

МЕТОД СКОЛЬЗЯЩИХ ОКОН В ЗАДАЧАХ ОБНАРУЖЕНИЯ ТОЧЕК ВЫБЕГА И ДИАГНОСТИКИ

В.И.Котюков, А.А.Наумов

Сибирский государственный университет путей сообщения,
Новосибирск, Россия, тел. (383-2) 287601
Новосибирский государственный технический университет,
Новосибирск, Россия, e-mail: naumov@fb.nstu.ru

В работе предлагаются вниманию результаты исследований в области разработки новых эффективных методов для обнаружения моментов времени, в которые начинаются те или иные изменения в системе (или так называемых точек выбега) [1]-[4].

1. Постановка задачи.

Пусть

$$S(t) = f(t) + \varepsilon(t), \quad (1)$$

где $S: R \rightarrow R$ сигнал, $f(t)$ – нелинейная функция, $f: R \rightarrow R$, t – время, $\varepsilon: R \rightarrow R$ шумовой сигнал. Задано множество характерных элементов диагнозов:

$$H^e = \{ \eta_i^e(t) \}, \quad i = 1, 2, \dots, r^e \quad (2)$$

Проблема состоит в следующем. Мы должны обнаружить элементы множества H^e в сигнале $S(t)$ и, тем самым, поставить диагноз относительно текущего состояния системы или окружающей ее среды. В этой работе мы рассматриваем методы с небольшим временем обнаружения, и поэтому соответствующие алгоритмы можно использовать в системах, работающих в реальном масштабе времени. Такие задачи возникают в различных областях: медицины, экономики, физики и т. д. (см., например, [5]-[12]).

Типичные представители множества H^e для диагностических задач в кардиологии – это фрагменты QRST-, QRS-, St-долей в ECG-сигналах (см. [8]-[12]). В Табл. 1 представлена информация об этих фрагментах, которая может иметь практическую ценность. Формальное описание некоторых характеристических элементов множества H^e приведено в табл. 2.

Таблица 1.


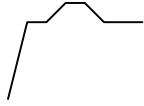

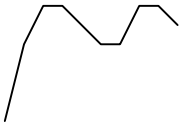
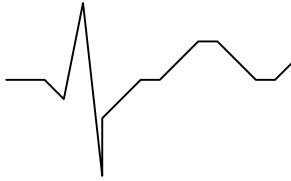
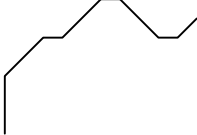
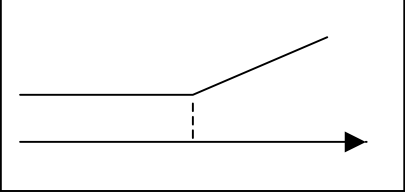
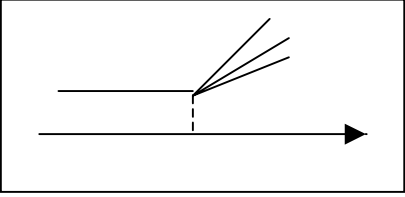
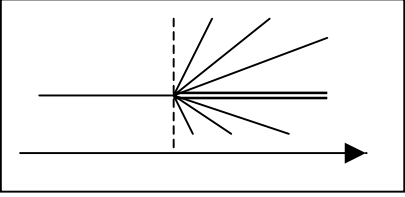
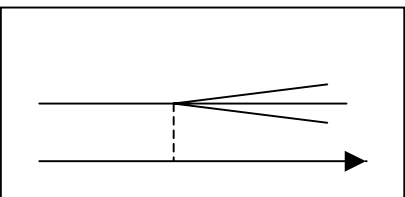
N	Фрагмент сигнала $S(t)$	Характеристика фрагмента	Характеристический элемент из множества H^e
1.		Нормальное значение T	$\eta_1^e(t)$: 
2.		Предельное отрицательное значение T	$\eta_2^e(t)$: 
3.		Высокое положительное значение T	$\eta_3^e(t)$: 

Таблица 2.

