УДК: 519.7

© 2021 г. **А.В. Лапко**^{1,3}, д-р техн. наук,

В.А. Лапко^{1,3}, д-р техн. наук,

С.Т. Им^{2,3}, канд. техн. наук,

А.В. Бахтина³,

В.Л. Авдеенок³,

В.П. Тубольцев³,

¹(Институт вычислительного моделирования СО РАН, Красноярск),

²(Институт леса им. В.Н. Сукачева СО РАН, Красноярск),

³(Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнева, Красноярск)

СИНТЕЗ И АНАЛИЗ СТАТИСТИЧЕСКОЙ МОДЕЛИ ДИНАМИКИ ВРЕМЕННЫХ СИСТЕМ С ДИСКРЕТНЫМ КОНТРОЛЕМ В УСЛОВИЯХ АПРИОРНОЙ НЕОПРЕДЕЛЕННОСТИ*

Предлагается статистическая модель временной системы с дискретным контролем, что является одним из актуальных направлений современной теории принятия решений. Подобные условия характерны при обработке информации дистанционного зондирования природных объектов. Перспективное направление преодоления возникающих проблем состоит в использовании структурных методов анализа статистических данных, когда под состоянием системы понимается множество значений случайных величин, обладающих похожими свойствами. Подобное определение позволяет упростить задачу оценивания состояний исследуемой системы с использованием непараметрических алгоритмов распознавания образов, открывающих возможность прогнозирования их временной динамики на основе преобразования множеств случайных величин. Проводится исследование свойств статистической модели временной динамики рассматриваемых объектов. В качестве показателя эффективности использу-

* Исследование выполнено при финансовой поддержке РФФИ, правительства Красноярского края и Красноярского краевого фонда науки в рамках научного проекта № 20-41-240001.

ется вероятность смены макросостояний системы из начального интервала времени в конечный. Исследуются асимптотические свойства указанного показателя эффективности и сопоставляются их результаты с предельной теоремой Маркова для цепных зависимостей.

Ключевые слова: временная система с дискретным контролем, структурный анализ данных, непараметрические алгоритмы распознавания образов, траектория состояний временной системы, асимптотические свойства, оценки вероятности смены состояний временной системы, предельная теорема Маркова для цепных зависимостей.

DOI: 10.22250/isu.2021.68.56-70

Введение

Исследование временных систем с дискретным контролем в условиях априорной неопределенности является одним из актуальных направлений современной теории принятия решений. К таким системам относятся, например, природные объекты, временная динамика которых зависит от климатических факторов [1 – 3]. При этом возникают задачи обнаружения закономерностей смены состояний объектов исследования в смежные моменты времени, определения динамики состояний объектов исследования в дискретные интервалы времени их контроля и исследования асимптотических свойств статистической модели изучаемой временной системы. Эти задачи составляют содержание структурного анализа данных. Категория «структурный анализ данных» была введена профессором А.А. Дорофеюком [4]. Под структурным анализом данных понимается не только обнаружение компактных групп наблюдений (классов), но и их преобразование при исследовании динамики объектов в дискретные интервалы времени в условиях априорной неопределенности. Это определение позволяет упростить задачу оценивания состояний исследуемых объектов с использованием непараметрических алгоритмов автоматической классификации и распознавания образов, открывающих возможность прогнозирования их временной динамики на основе преобразования обнаруженных множеств случайных величин.

Основная идея структурного прогнозирования состоит в декомпозиции области значений переменных исследуемых объектов с использованием алгоритмов автоматической классификации. Число классов выбирается с помощью человеко-машинной процедуры. На основе алгоритмов распознавания образов определяются закономерности переходов между обнаруженными классами в смежные моменты времени. Подобная процедура выполняется для всех моментов времени наблюдения за объектом. В качестве прогнозной модели используется марковская цепь, которая определяется матрицей пере-

ходных вероятностей [5].

Временные процессы с дискретным контролем в условиях априорной неопределенности ранее рассматривались при исследовании медико-биологических объектов. Предполагалось, что состояние временного процесса определяется его состоянием в предыдущий момент времени и внешними воздействиями на него в интервале времени между ними [1-2]. Под состоянием временной системы в конкретный интервал времени понимается область значений ее переменных, соответствующих одномодальным фрагментам плотности вероятности. Эти состояния формируются с использованием алгоритмов автоматической классификации. Взаимосвязи между состояниями в смежные моменты времени определяются в пространстве внешних воздействий на основе методов распознавания образов.

Цель данной статьи состоит в развитии и обобщении методики синтеза и анализа статистической модели временной системы с дискретным контролем в предположении, что ее состояние априори известно.

Постановка задачи

Исследуются объекты, состояния которых контролируются в дискретные моменты времени t и определяются преобразованием

$$y(t) = \varphi_t(y(t-1), x(t)), \quad t = \overline{1, M}. \quad (1)$$

Состояние объекта $y(t)$ в момент времени t зависит от его предыдущего состояния $y(t-1)$ и воздействия $x(t)$ на объект в интервале $(t-1, t)$. Вид преобразования (1) является нестационарным и неизвестным. Количество состояний в момент времени t является конечным

$$y(t) = (S_{jt}, j = \overline{1, N_t}), \quad t = \overline{1, M}.$$

Исходную информацию о динамике объекта (1) составляют статистические данные $V = (y^i(t-1), y^i(t), x^i(t), t = \overline{1, M}, i = \overline{1, n})$.

