соответственно, причём класс Ω_l наблюдается с вероятностью q_l , $\sum_{l=1}^L q_l = 1$; здесь n – длительность наблюдения сигнала. Зададим упорядоченный набор $N-1$ чисел $G = \{g_1, \dots, g_{N-1}\}$, определяющий разбиение числовой прямой на N непересекающихся промежутков (ячеек):

$$-\infty = g_0 < g_1 < g_2 < \dots < g_{N-1} < g_N = +\infty; \quad (2)$$

$$\mathfrak{R} = \bigcup_{i=1}^N G_i, \quad G_i = (g_{i-1}, g_i]. \quad (3)$$

С помощью (2), (3) осуществим преобразование сигналов (1) с пространством значений $X = \mathfrak{R}$ в дискретные временные ряды с конечным множеством значений $A = \{1, 2, \dots, N\}$:

$$y_t^{(l)} = \begin{cases} 1, & \text{если } x_t^{(l)} \in G_1, \\ \dots & t \in \mathfrak{T}; \quad l = 1, \dots, L. \\ N, & \text{если } x_t^{(l)} \in G_N, \end{cases} \quad (4)$$

Будем описывать временной ряд $\{y_t^{(l)}\}$ ($l = 1, \dots, L$) однородной цепью Маркова s -го порядка ($s \geq 1$) с пространством состояний A [12]. Если $s = 1$, то цепь Маркова называется простой; если $s > 1$ – сложной.

Вероятностная модель класса Ω_l задаётся s -мерной матрицей начального распределения вероятностей ($l = 1, \dots, L$):

$$\pi^{(l)} = (\pi_{i_1, \dots, i_s}^{(l)}), \quad \pi_{i_1, \dots, i_s}^{(l)} = P\{y_{t-1}^{(l)} = i_s, \dots, y_{t-s}^{(l)} = i_1\}, \quad t > s, \quad (5)$$

и $(s+1)$ -мерной матрицей вероятностей одношаговых переходов:

$$P^{(l)} = (p_{i_1, \dots, i_s, i_{s+1}}^{(l)}), \quad p_{i_1, \dots, i_s, i_{s+1}}^{(l)} = P\{y_t^{(l)} = i_{s+1} \mid y_{t-1}^{(l)} = i_s, \dots, y_{t-s}^{(l)} = i_1\}. \quad (6)$$

Будем предполагать, что начальное распределение вероятностей $\pi^{(l)}$ является стационарным распределением. Выполняются следующие соотношения нормировки ($l = 1, \dots, L$):

$$\sum_{i_1, \dots, i_s=1}^N \pi_{i_1, \dots, i_s}^{(l)} = 1, \quad (7)$$

$$\sum_{i_{s+1}=1}^N p_{i_1, \dots, i_s, i_{s+1}}^{(l)} = 1, \quad i_1, \dots, i_s \in A.$$

Отметим, что для оценивания порядка s цепи Маркова по наблюдениям могут использоваться известные методы [12].

Классы $\{\Omega_l\}$ различаются вероятностными характеристиками (5), (6). Требуется построить решающее правило (алгоритм распознавания) для классификации в L классов наблюдаемой в течение n_0 единиц времени реализации сигнала x_1, \dots, x_{n_0} , преобразованной согласно (4) в дискретный временной ряд $Y = (y_1, y_2, \dots, y_{n_0})' \in A^{n_0} : d = d(Y), d \in \{1, 2, \dots, L\}$. В качестве показателя точности решающего правила (РП) будем использовать вероятность ошибочной классификации сигнала:

$$r = \mathbf{P}\{d \neq \nu\} \rightarrow \min_{d(\cdot)}, \quad (8)$$

где $\nu \in \{1, 2, \dots, L\}$ – дискретная случайная величина с распределением вероятностей $\mathbf{P}\{\nu = l\} = q_l$, указывающая истинный ненаблюдаемый номер класса регистрируемого стохастического сигнала.

