2.3. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И ТЕЛЕКОММУНИКАЦИИ

Системный анализ, управление и обработка информации (технические науки) (2.3.1)

УДК 519.27
DOI: 10.24160/1993-6982-2022-5-112-120

Параметрический МА-алгоритм обнаружения разладки гауссовского временного ряда по математическому ожиданию

Г.Ф. Филаретов, А.А. Ларин, В.А. Локтюшов

Решена задача обнаружения в реальном масштабе времени разладки гауссовского временного ряда, связанной со спонтанным скачкообразным изменением его математического ожидания. Для решения данной задачи предложено использование контролирующнго алгоритма параметрического типа на базе метода скользящего среднего (Moving Average) или МА-алгоритма. Отмечено, что, несмотря на то, что данный алгоритм известен, практически начиная с первых работ в области статистического контроля, его свойства, возможности и эффективность по сравнению с другими алгоритмами обнаружения разладки до сих пор (по ряду причин) изучены очень слабо.

Цель настоящей работы — всестороннее исследование характеристик МА-алгоритма, обеспечивающее решение задачи синтеза оптимальной процедуры обнаружения разладки. Исследование проводили с помощью имитационного моделирования. Приведены структура и поэлементное описание программы имитационного эксперимента, в полной мере воспроизводящей работу МА-алгоритма в реальном масштабе времени. С ее помощью установлено, что традиционный способ задания решающего порога h контролирующей процедуры, значение которого должно обеспечивать заданную величину среднего времени между ложными тревогами, когда подается сигнал о появлении разладки (сигнал тревоги), хотя в действительности объект остается в состоянии «норма», несостоятелен. Для фиксированного ряда указанных величин получены корректные значения порога h в зависимости от ширины окна скользящего усреднения N контролирующего МА-алгоритма. Аналогичным образом определены величины среднего времени запаздывания подачи сигнала тревоги в случае появления разладки заданного фиксированного уровня δ. На основе полученных данных найдены зависимости показателя эффективности контролирующей процедуры от N для набора различных величин среднего времени между ложными тревогами и δ. Доказано, что для каждого такого набора существует значение N, при котором контролирующая процедура обладает наибольшей эффективностью. Проведено сопоставление данной оптимальной процедуры с аналогичной процедурой алгоритма кумулятивных сумм (АКС), показавшее, что МА-алгоритм по своей эффективности в целом лишь незначительно уступает АКС, а в некоторых вариантах даже его превосходит.

Ключевые слова: разладка временно́го ряда, обнаружение разладки в реальном времени, алгоритмы обнаружения, алгоритм скользящего среднего, вероятностные характеристики МА-алгоритма.

Для цитирования: Филаретов Г.Ф., Ларин А.А., Локтюшов В.А. Параметрический МА-алгоритм обнаружения разладки гауссовского временного ряда по математическому ожиданию // Вестник МЭИ. 2022. № 5. С. 112—120. DOI: 10.24160/1993-6982-2022-5-112-120.

A Parametric MA-Algorithm for Gaussian Time Series Change Point Detection from Mathematical Expectation

G.F. Filaretov, A.A. Larin, V.A. Loktyushov

The problem of online detection of a Gaussian time series change connected with a spontaneous abrupt change in its mathematical expectation is solved. For solving the problem, it is proposed to use a parametric controlling algorithm based on the Moving Average method or the MA- algorithm. It is noted that, despite the fact that this algorithm has been known almost since the first works in the field of statistical control, its properties, capabilities, and efficiency in comparison with other change detection algorithms are still (for a number of reasons) have been studied to a very poor extent. The purpose of this work is to thoroughly study the MA-algorithm's characteristics with the aim to synthesize an optimal time series change detection procedure. The study was carried out using a simulation modeling method. The article presents the structure and element-wise description of the simulation experiment program, which fully replicates the MA-algorithm operation in the online mode. By using the developed program, it has been found that the conventional way of setting the control procedure deciding threshold h, the value of which should provide the preset value of the average time between false alarms, when an alarm signal about the appearance of a change is generated, although in reality the object remains in the "normal" state, is invalid. For a fixed series of the above-mentioned quantities, the correct values of the threshold h are obtained depending on the width of the moving averaging window N of the controlling MA-algorithm. Similarly, the values of the average delay time of producing an alarm signal when a change with the given fixed level δ appears. Based on the data obtained, dependences of the control procedure efficiency indicator N for a set of different values of average time between false alarms and δ have been found. It has been shown that for each such set there is a value of N at which the control procedure is most effective. This optimal procedure is compared with a similar procedure of the cumulative sum algorithm (a CUSUM algorithm), and it has been shown from the comparison results that the MA-algorithm is in general only slightly inferior to the CUSUM algorithm in efficiency, and in some cases even outperforms it.