Необходимо по статистическим данным V в принятых условиях осуществить синтез модели преобразования (1), которое позволяет по информации о начальном состоянии объекта $y(1)$ и значениях внешних воздействий $x(t)$, $t = \overline{2, M}$ оценить траекторию смены состояний $y(t)$, $t = \overline{2, M}$ исследуемого объекта.

Методика синтеза модели объекта исследования

Предлагаемая методика предполагает выполнение следующих действий:

1. Отметим, что закономерность смены состояний исследуемого объ-

екта из начального $y'(t-1)$ в одно из состояний множества $y(t) = (S_{jt}, j = \overline{1, N_t})$ при воздействии $x(t)$ определяется алгоритмом распознавания образов

$$m_{t-1,t}(y'(t-1), x(t)).$$

Для его восстановления будем использовать информацию $V_{t-1,t} = (y'(t-1), x^i(t), i \in I_{t-1})$ из исходных статистических данных V . Решающее правило $m_{t-1,t}(y'(t-1), x(t))$ позволяет при условии $y'(t-1) \in y(t-1)$ и воздействии $x(t)$ определить переход состояния системы из $y'(t-1)$ в одно из состояний $y'(t) \in y(t)$. Здесь I_{t-1} – множество номеров ситуаций из V , определяющих переход системы из состояния $y'(t-1)$ в одно из состояний $y(t)$.

2. Повторим этап 1 для всех состояний $y'(t-1) \in y(t-1)$, что позволит определить решающие правила $m_{t-1,t}(y'(t-1), x(t))$, $y'(t-1) \in y(t-1)$, используя исходную статистическую информацию V .

3. Этапы 1, 2 повторим для всех моментов времени $t = \overline{2, M}$.

4. Тогда временная динамика процесса изменения состояний объекта из исходного $y'(1) = S_{q1}$ при воздействиях $x(t)$, $t = \overline{2, M}$ представляется следующей логической схемой

$$(t=1) \xrightarrow{1} R_{qt}^{v,t+1}(x(t+1)) (t=t+1) (q=v) v \xrightarrow{1}, \quad (2)$$

где $R_{qt}^{v,t+1}(x(t+1))$ – алгоритм, определяющий переход объекта исследования из состояния S_{qt} в состояние $S_{v,t+1} \in y(t+1)$ под воздействием значений $x(t+1)$. При выполнении логического условия $v: (t \leq M-1)$ происходит переход по стрелке. Структура модели (2) представлена на рис. 1.

Рис. 1. Структура модели (2) временной системы с дискретным контролем.

Для сравнения различных вариантов построения статистической модели (2) необходимо определить показатель ее эффективности, который представляет собой интегральную характеристику степени адаптации исследуемой системы к выполнению поставленных перед нею задач. Показатель эффективности должен достаточно полно отражать основные особенности и свойства структуры модели, стохастический характер взаимодействия между ее элементами и внешней средой.

К подобному показателю эффективности относится, например, вероятность $P(S_{jt}, S_{qM})$ перехода исследуемой системы из состояния S_{jt} в S_{qM} :

$$P(S_{jt}, S_{qM}) = \sum_{v \in I_{t+1}} P_{jt}^{v, t+1} P(S_{v, t+1}, S_{qM}), \quad (3)$$

где $P_{jt}^{v, t+1}$ – условная вероятность перехода между состояниями S_{jt} , $S_{v, t+1}$ в смежные моменты времени.

Непараметрический алгоритм восстановления закономерностей смены состояний исследуемой системы в смежные интервалы времени

Закономерность перехода между состоянием S_{jt} и конечным множеством состояний S_{t+1} в смежные интервалы времени под воздействием $x(t+1)$ адекватно оценивается алгоритмом распознавания образов. Для его восстановления в условиях априорной неопределенности будем использовать непараметрический алгоритм распознавания образов. Исходная информация представляется статистическими данными

$$V_{jt}^{t+1} = (x^i(t+1), \sigma_{jt}^{t+1}(i), i \in I_{jt}),$$

где I_{jt} – множество номеров элементов выборки V , определяющих состояние S_{jt} . Значения $\sigma_{jt}^{t+1}(i)$ определяют информацию о переходе исследуемой системы под воздействием $x(t+1)$ из состояния S_{jt} в одно из состояний $y(t+1) = (S_{q, t+1}, q = \overline{1, N_{t+1}})$.