2. Байесовское решающее правило и его точность

Рассмотрим сначала случай, когда параметры (5), (6) классов $\{\Omega_l\}$ известны.

В теореме 1 построено оптимальное решающее правило (РП) классификации реализаций цепей Маркова s -го порядка.

Теорема 1. *Оптимальное в смысле минимума вероятности ошибочной классификации (8) байесовское решающее правило классификации реализации цепи Маркова $Y = (y_1, y_2, \dots, y_{n_0})$ для модели (5), (6) имеет вид*

$$d(Y) = \arg \max_{1 \leq l \leq L} \left(\ln q_l + \ln \pi_{y_1, \dots, y_s}^{(l)} + \sum_{i_1, \dots, i_{s+1} \in A} n_{i_1, \dots, i_s, i_{s+1}} \ln p_{i_1, \dots, i_s, i_{s+1}}^{(l)} \right); \tag{9}$$

$$n_{i_1, \dots, i_s, i_{s+1}} = \sum_{t=1}^{n_0-s} \mathbf{I}(y_t = i_1, \dots, y_{t+s-1} = i_s, y_{t+s} = i_{s+1}), \quad i_1, \dots, i_s, i_{s+1} \in A, \tag{10}$$

где $\mathbf{I}(E)$ – индикатор события E , $\{n_{i_1, \dots, i_s, i_{s+1}}\}$ – множество частот $(s+1)$ состояний, вычисленных по классифицируемой реализации Y .

Доказательство непосредственно следует из [2, 7] при использовании логарифмической функции правдоподобия для модели (5), (6).

Следствие 1. *В асимптотике растущей длительности наблюдения $n_0 \rightarrow \infty$ байесовское РП (9), (10) допускает представление*

$$d(Y) = \arg \max_{1 \leq l \leq L} \sum_{i_1, \dots, i_{s+1} \in A} \hat{\Pi}_{i_1, \dots, i_s, i_{s+1}} \ln p_{i_1, \dots, i_s, i_{s+1}}^{(l)}, \quad \hat{\Pi}_{i_1, \dots, i_s, i_{s+1}} = \frac{n_{i_1, \dots, i_s, i_{s+1}}}{n_0}. \tag{11}$$

Следствие 2. *В случае двух классов ($L = 2$) байесовское РП (11) имеет вид*

$$d(Y) = \mathbf{I} \left(\sum_{i_1, \dots, i_{s+1} \in A} \hat{\Pi}_{i_1, \dots, i_s, i_{s+1}} \ln \frac{p_{i_1, \dots, i_s, i_{s+1}}^{(2)}}{p_{i_1, \dots, i_s, i_{s+1}}^{(1)}} > 0 \right) + 1. \tag{12}$$

Исследуем теперь вероятность ошибочной классификации байесовского РП в асимптотике растущей длительности наблюдения $n_0 \rightarrow \infty$. Рассмотрим случай двух классов. Совместно с асимптотикой $n_0 \rightarrow \infty$ будем исследовать наиболее трудный случай сближающихся классов:

$$p_{i_1, \dots, i_s, i_{s+1}}^{(2)} = p_{i_1, \dots, i_s, i_{s+1}}^{(1)} (1 + \omega_{i_1, \dots, i_s, i_{s+1}} \varepsilon) \text{ и } p_{i_1, \dots, i_s, i_{s+1}}^{(2)} \rightarrow p_{i_1, \dots, i_s, i_{s+1}}^{(1)} \text{ при } \varepsilon \rightarrow 0, \tag{13}$$

где $\omega_{i_1, \dots, i_s, i_{s+1}}$ – весовые коэффициенты ($\sum_{i_{s+1} \in A} \omega_{i_1, \dots, i_s, i_{s+1}} p_{i_1, \dots, i_s, i_{s+1}}^{(1)} = 0$), $i_1, \dots, i_{s+1} \in A$. Асимптотика сближающихся классов при $\varepsilon = \varepsilon(n_0)$ позволяет оценить «расстояние» между параметрами классов, при котором классы различимы при классификации реализаций цепей Маркова длительности n_0 .