Key words: time series change, time series change online detection, detection algorithms, moving average algorithm, MA-algorithm probabilistic characteristics.

For citation: Filaretov G.F., Larin A.A., Loktyushov V.A. A Parametric MA-Algorithm for Gaussian Time Series Change Point Detection from Mathematical Expectation. Bulletin of MPEI. 2022;5:112—120. (in Russian). DOI: 10.24160/1993-6982-2022-5-112-120.

Введение

Проблема разработки и совершенствования различных методов обнаружения внезапного изменения статистических характеристик (разладки) временных рядов остается актуальной в течение достаточно длительного периода. Как отмечено в обзоре [1], за рубежом наблюдается экспоненциальный рост числа публикаций по данной тематике. Во многом этот факт объясняется тем, что такого рода методы всё чаще находят своё применение в различных диагностических комплексах, системах экологического, технологического, медицинского мониторинга, в других областях, когда требуется оперативно, т. е. в ритме с поступлением измерительных данных, выявить анормальное поведение контролируемого объекта [2 — 4].

К настоящему времени предложено весьма большое число различных методов обнаружения разладки в реальном времени или, как их ещё называют, последовательных методов. Все они основаны на вычислении по наблюдаемым дискретным значениям x_i (i = 1, 2, ..., n…) в темпе с поступлением значений некоторой решающей функции $g_i = g_i(x_i, x_{i-1}, x_{i-2}, \dots)$. Данную функцию находят рекуррентно, исходя из значения решающей функции на предыдущем шаге и очередного измеренного значения x_i. Полученное значение g_i затем сравнивают с заданным пороговым уровнем h. Если $g_i < h$, то считается, что разладка отсутствует, и процесс контроля продолжается. Если же $g_i \geq h$, то подаётся сигнал о наличии разладки (сигнал тревоги), тогда следует проанализировать сложившуюся ситуацию. Если по объективным показателям действительно имеет место разладка, то необходимо временно приостановить контролирующую процедуру, выяснить и устранить причины анормального поведения объекта контроля и затем вновь запустить контролирующую процедуру.

Однако в силу стохастичности контролируемого процесса сигнал тревоги может появиться и тогда, когда на самом деле разладка отсутствует. Это ситуация, так называемой, ложной тревоги. Проведение контролирующей процедуры в этом случае целесообразно продолжить, но, конечно, желательно, чтобы подобная ситуация возникала как можно реже. С учетом данного обстоятельства для описания потребительских свойств конкретного контролирующего алгоритма, реализующего тот или иной метод обнаружения разладки, вводится такой показатель как среднее время между ложными тревогами $\overline{T}_{\text{лт}}$. Его значение легко увеличить путем повышения порогового уровня h, однако при этом, к сожалению, вырастет другой важный показатель контролирующего алгоритма — среднее время запаздывания $\overline{\tau}_{\text{зап}}$ в обнаружении разладки определенной (заданной) величины δ, характеризующей степень отклонения контролируемого параметра временного ряда от нормы. Показатель $\overline{\tau}_{\text{зап}}$ характеризует быстродействие алгоритма обнаружения разладки, и желательно, чтобы он имел наименьшее значение.

Для сопоставления свойств различных алгоритмов обнаружения разладки удобен комплексный показатель их эффективности $E = \overline{T}_{\text{лт}} / \overline{\tau}_{\text{зап}}$. При этом, естественно, необходима одинаковая исходная информация о вероятностных свойствах контролируемых процессов и характере разладки. Предположим, что контролируемый временно́й ряд является гауссовским с некоррелированными отсчетами и постоянной дисперсией σ_X^2, а разладка связана со скачкообразным изменением значения математического ожидания от уровня m_0 до уровня $m_1 > m_0$, тогда $\delta = \dfrac{m_1 - m_0}{\sigma_X} > 0$. Пусть (для простоты) $m = 0$, следовательно, $\delta = m_1/\sigma_X$.

Постановка задачи исследования

Среди последовательных алгоритмов обнаружения разладки временных рядов в настоящее время наибольшее распространение на практике получили следующие алгоритмы: Шухарта, кумулятивных сумм (АКС) или CUSUM-алгоритм, экспоненциального взвешенного скользящего среднего или EWMA-алгоритм (EWMA — Exponentially Weighted Moving Average), скользящего среднего или MA-алгоритм (MA — Moving Average) [5].