По выборке V_{jt}^{t+1} осуществить синтез решающего правила $R_{jt}^{t+1}(x(t+1))$ с использованием непараметрического алгоритма распознавания образов [6–10]

$$\begin{aligned} \bar{R}_{jt}^{t+1}(x(t+1)): S_{jt} \rightarrow S_{v, t+1}, \text{ если} \\ \bar{P}_{jt}^{v, t+1} \bar{p}_{jt}^{v, t+1}(x(t+1)) = \max_{q=1, N_{t+1}} \bar{P}_{jt}^{q, t+1} \bar{p}_{jt}^{q, t+1}(x(t+1)). \end{aligned} \quad (4)$$

В непараметрическом решающем правиле (4) значение

$\bar{P}_{jt}^{q,t+1} = n_{jt}^{q,t+1} / n_{jt}$ является оценкой априорной вероятности перехода системы из состояния S_{jt} в $S_{q,t+1}$; n_{jt} – количество элементов выборки V , определяющих состояние S_{jt} ; $n_{jt}^{q,t+1}$ – количество элементов выборки V_{jt} , для которых выполняется переход из состояния S_{jt} в $S_{q,t+1}$. Ядерная оценка условной плотности вероятности $\bar{p}_{jt}^{q,t+1}(x(t+1))$ типа Розенблатта – Парзена, характеризующая закономерность смены состояния системы $S_{jt} \rightarrow S_{q,t+1}$ при воздействии $x(t+1)$, представляется выражением [11 – 15]

$$\bar{p}_{jt}^{q,t+1}(x(t+1)) = \left(n_{jt}^{q,t+1} \prod_{r=1}^k c_{jt}^{t+1}(r) \right)^{-1} \sum_{i \in I_{jt}^{q,t+1}} \prod_{r=1}^k \Phi \left(\frac{x_r(t+1) - x_r^i(t+1)}{c_{jt}^{t+1}(r)} \right), \quad (5)$$

$$q = \overline{1, N_{t+1}}.$$

Здесь $I_{jt}^{q,t+1}$ – множество номеров элементов выборки V_{jt} , определяющих смену состояний $S_{jt} \rightarrow S_{q,t+1}$.

В оценке плотности вероятности (5) ядерные функции $\Phi(u_r)$ удовлетворяют условиям H :

$$0 \leq \Phi(u_r) < \infty, \quad \Phi(u_r) = \Phi(-u_r), \quad \int \Phi(u_r) du_r = 1, \quad \int u_r^2 \Phi(u_r) du_r = 1,$$

$$\int u_r^m \Phi(u_r) du_r < \infty, \quad 0 \leq m < \infty, \quad r = \overline{1, k},$$

а $c_{jt}^{t+1}(r)$ – последовательность коэффициентов размытости ядерных функций таких, что $c_{jt}^{t+1}(r) \rightarrow 0$, $n_{jt}^{q,t+1} \prod_{r=1}^k c_{jt}^{t+1}(r) \rightarrow \infty$ при $n_{jt}^{q,t+1} \rightarrow \infty$. Здесь и далее бесконечные пределы интегрирования опускаются.

Оптимизация непараметрического алгоритма (4), определяющего закономерность смены смежных по времени состояний исследуемой системы, по коэффициентам размытости ядерных функций $c_{jt}^{t+1}(r)$, $r = \overline{1, k}$ осуществляется в режиме «скользящего экзамена» из условия минимума оценки вероятности ошибки распознавания образов

$$\bar{\rho}_{jt}^{t+1} = \frac{1}{n_{jt}} \sum_{il \in I_{jt}} 1(\sigma_{jt}^{t+1}(il), \bar{\sigma}_{jt}^{t+1}(il)).$$

Здесь индикаторная функция

$$1(\sigma_{jt}^{t+1}(il), \bar{\sigma}_{jt}^{t+1}(il)) = \begin{cases} 0, & \text{если } \sigma_{jt}^{t+1}(il) = \bar{\sigma}_{jt}^{t+1}(il) \\ 1, & \text{если } \sigma_{jt}^{t+1}(il) \neq \bar{\sigma}_{jt}^{t+1}(il), \end{cases}$$

где $\bar{\sigma}_{jt}^{t+1}(il)$ – решение непараметрического алгоритма (4) о переходе систе-

мы из состояния S_{jt} в одно из состояний S_{t+1} под воздействием $x(t+1)$. При совпадении значений $\sigma_{jt}^{t+1}(il)$, $\bar{\sigma}_{jt}^{t+1}(il)$ ошибка в принятии решений непараметрическим алгоритмом (4) отсутствует. В процессе формирования решения $\bar{\sigma}_{jt}^{t+1}(il)$ ситуация $x^{il}(t+1)$ исключается из обучения в непараметрической статистике (5).

Для упрощения задачи оптимизации непараметрического алгоритма распознавания образов (4) имеется возможность полагать, что $c_{jt}^{t+1}(r) = c \bar{\delta}_{jt}(r)$. Оценка среднеквадратического отклонения $\bar{\delta}_{jt}(r)$ случайной величины $x_r(t+1)$ вычисляется по статистическим данным V_{jt} объема n_{jt} , которые определяют состояние S_{jt} . Данное условие основывается на утверждении: значение коэффициента размытости зависит от интервала значений случайной величины.