Введем следующие обозначения:

$$a_l = (-1)^l \sum_{i_1, \dots, i_s \in A} \pi_{i_1, \dots, i_s}^{(l)} \sum_{i_{s+1} \in A} p_{i_1, \dots, i_s, i_{s+1}}^{(l)} \ln \frac{p_{i_1, \dots, i_s, i_{s+1}}^{(2)}}{p_{i_1, \dots, i_s, i_{s+1}}^{(1)}};$$

$$\sigma_l^2 = \sum_{\substack{i_1, \dots, i_s, i_{s+1} \\ j_1, \dots, j_s, j_{s+1} \in A}} \ln \frac{P_{i_1, \dots, i_s, i_{s+1}}^{(2)}}{P_{i_1, \dots, i_s, i_{s+1}}^{(1)}} \ln \frac{P_{j_1, \dots, j_s, j_{s+1}}^{(2)}}{P_{j_1, \dots, j_s, j_{s+1}}^{(1)}} b_{i_1, \dots, i_s, i_{s+1}, j_1, \dots, j_s, j_{s+1}}^{(l)};$$

$$b_{i_1, \dots, i_s, i_{s+1}, j_1, \dots, j_s, j_{s+1}}^{(l)} = \pi_{i_1, \dots, i_s}^{(l)} P_{i_1, \dots, i_s, i_{s+1}}^{(l)} \left(\prod_{t=1}^s \mathbf{I}(i_t = j_t) - \pi_{j_1, \dots, j_s}^{(l)} P_{j_1, \dots, j_s, j_{s+1}}^{(l)} \right) +$$

$$+ P_{i_1, \dots, i_s, i_{s+1}}^{(l)} P_{j_1, \dots, j_s, j_{s+1}}^{(l)} \left(\pi_{i_1, \dots, i_s}^{(l)} c_{i_2, \dots, i_{s+1}, j_1, \dots, j_s}^{(l)} + \pi_{j_1, \dots, j_s}^{(l)} c_{j_2, \dots, j_{s+1}, i_1, \dots, i_s}^{(l)} \right);$$

$$c_{i_1, \dots, i_s, j_1, \dots, j_s}^{(l)} = \sum_{t=1}^{\infty} \left(\tilde{p}_{i_1, \dots, i_s, j_1, \dots, j_s}^{(l)}(t) - \pi_{j_1, \dots, j_s}^{(l)} \right) < \infty,$$

где a_j – дивергенция Кульбака [12] между классами Ω_1, Ω_2 ; $\tilde{p}_{i_1, \dots, i_s, j_1, \dots, j_s}^{(l)}(t)$ – вероятность перехода цепи Маркова для класса Ω_l за t шагов из последовательности состояний (i_1, \dots, i_s) в последовательность состояний (j_1, \dots, j_s) .

Теорема 2. Если имеет место асимптотика растущей длительности наблюдения $n_0 \rightarrow \infty$ и сближающихся классов (13) при $\varepsilon = c/\sqrt{n_0}$, тогда вероятность ошибочной классификации (8) байесовского РП (12) имеет предельное значение:

$$r \rightarrow r_0 = \Phi\left(-\frac{\Delta}{2}\right), \quad \Delta = \frac{c\mu}{\sqrt{V}} > 0, \quad \mu = \sum_{\substack{i_1, \dots, i_s, i_{s+1} \\ j_1, \dots, j_s, j_{s+1} \in A}} \omega_{i_1, \dots, i_s, i_{s+1}}^2 \pi_{i_1, \dots, i_s}^{(1)} P_{i_1, \dots, i_s, i_{s+1}}^{(1)},$$

$$V = \sum_{\substack{(i_1, \dots, i_s, i_{s+1}) \\ (j_1, \dots, j_s, j_{s+1}) \in A_{\Pi}}} \left(\omega_{i_1, \dots, i_s, i_{s+1}} - \omega_{N, \dots, N} \right) b_{i_1, \dots, i_s, i_{s+1}, j_1, \dots, j_s, j_{s+1}}^{(1)} \left(\omega_{j_1, \dots, j_s, j_{s+1}} - \omega_{N, \dots, N} \right) > 0,$$