Алгоритм Шухарта — исторически наиболее ранний алгоритм, который можно отнести к числу предвестников алгоритмов обнаружения разладки [6 — 8]. Его первоначальный вариант (контрольная карта Шухарта) предназначался для решения совсем другой задачи, а именно, для контроля стабильности технологического процесса массового производства некоторых изделий. Для контроля стабильности предлагалось осуществлять периодический отбор N изделий с последующим измерением значений контролируемой переменной x_l, вычислением среднего $\overline{X} = 1\Big/ N\sum_{l=1}^{N} x_l$ и его сопоставлением с номинальным заданным значением контролируемого параметра m_0. Для такого сопоставления выполняли расчет решающей статистики g и ее сравнение с критическим пороговым значением h при выбранном уровне значимости p. Для гауссовской последовательности и одностороннего варианта, когда отклонение от номинала могло быть только в сторону увеличения, решающая статистика g выглядит следующим образом:

$$g = \frac{\overline{X} - m_0}{\sigma_X}\sqrt{N}.$$

Процесс считался стабильным, если $g < h = z_{1-p}$, где z_{1-p} — соответствующий квантиль гауссовского распределения вероятностей.

Если использовать данный алгоритм в реальном масштабе времени, то необходимы предварительное выделение групп данных на примыкающих временных интервала длиной N и последующее вычисление соответствующих средних значений $\overline{X}$ в каждой группе. На практике значения g будут обновляться на каждом такте, кратном N, когда самое раннее значение суммы $\overline{X}$ удаляется и заменяется на новое. Такую процедуру трудно назвать последовательной, что впрочем неудивительно, поскольку в принципе карты Шухарта и не предназначались для решения задачи обнаружения разладки. Именно поэтому никакие показатели типа $\overline{T}_{\text{лт}}$ или $\overline{\tau}_{\text{зап}}$ применительно к различным классическим видам карт Шухарта долгое время не рассматривались.

Позднее ситуация несколько изменилась. Появились публикации, связанные с возможным использованием карт Шухарта для обнаружения разладки. В [9] предложено определять значение $\overline{T}_{\text{лт}}$ по формуле $\overline{T}_{\text{лт}} = 1/p$ (см. также [8]), где p — вероятность превышения решающей статистики g порога h. Значения h, установленные таким способом и обеспечивающие получение заданных величин $\overline{T}_{\text{лт}}$ для алгоритма Шухарта, приведены в табл. 1.

Таблица 1

Значения порога h для получения заданных величин $\overline{T}_{\text{лт}}$

h	2,054	2,326	2,652	2,878	3,090
$\overline{T}_{\text{лт}}$	50	100	250	500	1000

Алгоритм кумулятивных сумм (CUSUM-алгоритм). В общем виде решающую функцию g_i, используемую в АКС для случая независимых наблюдений, запишем следующим образом [10, 11]:

$$g_i = \max[0;\, g_{i-1} + z(x_i)];\ g_0 = 0, \tag{1}$$

где $z(x_i)$ — приращение решающей функции, подобное имеющему место в стандартном последовательном критерии отношения вероятностей Вальда [12]:

$$z(x_i) = \log[f_{1X}(x_i)/f_{0X}(x_i)],$$

где $f_{0X}(x_i)$, $f_{1X}(x_i)$ — функции плотности распределения вероятностей контролируемого процесса $X(t)$ до и при наличии разладки.

Из соотношения (1) следует, что решающая функция не может принимать отрицательные значения, что во многих случаях обеспечивает бо́льшее быстродействие АКС по сравнению с другими алгоритмами и его широкое использование при решении различных прикладных задач.

Алгоритм экспоненциального взвешенного скользящего среднего или *EWMA-алгоритм* [13] основан на стандартной формуле экспоненциального сглаживания, когда очередное значение решающей функции определяется как

$$g_i = (1 - \lambda)g_{i-1} + \lambda x_i,$$

где λ — параметр сглаживания.

Алгоритм скользящего среднего или *MA- алгоритм* имеет решающую функцию следующего вида [5]:

$$g_i = \frac{x_i + x_{i-1} + \ldots + x_{i-N+1}}{N};\quad i \geq N. \tag{2}$$

EWMA- и MA-алгоритмы как инструменты обнаружения разладки, к сожалению, изучены сравнительно слабо. Полученные к настоящему времени результаты не доведены до уровня, позволяющего синтезировать контролирующий алгоритм с заданными характеристиками. Для МА-алгоритма это, скорее всего, объясняется тем, что его ошибочно считают некоторым частным случаем алгоритма Шухерта.

Цель настоящей работы — детальное исследование МА-алгоритма, имея в виду получение в конечном итоге исчерпывающей информации, необходимой для синтеза процедуры обнаружения разладки в реальном масштабе времени. Рабочим инструментом получения является имитационное моделирование.