N	Графическое представление характеристического элемента	Формальное представление характеристического элемента
1		$\eta_4^e(t) = \begin{cases} \mu_0, & t \leq 0 \\ \mu_0 + \alpha t, & t > 0 \end{cases}$ $D\eta_4^e(t) = \{ \eta_4^e(t) \mid \alpha \neq 0 \}$

2		$\eta_5^e(t) = \begin{cases} \mu_0, t \leq 0 \\ \mu_0 + \alpha t, t > 0 \end{cases}$ $D\eta_5^e(t) = \{ \eta_5^e(t) \mid \alpha_1 \leq \alpha \leq \alpha_2 \}$
3		$\eta_6^e(t) = \begin{cases} \mu_0, t \leq 0 \\ \mu_0 + \alpha t, t > 0 \end{cases}$ $D\eta_6^e(t) = \{ \eta_6^e(t) \mid -\infty < \alpha \leq -\alpha_1, \alpha_2 < \alpha < +\infty, \alpha_1 > 0, \alpha_2 > 0 \}$
4		$\eta_7^e(t) = \begin{cases} \mu_0, t \leq 0 \\ \mu_0 + \alpha t, t > 0 \end{cases}$ $D\eta_7^e(t) = \{ \eta_7^e(t) \mid -\alpha_1 < \alpha \leq \alpha_1, \alpha_1 > 0 \}$

В табл. 2 $\mu_0, \mu_1, \alpha_1, \alpha_2, \alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$ - вещественные числа, $D\eta_i^e(t)$ - область допустимых значений для элементов $\eta_i^e(t)$.

2. Формальная постановка задачи обнаружения точек выбега.

Пусть множество H^f будет сформировано из элементов H^e , обнаруженных в сигнале $S(t)$. Будем писать $\eta_i^f(\hat{t}_j, \delta, \beta)$ для обозначения того факта, что элемент $\eta_i^e(t)$ был обнаружен в δ -окрестности момента времени \hat{t}_j , с вероятностью обнаружения β .

Тогда во множество H^f будут включены оценки элементов из H^e , которые, возможно, будут иметь одинаковые индексы i (элементов $\eta_i^e(t)$), но различные значения \hat{t}_j и различные оценки $\hat{\mu}_0, \hat{\mu}_1, \hat{\alpha}_1, \hat{\alpha}_2$. Т.е., фактически, мы будем иметь:

$$\eta_i^f(\hat{t}_j, \delta, \beta) = \hat{\eta}_i^e(t),$$

$$\hat{\eta}_i^e(t) \in D\eta_i^e(t), \Delta_{\hat{\eta}_i^e(t)}^\beta \subset D\eta_i^e(t), \quad \forall t \in (\hat{t}_j - \delta; \hat{t}_j + \delta),$$

где $\Delta_{\hat{\eta}_i^e(t)}^\beta$ доверительный интервал для оценки $\hat{\eta}_i^e(t)$ с доверительной вероятностью β . Оценка $\hat{t}_j = \hat{t}_{cp,j}$ - это оценка времени обнаружения характеристического элемента (или времени начала выбега, времени начала изменений в системе) и фактически она будет найдена с запаздыванием, т.е. в некоторый момент времени $\hat{t}_{cp,d,j} = \hat{t}_{cp,j} + \hat{t}_{d,j}$, где $\hat{t}_{d,j}$ - оценка времени запаздывания.

3. Метод двух окон для нахождения точек выбега.

Пусть, без умаления общности,

$$Y_n = S(t = n\Delta t) = \eta^e(n) + \varepsilon_n, \quad n \in \{-l, -l+1, \dots, 0, 1, 2, \dots, q\} = N_1, \quad (3)$$

$$t = n\Delta t,$$

есть дискретный ряд, где ε_n - реализация случайной переменной, для которой

$$E(\varepsilon_i) = 0, \quad D(\varepsilon_i) = \sigma^2 > 0, \quad E(Y_i Y_j) = \begin{cases} 0, & i \neq j; \\ \sigma^2, & i = j. \end{cases} \quad (4)$$

В этом случае $\hat{t}_{cp} = \hat{n}_{cp} = 0$. В области $n = 0$ строим два окна с наборами данных:

$$\bar{y}'_1 = (Y_{-n_1+1}, Y_{-n_1+2}, \dots, Y_0) \quad \text{и} \quad \bar{y}'_2 = (Y_1, Y_2, \dots, Y_{n_2})$$

для окон W_{1,n_1} и W_{2,n_2} соответственно (n_1 и n_2 - длины окон).