По аналогии осуществляется синтез непараметрических алгоритмов распознавания образов $\bar{R}_{jt}^{t+1}(x(t+1))$, $j = \overline{1, N_t}$, $t = \overline{1, M-1}$, формирующих структуру статистической модели временной системы с дискретным контролем (2).

Для оптимизации непараметрических алгоритмов распознавания образов в условиях больших объемов статистических данных целесообразно использовать процедуры быстрого выбора коэффициентов размытости ядерных функций [16 – 24].

Асимптотические свойства статистических оценок показателей эффективности моделей развивающихся систем

Основу статистической оценки $\bar{P}(S_{jt}, S_{qM})$ составляют статистики $\bar{P}_{jt}^{v,t+1}$, $v \in I_{t+1}$, $j \in I_t$, которые определяют последовательность действий при исследовании его асимптотических свойств. Сначала доказывается асимптотическая несмещенность и состоятельность оценки вероятности смены состояний системы $\bar{P}_{jt}^{v,t+1}$, проводится анализ их условий сходимости.

В процессе реализации намеченного плана будем пользоваться обозначениями, принятыми выше.

Пусть $N_{t+1}(jt)$ – количество состояний из множества S_{t+1} , в которые возможен переход из S_{jt} , $N_{t+1}(jt) \leq N_{t+1}$.

Обозначим через $A_{jt}^{t+1}(x(t+1)) = (p_{jt}^{q,t+1}(x(t+1)))^{(2)}$ – функционал от вто-

рой производной условной плотности вероятности $p_{jt}^{q,t+1}(x(t+1))$ по $x(t+1)$.

Убывающая с ростом объема выборки n последовательность J имеет скорость сходимости $\beta > 0$, если $n^\beta J(n) = 1$.

Лемма. Пусть:

1) плотность вероятности $p_{jt}^{q,t+1}(x(t+1))$ ограничена и непрерывна со всеми производными до порядка два включительно;

2) кроме того,

$$\int (A_{jt}^{t+1}(x(t+1)))^2 dx(t+1) \neq 0;$$

3) коэффициенты размытости ядерных функций оптимальны в смысле минимума асимптотических среднеквадратических отклонений между непараметрической оценкой уравнения разделяющей поверхности $\bar{f}_{jt}^{q,t+1}(x(t+1))$ и байесовской решающей функцией $f_{jt}^{q,t+1}(x(t+1))$;

4) ядерные функции $\Phi(u)$ удовлетворяют условиям H . Тогда при $n_{jt}^{q,t+1} = n_{jt}^\alpha : \alpha = 0,53, n \rightarrow \infty$:

а) $E(\bar{P}_{jt}^{q,t+1} - P_{jt}^{q,t+1}) \sim O(N_{t+1}(jt) n^{-0.18}),$

б) $E(\bar{P}_{jt}^{q,t+1} - P_{jt}^{q,t+1})^2 \sim O(n_{jt}^{-0.36}(1 + N_{t+1}^2(jt))),$

в) при дополнительном условии $N_{t+1}(jt) = n_{jt}^{k1}, k < 0,18,$
 $V(1) \sim O(n_{jt}^{-0.18+k1}), V(2) \sim O(n_{jt}^{-0.36+2k1}).$

Для доказательства леммы, которое основано на использовании свойств оценок типа $\bar{p}_{jt}^{q,t+1}(x(t+1))$, используем результаты работ [12, 15, 25, 26].

Полученные выводы позволяют перейти непосредственно к исследованию асимптотических свойств непараметрической оценки $\bar{P}(S_{jt}, S_{qM})$ показателя эффективности (3) исследуемой системы. Имеет место следующее утверждение.

Теорема. Пусть:

1) выполняются условия и утверждения леммы;

2) $n_{vh}^{\varepsilon,h+1} = n_{vh}^\alpha \rightarrow \infty, \alpha = 0,53, v \in I_h, \varepsilon \in I_{h+1}, h = \overline{t, M-1};$

3) количество уровней $(M-t)$ и число $N_h, h = \overline{t+1, M-1}$ состояний структуры системы ограничены.

Тогда:

1) статистика $\bar{P}(S_{jt}, S_{qM})$ является асимптотически несмещенной и состоятельной оценкой вероятности $P(S_{jt}, S_{qM})$ перехода системы в состояние S_{qM} из начального S_{jt} ;

2) при дополнительных предположениях

$$n_{vh} = n, \quad v \in I_h, \quad h = \overline{t, M-1},$$

$$N_{h+1}(vh) = n^{k1}, \quad v \in I_h, \quad h = \overline{t, M-2}, \quad N_h = n^{\gamma k1},$$

$$\gamma \geq 1, \quad h = \overline{t+1, M-1}, \quad m = M - t:$$

а) статистика $\bar{P}(S_{jt}, S_{qM})$ сходится к $P(S_{jt}, S_{qM})$ в среднеквадратическом смысле, если

$$\beta_1 > 2k1\gamma(1 - 2^{-m}),$$

где β_1 – скорость сходимости оценок вероятностей переходов между состояниями системы,