Планирование имитационного эксперимента

Для получения достоверных результатов имитационный эксперимент по исследованию МА-алгоритма обнаружения разладки должен отвечать ряду требований. Важнейшим из них является требование скрупулезного воспроизведения процесса функционирования МА-алгоритма в реальном времени, начиная с процесса запуска контролирующей процедуры вплоть до появления сигнала тревоги. Кроме того, должна быть обеспечена высокая точность получения конечных результатов за счет многократного повторения имитационных опытов с последующей статистической обработкой их результатов. В настоящем исследовании использована весьма высокая кратность повторения L, равная $10^4...10^5$. И, наконец, следует предусмотреть возможность задания значений сглаживающего окна N, решающего порога h, величины разладки δ в широком диапазоне, интересном для пользователей. В данном случае использованы значения N = 1...16 и δ из ряда значений, рекомендованных в методических материалах [14]: 0,5; 1,0; 1,5; 2,0; 2,5; 3,0. Что касается порога h, то его можно устанавливать по желанию экспериментатора, исходя из условия получения различных значений $\overline{T}_{\text{лт}}$ (в настоящем варианте в диапазоне от 100 до 1000).

Суть работы алгоритма имитационного моделирования заключается в L-кратном запуске процесса имитации работы МА-алгоритма. При каждом j-м запуске (j = 1, 2, … , L) фиксируется номер такта i, на котором среднее значение решающей функции g_i достигает или превосходит порог h: $g_i \ge h$. На этом j-й запуск заканчивается. Зафиксированное в нем значение i запоминается как T_j.

Всю последовательность операций на каждом j-м запуске можно разделить на две составные части — предварительную и основную.

Предварительная часть включает в себя следующие операции:

● генерацию N гауссовских случайных значений $x_i^* \ \left(i=\overline{1,N}\right)$ с $m_X = 0$ и дисперсией $\sigma_X^2 = 1$;

● заполнение стека значениями $x_1^*, x_2^*, \ldots, x_N^*$;

● вычисление начальных значений решающей функции $g_{0k} = \dfrac{1}{k}\sum_{i=1}^{k} x_i^*$; $k = 1,2,...,N_p$ и сравнение g_{0k} с решающим порогом h:

— если для какого-либо k (k = 1,2, …, N) $g_{0k} > h$, то предварительная часть реализуется повторно, начиная с первого пункта;

— если все $g_{0k} < h$, то предварительная часть завершается, и осуществляется переход к основной части алгоритма.

Фактически с помощью предварительной части исключается влияние переходного процесса, связанного с первоначальным заполнением стека длиной N, на конечные результаты моделирования.

Основная часть. На каждом такте i основной части проводятся:

● генерация значения x_i (i = 1, 2, ….) с заданным значением $m_X \ge 0$ и $\sigma_X^2 = 1$. Варианты задания m_X: $m_X = 0$ — для установления оценки $\overline{T}_{\text{лт}}$ и $m_X > 0$ — для определения при различных величинах разладки. Обновление содержимого стека: самое раннее значение из стека исключается, а очередное помещается на последнее место в стеке;

● вычисление текущего значения решающей функции g_i на i-м такте с помощью соотношения (2);

● сравнение значения g_i с решающим порогом h:

— если $g_i < h$, то тогда следует перейти к следующему (i +1)-ому такту данного j-го запуска;

— если $g_i \ge h$, то данный j-й запуск завершается, фиксируется номер такта i, запоминается значение T_j, равное зафиксированному значению i;

— после этого осуществляется переход к следующему (j + 1)-ому запуску, начиная с предварительной части.

Имитационный эксперимент заканчивается при завершении L-го запуска. Его результаты — значения T_j (j = 1, 2, …, L).

В обработку результатов имитационного эксперимента входят:

● определение минимального и максимального значений T_j;

● вычисления среднего значения, дисперсий и среднеквадратического отклонения:

$$\overline{T} = 1/L\sum_{j=1}^{L} T_j;\ \sigma_{T_j}^2 = 1/L\sum_{j=1}^{L}(T_j - \overline{T})^2;$$

$$\sigma^2\{\overline{T}\} = 1/L\,\sigma_{T_j}^2;\ \sigma\{\overline{T}\}.$$

Полученные результаты фиксируются как исходный материал для последующего анализа и интерпретации.

Результаты имитационного эксперимента

При проведении имитационного эксперимента в первую очередь проанализирован вопрос о применимости значений порога h для получения заданных величин $\overline{T}_{\text{лт}}$, используемых в алгоритме Шухарта и приведенных в табл. 1, для МА-алгоритма. Выяснилось, что такой способ задания порога h здесь неприменим.