В окне W_{1,n_1} будем находить оценки параметров μ_0 и α_1 . В окне W_{2,n_2} - оценку α_2 . Для окна W_{1,n_1} определим матрицу:

$$\Phi = \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ (-n_1+1) & (-n_1+2) & \dots & 1 & 0 \end{pmatrix}'.$$

В этом случае МНК- оценки параметров μ_0 и α_1 и дисперсии этих оценок будут иметь вид:

$$\begin{pmatrix} \hat{\mu}_0 \\ \hat{\alpha}_1 \end{pmatrix} = (\Phi' \Phi)^{-1} \Phi' \bar{y}'_1 = F \bar{y}'_1, \quad D \begin{pmatrix} \hat{\mu}_0 \\ \hat{\alpha}_1 \end{pmatrix} = \sigma^2 (\Phi' \Phi)^{-1} = \sigma^2 G. \quad (5)$$

Отметим, что матрицы F и G следует вычислять только один раз (до начала работы алгоритма) и затем пользоваться ими при оценивании параметров и дисперсии. Этот прием известен в литературе как метод фиксированной системы координат, когда матрица оценивателя не изменяется при изменении временных отсчетов обрабатываемого временного ряда. МНК- оценка параметра α_2 для окна W_{2,n_2} и соответствующая дисперсия будут иметь вид:

$$\hat{\alpha}_2 = \frac{\sum_{i=1}^{n_2} i \tilde{Y}_i}{\sum_{i=1}^{n_2} i^2}; \quad (6)$$

$$\hat{\sigma}_{\hat{\alpha}_2} = \sigma / \sqrt{\sum_{i=1}^{n_2} i^2}, \quad \tilde{Y}_i = \tilde{Y}_i - \hat{\mu}_0, \quad i=1,2,\dots,n_2.$$

Примечание.

Если $\Delta t \neq 1$, то вместо (6) нужно использовать формулы:

$$\hat{\alpha}_2 = \frac{\sum_{i=1}^{n_2} i \Delta t \tilde{Y}_i / \sum_{i=1}^{n_2} (i \Delta t)^2}{\Delta t} = \frac{1}{\Delta t} \frac{\sum_{i=1}^{n_2} i \tilde{Y}_i / \sum_{i=1}^{n_2} i^2}{\Delta t} \quad \text{и} \quad \hat{\sigma}_{\hat{\alpha}_2} = \sigma / (\Delta t \sqrt{\sum_{i=1}^{n_2} i^2}).$$

Для окна W_{2,n_2} его длина n_2 изменяется в некотором диапазоне $[n_2^{\min}; n_2^{\max}]$ и то значение $n_2 = n_2^0$, при котором выполнится включение $\Delta_{\hat{\alpha}_2}^\beta \subset [\alpha_{21}; \alpha_{22}]$ и будет соответствовать моменту обнаружения элемента $\hat{\eta}^e(t)$ в сигнале $S(t)$ и, следовательно, $n_2^0 = \hat{t}_{d,j}$. В табл. 3 приведены оценки $\hat{t}_{d,j}$ для некоторых элементов $\hat{\eta}_i^e(t)$.

Таблица 3.

N	Характеристические элементы $\eta_i^e(t)$ и области $D\eta_i^e$	Оценка $\hat{t}_{d,j}$
1	$\eta_4^e(t), D\eta_4^e(t)$	$\hat{t}_{d,j} = \min\{n_2 \in N_2 = \{1, 2, \dots, n_2, \max\} \mid \left \sum_{i=1}^{n_2} i \tilde{Y}_i \right > k_{\alpha_2}^\beta \sigma \sqrt{\sum_{i=1}^{n_2} i^2}\}$
2	$\eta_5^e(t), D\eta_5^e(t)$	$\hat{t}_{d,j} = \min\{n_2 \in N_2 \mid \frac{\sum_{i=1}^{n_2} i \tilde{Y}_i / \sum_{i=1}^{n_2} i^2 - k_{\alpha_2}^\beta \sigma / \sqrt{\sum_{i=1}^{n_2} i^2} > \alpha_1$ $\text{и} \quad \frac{\sum_{i=1}^{n_2} i \tilde{Y}_i / \sum_{i=1}^{n_2} i^2 + k_{\alpha_2}^\beta \sigma / \sqrt{\sum_{i=1}^{n_2} i^2} < \alpha_2 \}$

3	$\eta_8^e(t), D\eta_8^e(t)$	$\hat{t}_{d,j} = \min\{n_2 \in N_2 \mid \sum_{i=1}^{n_2} i\tilde{Y}_i > k_{\alpha_2}^{\beta} \sigma \sqrt{\sum_{i=1}^{n_2} i^2} + (\hat{\alpha}_1 + k_{\alpha_1}^{\beta} \hat{\sigma}_{\hat{\alpha}_1}) \sum_{i=1}^{n_2} i^2, \text{ или}$ $\sum_{i=1}^{n_2} i\tilde{Y}_i < -k_{\alpha_2}^{\beta} \sigma \sqrt{\sum_{i=1}^{n_2} i^2} + (\hat{\alpha}_1 - k_{\alpha_1}^{\beta} \hat{\sigma}_{\hat{\alpha}_1}) \sum_{i=1}^{n_2} i^2\}$
4	$\eta_{10}^e(t), D\eta_{10}^e(t)$	$\hat{t}_{d,j} = \min\{n_2 \in N_2 \mid \alpha_{21} + k_{\alpha_2}^{\beta} \sigma / \sqrt{\sum_{i=1}^{n_2} i^2} <$ $< \frac{\sum_{i=1}^{n_2} i\tilde{Y}_i / \sum_{i=1}^{n_2} i^2 < \alpha_{22} - k_{\alpha_2}^{\beta} \sigma \sqrt{\sum_{i=1}^{n_2} i^2}\} .$

4. Результаты моделирования.