б) при $m \rightarrow \infty$ скорость сходимости $\bar{P}(S_{jt}, S_{qM})$ имеет предел

$$\beta_2 = (\beta_1 - 2k1\gamma)/(2(1 + \gamma)).$$

Доказательство. Производя тождественные преобразования, запишем математическое ожидание смещения $(\bar{P}(S_{jt}, S_{qM}) - P(S_{jt}, S_{qM}))$ с учетом выражения (3):

$$\begin{aligned} V(3) &= E(\bar{P}(S_{jt}, S_{qM}) - P(S_{jt}, S_{qM})) = \\ &= E \left\{ \sum_{v \in I_{t+1}} (\bar{P}_{jt}^{v, t+1} \bar{P}(S_{v, t+1}, S_{qM}) - P_{jt}^{v, t+1} P(S_{v, t+1}, S_{qM})) \right\} = \\ &= E \left\{ \sum_{v \in I_{t+1}} \bar{P}_{jt}^{v, t+1} (\bar{P}(S_{v, t+1}, S_{qM}) - P(S_{v, t+1}, S_{qM})) \right\} + \\ &+ E \left\{ \sum_{v \in I_{t+1}} P(S_{v, t+1}, S_{qM}) (\bar{P}_{jt}^{v, t+1} - P_{jt}^{v, t+1}) \right\}. \end{aligned}$$

Здесь E – знак математического ожидания.

Последовательно повторяя подобные выкладки, легко проверить, что

$$\begin{aligned} V(3) &= \sum_{v \in I_{t+1}} P(S_{v, t+1}, S_{qM}) E(\bar{P}_{jt}^{v, t+1} - P_{jt}^{v, t+1}) + \\ &+ E \left\{ \sum_{h=t+2}^{M-1} \sum_{\lambda_1 \in I_{t+1}} \dots \sum_{\lambda_h \in I_h} P(S_{\lambda_h}, S_{qM}) \bar{P}_{jt}^{\lambda, t+1} \dots \right. \\ &\left. \dots \bar{P}_{\lambda_{h-3}, h-3}^{\lambda_h, h-2} \bar{P}_{\lambda_{h-2}, h-2}^{\lambda_{h-1}, h-1} (\bar{P}_{\lambda_{h-1}, h-1}^{\lambda_h h} - P_{\lambda_{h-1}, h-1}^{\lambda_h h}) \right\} - \end{aligned}$$

$$- E \left\{ \sum_{\lambda_1 \in I_{t+1}} \dots \sum_{\lambda_{M-1} \in I_{M-1}} \bar{P}_{jt}^{\lambda_1, t+1} \dots \bar{P}_{\lambda_{M-3}, M-3}^{\lambda_{M-2}, M-2} \times \right. \\ \left. \times \bar{P}_{\lambda_{M-2}, M-2}^{\lambda_{M-1}, M-1} \left(\bar{P}_{\lambda_{M-1}, M-1}^{qM} - P_{\lambda_{M-1}, M-1}^{qM} \right) \right\}.$$

При завершении доказательства асимптотической несмещенности $\bar{P}(S_{jt}, S_{qM})$ обозначим

$$P_{\lambda_{h-2}, h-2}^{h-1} \geq \max_{\lambda_{h-1} \in I_{h-1}} \bar{P}_{\lambda_{h-2}, h-2}^{\lambda_{h-1}, h-1}, \quad h = \overline{t+1, M},$$

а $N_h = [I_h]$ – количество состояний $S_{\lambda_h h}$ системы в момент времени h . Заме-

чая, что $\sum_{\lambda_{h-1} \in I_{h-1}} \bar{P}_{\lambda_{h-2}, h-2}^{\lambda_{h-1}, h-1} = 1$, $h = \overline{t+2, M}$, оценим сверху $V(3)$:

$$V(3) \leq \sum_{v \in I_{t+1}} P(S_{v, t+1}, S_{qM}) E(\bar{P}_{qt}^{v, t+1} - P_{qt}^{v, t+1}) + \\ + \sum_{h=t+2}^{M-1} N_{h-1} P_{\lambda_{h-2}, h-2}^{\lambda_{h-1}, h-1} \sum_{\lambda_h \in I_h} P(S_{\lambda_h}, S_{iM}) E\{\bar{P}_{\lambda_{h-1}, h-1}^{\lambda_h h} - \\ - P_{\lambda_{h-1}, h-1}^{\lambda_h h}\} + P_{\lambda_{M-2}, M-2}^{M-1} \sum_{\lambda_{M-1} \in I_{M-1}} E\{\bar{P}_{\lambda_{M-1}, M-1}^{qM} - P_{\lambda_{M-1}, M-1}^{qM}\}. \quad (6)$$