Во всех вариантах наблюдалось увеличение значений $\overline{T}_{лт}$ с ростом N. В качестве примера в табл. 2 представлены данные о значениях $\overline{T}_{лт}$ в зависимости от N для порога h = 2,326, который, согласно табл. 1, должен обеспечивать получение $\overline{T}_{лт}$ = 100.

Наблюдаемое увеличение $\overline{T}_{лт}$ с ростом N, очевидно, обусловлено коррелированностью значений решающей функции g_i, тем большей, чем больше N. В этой связи компьютерная программа, реализующая алгоритм имитационного моделирования, была дополнена специальным модулем, осуществляющим поиск значения порога h, обеспечивающего заданную величину $\overline{T}_{лт}$. В результате для различных $\overline{T}_{лт}$ и N найдены необходимые значения порога h (табл. 3).

Указанные зависимости с высокой точностью можно аппроксимировать линейными функциями следующего вида:

для $\overline{T}_{лт}$ = 100: $h = 2{,}3808 - 0{,}0442\overline{T}_{лт}$;
для $\overline{T}_{лт}$ = 250: $h = 2{,}7022 - 0{,}0342\overline{T}_{лт}$;
для $\overline{T}_{лт}$ = 500: $h = 2{,}9225 - 0{,}0289\overline{T}_{лт}$;
для $\overline{T}_{лт}$ =1000: $h = 3{,}1408 - 0{,}0264\overline{T}_{лт}$.

При оценке быстродействия МА-алгоритма для всех вариантов значений N, δ и $\overline{T}_{лт}$ определено среднее время запаздывания в обнаружении разладки $\overline{\tau}_{зап}$. Результаты расчетов даны в табл. 4 — 7.

Анализируя данные, приведенные в табл. 4 — 7, можно убедиться, что для некоторой фиксированной

Таблица 2

Зависимость $\overline{T}_{лт}$ от N для порога h = 2,326 ($\overline{T}_{лт}$ = 100)

N	1	2	4	6	8	10	12	14	16
$\overline{T}_{лт}$	100	112	150	188	225	261	297	333	366

Таблица 3

Значения порога h для различных $\overline{T}_{лт}$ и N

$\overline{T}_{лт}$	N								
	1	**2**	**4**	**6**	**8**	**10**	**12**	**14**	**16**
100	2,326	2,281	2,158	2,051	1,959	1,874	1,800	1,729	1,677
250	2,652	2,622	2,529	2,445	2,372	2,307	2,250	2,193	2,148
500	2,878	2,855	2,781	2,709	2,646	2,589	2,546	2,494	2,446
1000	3,090	3,071	3,019	2,957	2,889	2,831	2,785	2,756	2,713

Таблица 4

Значения $\overline{\tau}_{зап}$ как функции N и δ при $\overline{T}_{лт}$ = 100

N	δ					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
1	29,709	10,814	4,885	2,684	**1,755**	**1,331**
2	21,711	7,294	3,605	**2,345**	1,841	1,594
3	18,729	6,387	**3,452**	2,49	2,058	1,804
4	16,806	6,03	3,531	2,683	2,249	1,966
5	15,716	5,911	3,684	2,897	2,412	2,098
6	14,980	**5,905**	3,868	3,043	2,551	2,215
7	14,542	6,011	4,051	3,196	2,671	2,324
8	14,234	6,102	4,219	3,329	2,774	2,407
9	13,969	6,244	4,366	3,446	2,881	2,487
10	13,822	6,383	4,521	3,552	2,968	2,535
11	13,742	6,545	4,633	3,664	3,041	2,634
12	13,652	6,711	4,774	3,748	3,122	2,677
13	13,636	6,85	4,97	3,829	3,181	2,732
14	**13,643**	6,992	4,974	3,927	3,253	2,813
15	13,735	7,133	5,089	4,003	3,311	2,854
16	13,738	7,276	5,177	4,071	3,371	2,898

Таблица 5

Значения $\bar{\tau}_{зап}$ как функции N и δ при $\bar{T}_{лт}$ = 250

N	δ					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
1	63,935	20,465	8,094	3,884	2,652	1,577
2	43,441	11,999	4,982	2,838	2,094	1,762
3	34,653	9,53	4,343	2,869	2,301	1,992
4	30,257	8,474	4,289	3,071	2,527	2,195
5	26,953	7,915	4,381	3,282	2,731	2,366
6	24,977	7,709	4,542	3,445	2,913	2,514
7	23,703	7,644	4,751	3,696	3,067	2,644
8	22,392	7,671	4,946	3,868	3,209	2,757
9	21,674	7,759	5,143	4,013	3,327	2,867
10	21,106	7,904	5,339	4,167	3,448	2,971
11	20,608	8,064	5,524	4,305	3,559	3,057
12	20,279	8,213	5,709	4,431	3,672	3,1415
13	19,925	8,368	5,836	4,549	3,754	3,215
14	19,754	8,544	5,983	4,665	3,846	3,286
15	19,629	8,749	6,117	4,767	3,929	3,361
16	19,563	8,922	6,264	4,874	4,005	3,432