где $\Phi(\cdot)$ – функция распределения стандартного нормального закона, $A_{\Pi} = A^s \setminus \{(N, \dots, N)\}$.

Доказательство следует из [7] при использовании свойств сложных цепей Маркова из [12].

Вычислительная сложность решающих правил (9)-(12) есть $O(n_0)$.

3. Подстановочное байесовское решающее правило в условиях априорной неопределенности

На практике параметры классов (5), (6) зачастую неизвестны, и наблюдается классифицированная обучающая выборка:

$$\mathbf{Y} = (Y^{(1)}, Y^{(2)}), \quad Y^{(l)} = (y_1^{(l)}, \dots, y_{n_l}^{(l)}), \quad l \in \{1, 2\},$$

где $Y^{(l)}$ – реализация цепи Маркова длины n_l из класса Ω_l . Предполагается, что реализации $Y, Y^{(1)}, Y^{(2)}$ независимы в совокупности. В этом случае по обучающей выборке могут быть построены оценки максимального правдоподобия [13] неизвестных матриц вероятностей одношаговых переходов $P^{(l)}$:

$$\tilde{P}^{(l)} = \left(\tilde{P}_{i_1, \dots, i_s, i_{s+1}}^{(l)} \right): \tilde{P}_{i_1, \dots, i_s, i_{s+1}}^{(l)} = \frac{n_{i_1, \dots, i_s, i_{s+1}}^{(l)}}{n_{i_1, \dots, i_s}^{(l)}}, \quad i_1, \dots, i_s, i_{s+1} \in A, \quad l \in \{1, 2\};$$

$$n_{i_1, \dots, i_s, i_{s+1}}^{(l)} = \sum_{t=1}^{n_l-s} \mathbf{I}(y_t^{(l)} = i_1, \dots, y_{t+s-1}^{(l)} = i_s, y_{t+s}^{(l)} = i_{s+1}).$$

Тогда подстановочное РП может быть получено из байесовского РП (12) заменой неизвестных параметров их оценками максимального правдоподобия:

$$d(Y; \mathbf{Y}) = \mathbf{I} \left(\sum_{i_1, \dots, i_{s+1} \in A} \hat{\Pi}_{i_1, \dots, i_{s+1}} \ln \frac{\hat{P}_{i_1, \dots, i_{s+1}}^{(2)}}{\hat{P}_{i_1, \dots, i_{s+1}}^{(1)}} > 0 \right) + 1. \quad (14)$$

Введем следующие обозначения:

$$\sigma_{i_1, \dots, i_s, j_{s+1}, j_1, \dots, j_s, j_{s+1}}^{(l)} = \frac{1}{\pi_{i_1, \dots, i_s}^{(l)}} \left(\prod_{t=1}^s \mathbf{I}(i_t = j_t) \left(\mathbf{I}(i_{s+1} = j_{s+1}) P_{i_1, \dots, i_s, j_{s+1}}^{(l)} - P_{i_1, \dots, i_s, j_{s+1}}^{(l)} P_{j_1, \dots, j_s, j_{s+1}}^{(l)} \right) \right),$$

где $i_1, \dots, i_{s+1}, j_1, \dots, j_{s+1} \in A$.