Отсюда, согласно первому утверждению леммы и в силу ограниченности значений N_{h-1} , $h = \overline{t+2, M-1}$, следует справедливость положения об асимптотической несмещенности статистической оценки показателя эффективности $\bar{P}(S_{jt}, S_{qM})$. Используя приведенную выше технику преобразований и неравенство Коши – Буняковского, оценим выражение

$$V(4) = E(\bar{P}(S_{jt}, S_{qM}) - P(S_{jt}, S_{qM}))^2 = \\ = E \left\{ \sum_{v \in I_{t+1}} \left[\bar{P}_{jt}^{v, t+1} (\bar{P}(S_{v, t+1}, S_{qM}) - P(S_{v, t+1}, S_{qM})) + \right. \right. \\ \left. \left. + P(S_{v, t+1}, S_{qM}) (\bar{P}_{jt}^{v, t+1} - P_{jt}^{v, t+1}) \right] \right\}^2 \leq \quad (7) \\ \leq \left[\sum_{v \in I_{t+1}} E(\bar{P}_{jt}^{v, t+1})^2 \sum_{v \in I_{t+1}} E(\bar{P}(S_{v, t+1}, S_{qM}) - P(S_{v, t+1}, S_{qM}))^2 + \right. \\ \left. + \sum_{v \in I_{t+1}} P^2(S_{v, t+1}, S_{qM}) \sum_{v \in I_{t+1}} E(\bar{P}_{jt}^{v, t+1} - P_{jt}^{v, t+1})^2 \right]^2,$$

где среднеквадратические отклонения $E(\bar{P}(S_{v, h+1}, S_{qM}) - P(S_{v, h+1}, S_{qM}))^2$, получаемые в процессе вычислений, допускают представление, аналогичное

(7), относительно вероятностных характеристик

$$\bar{P}_{v,h+1}^{v',h+2}, P_{v,h+1}^{v',h+2}, \bar{P}(S_{v',h+2}, S_{qM}), P(S_{v',h+2}, S_{qM}),$$

$v' \in I_{h+2}, v \in I_{h+1}, h = \overline{t, M-2}$, закономерностей перехода между множеством состояний S_{h+1}, S_{h+2} и состоянием S_{qM} . Таким образом, может быть сформировано ограничение сверху в виде нелинейного функционала $\bar{V}$ (4) от среднеквадратических отклонений

$$E(\bar{P}_{vh}^{v',h+1} - P_{vh}^{v',h+1})^2, v \in I_h, v' \in I_{h+1}, h = \overline{t, M-1}, v = j,$$

при $h = t, v' = 1$ и $h = M-1$.

Отсюда, предполагая ограниченность значений $(M-t), N_h, h = \overline{t+1, M-1}$ и учитывая утверждение 2) леммы, следует сходимость в среднеквадратическом статистики $\bar{P}(S_{jt}, S_{qM})$ и справедливость первой части теоремы.

С целью удобства дальнейшего анализа $\bar{V}$ (4) без существенной потери общности введем дополнительные предположения. Пусть объемы статистических выборок, определяющих состояния $S_{vh}, n_{vh} = n, v \in I_h, h = \overline{t, M-1}$, причем количество возможных переходов из S_{vh} в множество состояний S_{h+1} фиксировано значением $N_{h+1}(vh) = \bar{N}, v \in I_h, h = \overline{t, M-2}$.

Далее условимся, что

$$\bar{P} = \max_{h, h1} \left(\sum_{v \in I_{h+1}} P^2(S_{v, h+1}, S_{qM}), \sum_{v \in I_{h+1}} E(\bar{P}_{v'h1}^{v, h1+1})^2 < 1, v' \in I_{h1} \right),$$

$$h = \overline{t, M-2}, h1 = \overline{t, M-1}.$$

Асимптотические свойства аддитивного функционала $\bar{V}$ (4) определяются в основном слагаемым $\bar{v}$ (4) с максимальной степенью нелинейности и имеющим наименьшую скорость сходимости.

С учетом принятых условий нетрудно получить компактную запись $\bar{v}$ (4) для часто встречающейся в практике ситуации $N_h = N$:

$$\bar{v}(4) = 4^m n^{-2^m} \beta^* (1 + \bar{N}^2)^{2^m} d^{2^m} \bar{P}^{\sum_{h=1}^m 2^h} N^{\sum_{h=1}^m 2^h},$$

где $m = M-1$ – количество уровней структуры системы между t -м и M -м.

Для $\bar{P} < 1$ значение $4^m 2^{2^m} \bar{P}^{\sum_{h=1}^m 2^h}$ ограничено, что вполне обосновано при принятых предположениях. Поэтому исследование асимптотических свойств $\bar{v}$ (4) проведем путем анализа зависимости от n произведения

$$n^{-2^m \beta^*} (1 + \bar{N}^2)^{2^m} N^{\sum_{h=1}^m 2^h}.$$

Положим: $\bar{N} = n^{k_1}$ и $N = n^{\alpha_1}$, $\alpha_1 = \gamma k_1$, $\gamma \neq 1$. Тогда условие сходимости $\bar{v}(4)$ к нулю запишется в виде:

при $n \rightarrow \infty$

$$k_1 < \beta^* / \left(2 + \sum_{h=1}^m 2^{h-m} \right),$$

при использовании формулы геометрической прогрессии

$$k_1 < \beta^* / 2(1 + \gamma(1 - 2^{-m})).$$

Таким образом, скорость сходимости непараметрической оценки $\bar{P}(S_{jt}, S_{qM})$ определяется соотношением

$$\beta_2 = (\beta_1 - 2k_1\gamma(1 - 2^{-m})) / 2(1 + \gamma(1 - 2^{-m})),$$

где $\beta_1 = \beta^* - 2k_1$ – скорость сходимости статистики $\bar{P}_{vh}^{\varepsilon, h+1}$, $v \in I_h$, $\varepsilon \in I_{h+1}$, характеризующей вероятности переходов между состояниями смежных уровней структуры системы.