Таблица 6

Значения $\bar{\tau}_{зап}$ как функции N и δ при $\bar{T}_{лт}$ = 500

N	δ					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
1	114,52	33,212	11,791	5,261	2,971	**1,819**
2	73,451	17,504	6,402	3,359	**2,289**	1,878
3	56,897	12,947	5,278	**3,204**	2,477	2,14
4	47,087	10,988	**4,952**	3,355	2,724	2,357
5	41,198	9,979	4,961	3,574	2,953	2,549
6	37,001	9,434	5,081	3,816	3,153	2,717
7	34,106	9,153	5,259	4,026	3,337	2,871
8	31,224	**9,063**	5,473	4,231	3,503	3,009
9	30,365	9,066	5,673	4,424	3,654	3,133
10	29,139	9,121	5,906	4,592	3,774	3,245
11	28,061	9,237	6,117	4,763	3,921	3,351
12	27,258	9,368	6,315	4,895	4,027	3,449
13	27,254	9,613	6,552	5,091	4,194	3,572
14	26,075	9,761	6,689	5,179	4,264	3,647
15	25,542	9,928	6,852	5,311	4,367	3,715
16	**25,368**	10,133	7,012	5,437	4,458	3,808

комбинации показателей δ и $\bar{T}_{лт}$ существует такое значение N, при котором имеет место минимальная величина $\bar{\tau}_{зап}$. В таблицах 4 — 7 значения N, которые можно назвать оптимальными в смысле обеспечения наибольшего быстродействия алгоритма обнаружения разладки, выделены полужирным шрифтом. Суммарно выделенные оптимальные значения N для различных δ и $\bar{T}_{лт}$ продемонстрированы в табл. 8, а в табл. 9 представлены соответствующие им показатели эффективности $E = \bar{T}_{лт}/\bar{\tau}_{зап}$ оптимизированного МА-алгоритма.

Именно такой оптимизированный вариант МА-алгоритма имеет смысл сопоставлять по эффективности с CUSUM-алгоритмом.

Сопоставление эффективности оптимизированного МА- и CUSUM-алгоритмов обнаружения разладки

Выбор CUSUM-алгоритма в качестве объекта сравнения неслучаен. Во-первых, это наиболее изученный алгоритм, если речь идет о разладке по математическому ожиданию. Во-вторых, теоретически доказана (по

Таблица 7

Значения $\bar{\tau}_{зап}$ как функции N и δ при $\bar{T}_{лт}$ = 1000

N	**δ**					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
1	206,43	54,697	17,805	7,422	3,605	3,088
2	125,23	25,892	8,441	3,972	**2,542**	**2,006**
3	93,022	17,943	17,912	6,397	2,666	2,261
4	74,896	14,472	5,741	**3,646**	2,913	2,51
5	63,366	12,657	**5,599**	3,865	3,159	2,721
6	55,919	11,606	5,614	4,108	3,388	2,911
7	50,047	11,042	5,791	5,763	3,597	3,077
8	44,783	**10.552**	5,961	4,551	3,753	3,219
9	42,966	10,636	6,223	4,791	3,948	3,376
10	40,581	10,498	6,451	4,988	4,104	3,512
11	38,523	10,502	6,686	5,169	4,251	3,634
12	36,958	10,621	6,911	5,339	4,396	3,741
13	35,641	10,773	7,115	5,505	4,531	3,858
14	34,484	10,921	7,328	5,673	4,649	3,961
15	33,711	11,136	7,522	5,821	4,775	4,064
16	**33,359**	11,391	7,765	5,991	4,895	4,172

Таблица 8

Оптимальные значения N для различных δ и $\bar{T}_{лт}$

$\bar{T}_{лт}$	**δ**					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
100	13	6	3	2	1	1
250	16	7	4	2	2	1
500	16	8	4	3	2	1
1000	16	8	5	4	2	2

Таблица 9

Показатели эффективности $E = \bar{T}_{лт}/\bar{\tau}_{зап}$ оптимизированного МА-алгоритма для различных δ и $\bar{T}_{лт}$

$\bar{T}_{лт}$	**δ**					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
100	7,33	16,93	28,97	42,64	57,14	75,13
250	12,78	32,71	58,29	88,09	119,4	158,5
500	19,71	55,17	101,0	156,1	218,4	274,9
1000	29,98	94,77	178,6	274,3	393,4	498,5