Теорема 3. Если имеет место асимптотика растущей длительности наблюдения и сближающихся классов (13):

$$n_0, n_1, n_2 \rightarrow \infty, \quad \frac{n_l}{n_0} = \tilde{\lambda}_l > 0, \quad \varepsilon = \frac{c}{\sqrt{n_0}}, \quad 0 < c < \infty, \quad l \in \{1, 2\},$$

то вероятность ошибочной классификации (8) подстановочного РП (14) имеет предельное значение:

$$r \rightarrow \tilde{r} = \Phi \left(-\frac{\tilde{\Delta}_1}{2} \right) + \Phi \left(-\frac{\tilde{\Delta}_2}{2} \right), \quad \tilde{\Delta}_l = \frac{c\mu}{\sqrt{V + \tilde{V}_l}} > 0,$$

$$\tilde{V}_l = \frac{1}{\tilde{\lambda}_{3-l}} \sum_{i_1, \dots, i_s} \sum_{\substack{i_{s+1}, j_{s+1}=1 \\ j_1, \dots, j_s \in A}}^{N-1} \left(\omega_{i_1, \dots, i_s, j_{s+1}} - \omega_{i_1, \dots, i_s, N} \right) \sigma_{i_1, \dots, i_s, j_{s+1}, j_1, \dots, j_s, j_{s+1}}^{(l)} \left(\omega_{j_1, \dots, j_s, j_{s+1}} - \omega_{j_1, \dots, j_s, N} \right) > 0,$$

где μ, V определены в теореме 2.

Доказательство использует метод асимптотического разложения риска из [2, 7] и асимптотические свойства для цепей Маркова из [12].

Вычислительная сложность подстановочного РП (14) есть $O(n_0 + n_1 + n_2)$.

4. Численные результаты

Рассмотрим применение предложенного метода распознавания к диагностике коронарной ишемической болезни сердца с использованием статистической классификации акустических сигналов кровотока.

Известно [9], что коронарная ишемическая болезнь сердца характеризуется появлением высокочастотных шумов в акустическом сигнале кровотока. Сигнал кровотока является нестационарным сигналом и состоит из трех фрагментов: систола, сокращение сердца и диастола. Для диагностики коронарной ишемической болезни наиболее информативный фрагмент – диастола [9], которая является стационарным стохастическим сигналом. Диагностика болезни заключается в отнесении наблюдаемого стохастического сигнала (фрагмента диастолы) в один из $L = 2$ классов: Ω_1 – сигналы кровотока здоровых людей; Ω_2 – сигналы кровотока людей, имеющих коронарную ишемическую болезнь сердца. В [8] в качестве имитационной модели диастолы предложена модель авторегрессии второго порядка, наблюдаемой в шуме:

$$x_t^{(l)} = \eta_t^{(l)} + u_t^{(l)}, \quad t \in \mathfrak{T}, \quad l \in \{1, 2\}; \quad (15)$$

$$\eta_t^{(l)} = \theta_1^{(l)} \eta_{t-1}^{(l)} + \theta_2^{(l)} \eta_{t-2}^{(l)} + \xi_t^{(l)}, \quad (16)$$

где $\{\xi_t^{(l)}\}$ – независимые одинаково распределенные случайные величины с нормальным распределением вероятностей $N(0, D_l^2)$; $\{u_t^{(l)}\}$ – независимые одинаково распределенные случайные величины с нормальным распределением вероятностей $N(0, B_l^2)$, описывающие «шум»

при регистрации фрагмента диастолы $\{\eta_i^{(l)}\}$. Классы Ω_1, Ω_2 различаются параметрами модели (15), (16):

$$\Omega_1 : \theta^{(1)} = (1, -0.6)', D_1^2 = 15;$$

$$\Omega_2 : \theta^{(2)} = (1.2, -0.7)', D_2^2 = 35.$$

Эти параметры подбирались так, чтобы имитационная модель была адекватна реальным данным [9]. Значения дисперсий $\{B_i^2\}$ выбирались таким образом, чтобы отношение шум/сигнал принимало значения из следующего множества:

$$Q = \frac{B_i}{\sqrt{D_i}} \in \{0, 0.1, 0.3, 0.5\}.$$

Набор чисел $G = \{g_1, \dots, g_{N-1}\}$, задающий дискретизацию исходного стохастического сигнала, был выбран так, чтобы вероятности попадания случайных величин $\{x_i^{(l)}\}$ в каждый из N промежутков были равными. Для описания получаемых дискретных сигналов использовалась модель (5), (6) со следующими значениями параметров: $N \in \{4, 8\}$, $s \in \{1, 2\}$. На рис. 1 приведена реализация исходного стохастического сигнала $\{x_i^{(l)}\}$ и реализация преобразованного согласно (4) дискретного временного ряда $\{y_i^{(l)}\}$ с $N = 8$ состояниями при уровне шума $Q = 0$.

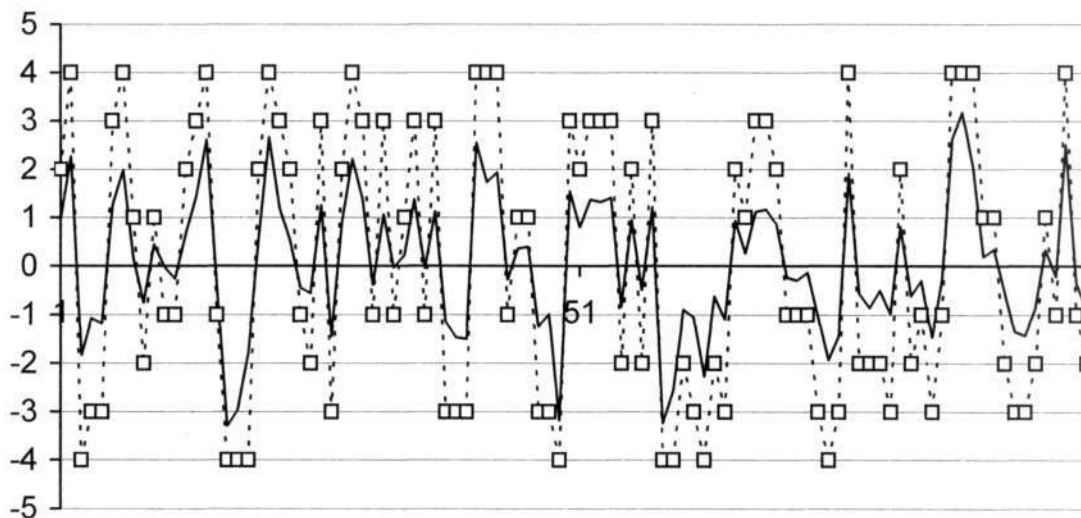


Рис. 1. Пример исходного (непрерывная линия) и полученного дискретного (пунктирная линия) сигналов кровотока здоровых людей (класс Ω_1)

Для анализа эффективности диагностики коронарной ишемической болезни с использованием подстановочного РП (14) был проведен вычислительный эксперимент по оцениванию условных вероятностей ошибочной классификации $r_1 = \mathbf{P}\{d = 2 | v = 1\}$ и $r_2 = \mathbf{P}\{d = 1 | v = 2\}$. Длина обучающих выборок составила $n_1 = n_2 = 100000$ (200 фрагментов диастолы длительностью 500 каждая). Длина классифицируемой реализации равна $n_0 = 500$ (один фрагмент диастолы). Оценивание вероятностей r_1, r_2 производилось по моделируемым сигналам 1 000 пациентов из класса Ω_1 и 1 000 пациентов из класса Ω_2 при различных уровнях шума Q . Значения оценок вероятности ошибочных классификаций приведены в табл. 1. С увеличением длины классифицируемой реализации n_0 оценки вероятностей r_1, r_2

стремились к 0. По результатам вычислительного эксперимента можно отметить высокую эффективность предложенного подхода применительно к диагностике коронарной ишемической болезни.