Исследуем характеристики метода скользящих окон для модели:

$$H^e = \{\eta_5^e(t) = \begin{cases} 0, t \leq 0 \\ \alpha t, t > 0 \end{cases}\};$$

$$D\eta^e(t) = \{\eta_5^e(t) \mid 0 < \alpha < +\infty\}, k_{\alpha} = 3.0, \alpha = 2.0, \varepsilon_i \propto \text{Rand}[-3\sigma; 3\sigma].$$

Параметр k_{α} определяет длину доверительного интервала

$\Delta_{\hat{\alpha}} = (\hat{\alpha} - k_{\alpha} \hat{\sigma}_{\hat{\alpha}}; \hat{\alpha} + k_{\alpha} \hat{\sigma}_{\hat{\alpha}})$. Результаты моделирования (на выборке из 1000 темных рядов) представлены в табл. 4.

Таблица 4.

σ	0.5	0.7	1.0	1.5	2.0	2.5	3.5	4.5	5.5	10.0
$\hat{t}_{d, \min}$	1	1	1	1	1	1	1	1	1	1
$\hat{t}_{d, \text{mean}}$	1.33	1.52	1.89	2.30	2.80	3.12	3.83	4.37	5.03	7.22
$\hat{t}_{d, \max}$	2	3	3	5	6	7	8	10	11	17

Здесь $\hat{t}_{d, \min}, \hat{t}_{d, \text{mean}}, \hat{t}_{d, \max}$ - минимальное, среднее и максимальные значения временной задержки. Можно заметить, что оценка $\hat{t}_{d, \max}$ ведет себя достаточно устойчиво по отношению к изменениям σ .

Таким образом, в работе предложены и исследованы новые алгоритмы, предназначенные для решения задач обнаружения точек выбега и основанные на методах скользящих окон и фиксированной системы координат.

Литература

1. *Basseville N., Nikiforov I.V.* Detection of Abrupt Changes, Prentice Hall, Englewood Cliffs, New Jersey, 1993.
2. *Daumer M., Bauer C., Lederer C.* Average Run Length and Mean Delay for Changepoint Detection : Robust Estimates for Threshold Alarms.// *Discussion Paper*, 95, TUM, 1997.
3. *Daumer M.* Adaptive Drifterkennung und Intelligente Alarmsysteme.// *Biomedizinische Technik*, 43, 1998, pp. 214-215.
4. *Daumer M., Bauer C., Lederer C.* Mean delay and average run length for changepoint detection: Robust estimates for threshold alarms.// In.: *Proceedings of the New Orleans Workshop on Statistical Modeling*, 1998.
5. *Chan K.S.* Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model.// *The Annals of Statistics*, 21, 1993, pp. 520-533.
6. *Chan K.S., Petruccioli J.D., Tong H., and Woolford S.W.* A Multiple Threshold AR(1) Model.// *Journal of Applied Probability*, 22, 1985, pp. 267-279.
7. *Tsay R.S.* Testing and Modeling Multivariate Threshold Models.// *Journal of the American Statistical Association*, 443, 1998, pp. 1188-1202.
8. *Bar F.W., Brugada P., Dassen W.* e.a. Prognostics Value of Q waves , R/S ration , loss of R wave voltage, ST-T segment abnormalities, electrical axis, low voltage and notching: Correlation of electrocardiogram and left ventriculogram.// *Journal Amer. coll. Cardiol.*, 4, 1984, pp. 17-27.
9. *Heinecker R.* EKG in Praxis und Klinik, Stuttgart, Georg Thieme Verlag, 1975.
10. *Heinicker R. Gonscha B.D.* EKG in Praxis und Klinik, Stuttgart, New York, 1992.
11. *Rapp H.J., Gasteiger P., Ackern K.* ST-trend- Analyse// *Anesthesist*, 42, 1993, pp. 652-656.
12. *Monitoring in Anästhesie und Intensivmedizin/ List W., Metzler H., Pasch T., Eds., Berlin, N.Y., Springer, 1995.*