Таблица 10

Показатели эффективности E_{CUSUM} CUSUM-алгоритма при различных δ и $\bar{T}_{\text{лт}}$

$\bar{T}_{\text{лт}}$	δ					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
100	6,71	16,53	29,24	43,86	59,88	76,92
250	11,92	32,05	59,10	91,58	128,2	168,9
500	19,28	54,64	103,3	162,9	230,4	306,7
1000	32,07	95,15	183,5	292,4	418,4	550,7

Таблица 11

Показатели относительной эффективности Є при различных δ и $\bar{T}_{\text{лт}}$

$\bar{T}_{\text{лт}}$	δ					
	0,5	**1,0**	**1,5**	**2,0**	**2,5**	**3,0**
100	1,090	1,024	0,990	0,972	0,954	0,976
250	1,072	1,021	0,986	0,962	0,931	0,938
500	1,022	1,010	0,978	0,958	0,948	0,896
1000	0,935	0,996	0,973	0,938	0,940	0,905

крайней мере в асимптотике при $\bar{T}_{\text{лт}} \to \infty$) его оптимальность в смысле минимизации среднего времени обнаружения разладки [15]. Показатели эффективности E_{CUSUM} для CUSUM-алгоритма при различных δ и $\bar{T}_{\text{лт}}$ приведены в табл. 10. Все они получены с помощью программной системы STATCONT [16].

Используя данные табл. 9 и 10, несложно рассчитать и показатели относительной эффективноси MA-алгоритма по сравнении с CUSUM-алгоритмом: $Є = E/ E_{\text{CUSUM}}$ (табл. 11).

Опираясь на данные, приведенные в табл. 11, констатируем, что MA-алгоритм в большинстве вариантов мало в чем уступает по эффективности CUSUM-алгоритму и даже несколько превосходит его при малых значениях δ и $\bar{T}_{\text{лт}}$. Следует отметить вычислительную простоту MA-алгоритма, отсутствие необходимости при формировании решающей функции базового (номинального) значения δ, что обязательно для CUSUM-алгоритма.

Выводы

Рассмотрена задача оперативного обнаружения разладки гауссовского временнóго ряда в виде скачкообразного изменения его математического ожидания.

Для решения указанной задачи предложен контролирующий алгоритм параметрического типа на базе метода скользящего среднего (Moving Average) или MA-алгоритм, до настоящего времени изученный весьма слабо.

С помощью имитационного моделирования для различных значений усредняющего окна N найдены основные характеристики MA-алгоритма: среднее время между ложными тревогами $\bar{T}_{\text{лт}}$ в зависимости от решающего порога h и среднее время запаздывания $\bar{\tau}_{\text{зап}}$ в зависимости от уровня (величины) разладки δ.

Установлена возможность синтеза MA-алгоритма, обладающего наибольшим быстродействием путем выбора оптимального значения N для различных комбинаций величин $\bar{T}_{\text{лт}}$ и δ, а также подготовлен необходимый справочный материал.

Сопоставлены характеристики оптимизированного MA-и аналогичного по своему назначению CUSUM-алгоритмов. Доказано, что MA-алгоритм по своей эффективности в целом лишь незначительно уступает CUSUM-алгоритму, а в некоторых случаях при малых значениях δ и $\bar{T}_{\text{лт}}$ даже его превосходит.

Литература

1. **Shafid Ahmad.** Bibliometric Analysis of EWMA and CUSUM Control Chart Schemes // Intern. J. Information Technol. and Electrical Eng. 2018. V. 7(2). Pp. 1—11.

2. **Носкова А.И., Токранова М.В.** Обзор автоматизированных систем мониторинга // Интеллектуальные технологии на транспорте. 2017. № 1. С. 42—47.

3. **Еремин Н.А. и др.** Информационная автоматизированная система мониторинга и анализа технологических данных объектов нефтегазодобычи // Автоматизация,

References

1. **Shafid Ahmad.** Bibliometric Analysis of EWMA and CUSUM Control Chart Schemes. Intern. J. Information Technol. and Electrical Eng. 2018;7(2):1—11.

2. **Noskova A.I., Tokranova M.V.** Obzor Avtomatizirovannykh Sistem Monitoringa. Intellektual'nye Tekhnologii na Transporte. 2017;1:42—47. (in Russian).

3. **Eremin N.A. i dr.** Informatsionnaya Avtomatizirovannaya Sistema Monitoringa i Analiza Tekhnologicheskikh Dannykh Ob'ektov Neftegazodobychi. Avtomatiza-

телемеханизация и связь в нефтяной промышленности. 2020. № 2. С. 11—20.