Таблица 1

Оценки условных вероятностей r_1, r_2 ошибочной классификации при различных значениях параметров модели цепи Маркова

N, s	Q			
	0	0.1	0.3	0.5
$N = 4, s = 1$	0,036	0,034	0,074	0,16
	0,021	0,015	0,048	0,128
$N = 4, s = 2$	0,029	0,048	0,089	0,163
	0,04	0,033	0,064	0,084
$N = 8, s = 1$	0,023	0,025	0,072	0,146
	0,011	0,019	0,037	0,097

Заключение

В работе для решения задачи распознавания стохастических сигналов предложено переходить от исходного непрерывного сигнала к дискретному сигналу, описываемому цепью Маркова s -го порядка, и задачу распознавания исходного сигнала решать в пространстве параметров цепей Маркова. Это позволяет свести задачи дискриминантного анализа различных классов стохастических сигналов к одной задаче - задаче дискриминантного анализа цепей Маркова s -го порядка.

Список литературы

1. Андерсон Т. Введение в многомерный статистический анализ. – М.: Физ.-мат. лит., 1963. – 500 с.
2. Харин Ю.С. Робастность в статистическом распознавании образов. – Мн.: БГУ, 1992. – 232 с.
3. Фукунага К. Введение в статистическую теорию распознавания образов. – М.: Наука, 1979. – 367 с.
4. McLachlan G.J. Discriminant Analysis and Statistical Pattern Recognition. – N.Y.: Wiley, 1992. – 526 p.
5. Репин В.Г., Тартаковский Г.П. Статистический синтез при априорной неопределённости и адаптация информационных систем. – М.: Сов. радио, 1977. – 432 с.
6. Сенин А.Г. Распознавание случайных сигналов. – М.: Наука, 1974. – 75 с.
7. Харин Ю.С., Костевич А.Л. Распознавание дискретных сигналов, описываемых цепями Маркова // Цифровая обработка сигналов и ее применения: Докл. 1 Междунар. конф. - М., 1998. - Т. 2. – С. 192-199.
8. Abramovich M.S., Zhuk E.E. Segmentation and Classification of the Blood Flow Signals // Proceedings of the VII International Conference «Pattern Recognition and Information Processing'2003». - Minsk, 2003. – P. 57–63.
9. Akay M. Wavelet Application in Medicine // IEEE Spectrum. – 1997. – Vol. 34. – № 5. – P. 50-56.
10. Kazakos D. The Bhattacharyya Distance and Detection Between Markov Chains // IEEE Transactions on Information Theory. – 1978. – Vol. IT-24. – № 6. – P. 747-754.
11. Koopmans L.H. Asymptotic Rate of Discrimination for Markov Processes // The Annals of Mathematical Statistics. – 1960. – Vol. 31. – P. 982-994.
12. Basawa I.V., Rao B.L.C. Statistical Inference for Stochastic Processes. – N.Y.: Academic Press, 1980. – 438 p.

13. Billingsley P. Statistical Methods in Markov Chains // The Annals of Mathematical Statistics. – 1961. – Vol. 32. – P. 12-40.

Поступила 03.02.04

*Белорусский государственный университет,
Минск, пр. Ф. Скорины, 4
e-mail: kharin@bsu.by*

Yu.S. Kharin, A.L. Kostevich

RECOGNITION OF STOCHASTIC SIGNALS USING MARKOV CHAINS OF ORDER S

For discriminant analysis of a stochastic signal, we propose to change the signal into a discrete signal described by a Markov chain of order s ; the problem of the initial signal recognition is considered in a space of the Markov chain parameters. The Bayesian decision rule and plug-in Bayesian decision rule for Markov chain classification are constructed. The misclassification probability of the decision rules under an asymptotic of contiguous classes is found. The proposed approach is successfully applied for diagnostics of the coronary ischemic disease by recognition of biomedicine signals.