Как и следовало ожидать, скорость сходимости $\bar{P}(S_{jt}, S_{qM})$ ниже, чем β_1 . С ростом числа уровней структуры m значение β_2 уменьшается и при $m \rightarrow \infty$ имеет предел

$$\bar{\beta}_2 = (\beta_1 - 2k_1\gamma) / 2(1 + \gamma).$$

Данный результат объясняется утверждением известной теоремы для однородных марковских цепей

$$\lim_{M \rightarrow \infty} P(S_{jt}, S_{qM}) = P_{qM}.$$

Если система, находясь в состоянии S_{vh} , $v \in I_h$ допускает переход во все состояния $S_{\varepsilon, h+1}$, $\varepsilon \in I_{h+1}$, $h = \overline{t, M-1}$ смежного уровня, то $\gamma = 1$. В данном случае

$$\bar{\beta}_2 = (\beta_1 - 2k) / 4,$$

причем с увеличением γ значение $\bar{\beta}_2$ уменьшается, т.е. скорость сходимости $\bar{P}(S_{jt}, S_{qM})$ убывает с ростом количества состояний N на каждом уровне структуры системы. С другой стороны, скорость сходимости β_2 при прочих одинаковых условиях возрастает по мере уменьшения количества возможных переходов между состояниями системы.

Утверждение б) теоремы согласуется с теоремой А.А. Маркова о предельных вероятностях для цепных зависимостей, в которой при заданной

матрице переходных вероятностей показано, что $\bar{P}(S_{jt}, S_{qM})$ при $M \rightarrow \infty$ стремится к априорной вероятности $P(S_{qM})$ [27]. В соответствии с теоремой существуют условия, когда скорость сходимости оценки $\bar{P}(S_{jt}, S_{qM})$ вероятности $P(S_{jt}, S_{qM})$ стремится к пределу. Данный вывод согласуется с предельной теоремой Маркова для цепных зависимостей и подтверждает корректность приведенного в работе доказательства.

Заключение

Статистическая модель временной системы с дискретным контролем ее переменных в условиях априорной неопределенности представляется множеством непараметрических алгоритмов распознавания образов в пространстве внешних воздействий. Ее применение позволяет прогнозировать траектории смены состояния системы в дискретные интервалы времени, которые соответствуют заданным последовательностям внешних воздействий. Основным показателем эффективности статистической модели изучаемой системы является вероятность смены ее состояний, принадлежащих различным интервалам времени. Этот показатель зависит от количества дискретных интервалов времени, числа состояний для каждого интервала времени, свойств непараметрических алгоритмов распознавания образов, определяющих закономерности переходов между состояниями в смежные интервалы времени, оценок априорных вероятностей смены состояний в смежные интервалы времени наблюдений за исследуемой системой. Оценка вероятности смены состояний системы, принадлежащих начальному и конечному интервалам времени, является асимптотически несмещенной и состоятельной. Причем с ростом количества интервалов наблюдений за системой скорость сходимости оценки вероятности стремится к пределу. Данный результат согласуется с предельной теоремой Маркова для цепных зависимостей.

Предложенные методические и алгоритмические средства открывают возможность создания программных средств синтеза и анализа непараметрических моделей временных систем с дискретным контролем их переменных в условиях статистических данных большого объема с элементами искусственного интеллекта, что создает основу для их применения при обработке данных дистанционного зондирования природных объектов, исследовании временной динамики медико-биологических и социально-экономических систем.

ЛИТЕРАТУРА

1. Лапко А.В., Новиков О.М., Поликарпов Л.С. Статистические методы моделирования и принятия решений в развивающихся медико-биологических системах. – Новосибирск: Наука, 1991.
2. А.В. Лапко, Л.С. Поликарпов, В.Т. Манчук и др. Автоматизация научных исследований в медицине. – Новосибирск: Наука, 1996.
3. Лапко А.В., Лапко В.А. Статистические модели пространственно распределенных временных процессов // Информатика и системы управления. – 2016. – Т. 50, № 4. – С. 69-75. DOI: 10.22250/isu.2016.50.69-75.
4. Дорофеюк А.А. Методология экспертно-классификационного анализа в задачах управления и обработки сложноорганизованных данных (история и перспективы развития) // Проблемы управления. – 2009. – №3.1. – С. 19-28.
5. Гольдовская М.Д., Дорофеюк Ю.А., Киселева Н.Е. Методы структурного анализа в прикладных задачах исследования временных рядов // Проблемы управления. – 2013. – № 3. – С. 33-40.
6. Лапко А.В., Лапко В.А. Анализ непараметрических алгоритмов распознавания образов в условиях пропуска данных // Автометрия. – 2008. – Т. 44, № 3. – С. 65-74.
7. Лапко А.В., Лапко В.А. Непараметрические алгоритмы распознавания образов в задаче проверки статистической гипотезы о тождественности двух законов распределения случайных величин // Автометрия. – 2010. – Т. 46, №6. – С. 47-53.
8. Лапко А.В., Лапко В.А., Егорочкин И.А. Непараметрические оценки смеси плотностей вероятности и их применение в задаче распознавания образов // Системы управления и информационные технологии. – 2009. – Т. 35, № 1. – С. 60-64.
9. Лапко А.В., Лапко В.А. Коллектив непараметрических решающих функций в двухальтернативной задаче распознавания образов // Системы управления и информационные технологии. – 2009. – Т. 37, № 3.1. – С. 156-160.
10. Лапко А.В., Лапко В.А. Непараметрическая оценка уравнения разделяющей поверхности в условиях больших выборок и ее свойства // Системы управления и информационные технологии. – 2010. – Т. 39, № 1.2. – С. 300-304.
11. Parzen E. On estimation of a probability density function and mode // Annals of Mathematical Statistics. – 1962. – Vol. 33, No3. – P. 1065-1076. DOI: 10.1214/aoms/1177704472.
12. Епанечников В.А. Непараметрическая оценка многомерной плотности вероятности // Теория вероятности и ее применения. – 1969. – Т.14, №1. – С. 156-161.
13. Лапко А.В., Лапко В.А. Синтез структуры смеси непараметрических оценок плотности вероятности многомерной случайной величины // Системы управления и информационные технологии. – 2011. – Т.43, № 1. – С. 12-15.
14. Лапко А.В., Лапко В.А. Регрессионная оценка плотности вероятности и ее свойства // Системы управления и информационные технологии. – 2012. – Т. 49, № 3.1. – С. 152-156.
15. Лапко А.В., Лапко В.А. Свойства непараметрической оценки плотности вероятности многомерных случайных величин в условиях больших выборок // Информатика и системы управления. – 2012. – Т. 32, № 2. – С. 121-126.
16. Silverman B.W. Density estimation for statistics and data analysis. – London: Chapman & Hall, 1986.

17. *Sheather S., Jones M.* A reliable data-based bandwidth selection method for kernel density estimation // *Journal of Royal Statistical Society Series B.* – 1991. – Vol. 53, No. 3. – P. 683-690. DOI: 10.1111/j.2517-6161.1991.tb01857.x.
18. *Sheather S.J.* Density Estimation // *Statistical Science.* – 2004. – Vol. 19, No. 4. – P. 588-597. DOI: 10.1214/088342304000000297.
19. *Terrell G.R., Scott D.W.* Oversmoothed nonparametric density estimates // *Journal of the American Statistical Association.* – 1985. – Vol. 80. – P. 209-214.
20. *Jones M.C., Marron J.S., Sheather S.J.* A brief survey of bandwidth selection for density estimation // *Journal of the American Statistical Association.* – 1996. – Vol. 91. – P. 401-407.
21. *Scott D.W.* Multivariate Density Estimation: Theory, Practice, and Visualization. – New Jersey: John Wiley & Sons, 2015.
22. *Ланко А.В., Ланко В.А.* Модифицированный алгоритм быстрого выбора коэффициентов размытости ядерных оценок многомерных плотностей вероятностей // *Измерительная техника.* – 2020. – №11. – С. 9-13. DOI: 10.32446/0368-1025it.2020-11-9-13.
23. *Ланко А.В., Ланко В.А.* Оценивание интеграла от квадрата производных симметричных плотностей вероятностей одномерных случайных величин // *Метрология.* – 2020. – №1. – С. 15-27. DOI: 10.32446/0132-4713.2020-1-15-27.
24. *Ланко А.В., Ланко В.А.* Оценивание нелинейного функционала от плотности вероятности при оптимизации непараметрических решающих функций // *Измерительная техника.* – 2021. – №1. – С. 14-20. DOI: 10.32446/0368-1025it.2021-1-14-20.
25. *Ланко А.В., Ланко В.А.* Многоуровневые непараметрические системы принятия решений. – Новосибирск: Наука, 1997.
26. *Ланко А.В., Ланко В.А.* Непараметрические методики анализа множеств случайных величин // *Автометрия.* – 2003. – Т. 39, № 1. – С. 54-61.
27. *Гнеденко Б.В.* Курс теории вероятностей. – М.: Наука, 1965.

E-mail:

Ланко Александр Васильевич – lapko@ict.krasn.ru;

Ланко Василий Александрович – lapko@ict.krasn.ru;

Им Сергей Тхекдеевич – stim@ksc.trans.ru;

Бахтина Анна Владимировна – anna-denisjuk@yandex.ru;

Авдеенок Валерий Леонидович – avdeyonok@gmail.com;

Тубольцев Виталий Павлович – vitalya.98@mail.ru.