4. **Funk P., Xiong N.** Why We Need to Move to Intelligent and Experience Based Monitoring and Diagnostic Systems // Proc. 23th Intern. Conf. Condition Monitoring and Diagnostic Eng. Management. 2010. Pp. 111—115.

5. **Roberts S.W.** A Comparison of Some Control Chart Procedures // Technometrcs. 1966. V. 8(3). Pp. 411—430.

6. **Shewhart W.A.** Quality Control Charts // Bell Syst. Techn. J. 1926. V. 5. Pp. 593—603.

7. **Murdoch J.** Control Charts. London: Macmillen Press, 1979.

8. **Адлер Ю.П., Максимова О.В., Шпер В.Л.** Контрольные карты Шухарта в России и за рубежом: краткий обзор современного состояния (статистические аспекты) // Стандарты и качество. 2011. № 7. С. 82—87; № 8. С. 82—87.

9. **Montgomery D.C.** Statistical Quality Control: a Modern Introduction. N.-Y.: John Wiley & Sons Inc., 2009.

10. **Page E.S.** Continuous Inspection Schemes // Biometrika. 1954. V. 41(1). Pp. 100— 115.

11. **Никифоров И.В.** Последовательное обнаружение изменения свойств временных рядов. М.: Наука, 1983.

12. **Айвазян С.А., Мхитарян В.С.** Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998.

13. **Roberts S.W.** Control Chart Tests Based on Geometric Moving Averages // Technometrics. 1959. V. 1(3). Pp. 239—250.

14. **Репин Д.С., Филаретов Г.Ф.** Методические аспекты исследования алгоритмов обнаружения разладки временных рядов // Информационные технологии в науке, образовании и управлении. 2020. № 1. С. 27—32.

15. **Ширяев А.Н.** Задача скорейшего обнаружения нарушения стационарного режима // Доклады АН СССР. 1961. Т. 138. № 5. С. 1039—1042.

tsiya, Telemekhanizatsiya i Svyaz' v Neftyanoy Promyshlennosti. 2020;2:11—20. (in Russian).

4. **Funk P., Xiong N.** Why We Need to Move to Intelligent and Experience Based Monitoring and Diagnostic Systems. Proc. 23th Intern. Conf. Condition Monitoring and Diagnostic Eng. Management. 2010:111—115.

5. **Roberts S.W.** A Comparison of Some Control Chart Procedures. Technometrcs. 1966;8(3):411—430.

6. **Shewhart W.A.** Quality Control Charts. Bell Syst. Techn. J. 1926;5:593—603.

7. **Murdoch J.** Control Charts. London: Macmillen Press, 1979.

8. **Adler Yu.P., Maksimova O.V., Shper V.L.** Kontrol'nye Karty Shukharta v Rossii i za Rubezhom: Kratkiy Obzor Sovremennogo Sostoyaniya (Statisticheskie Aspekty). Standarty i Kachestvo. 2011;7:82—87; 8:82—87. (in Russian).

9. **Montgomery D.C.** Statistical Quality Control: a Modern Introduction. N.-Y.: John Wiley & Sons Inc., 2009.

10. **Page E.S.** Continuous Inspection Schemes. Biometrika. 1954;41(1):100— 115.

11. **Nikiforov I.V.** Posledovatel'noe Obnaruzhenie Izmeneniya Svoystv Vremennykh Ryadov. M.: Nauka, 1983. (in Russian).

12. **Ayvazyan S.A., Mkhitaryan V.S.** Prikladnaya Statistika i Osnovy Ekonometriki. M.: YUNITI, 1998. (in Russian).

13. **Roberts S.W.** Control Chart Tests Based on Geometric Moving Averages. Technometrics. 1959;1(3): 239—250.

14. **Repin D.S., Filaretov G.F.** Metodicheskie Aspekty Issledovaniya Algoritmov Obnaruzheniya Razladki Vremennykh Ryadov. Informatsionnye Tekhnologii v Nauke, Obrazovanii i Upravlenii. 2020;1:27—32. (in Russian).

15. **Shiryaev A.N.** Zadacha Skoreyshego Obnaruzheniya Narusheniya Statsionarnogo Rezhima. Doklady AN SCSR. 1961;138;5:1039—1042. (in Russian).

Сведения об авторах:

Филаретов Геннадий Федорович — доктор технических наук, профессор кафедры управления и интеллектуальных технологий НИУ «МЭИ», e-mail: gefefi@yandex.ru
Ларин Андрей Александрович — аспирант кафедры управления и интеллектуальных технологий НИУ «МЭИ», e-mail: andrey.larin.a@yandex.ru
Локтюшов Владимир Алексеевич — студент кафедры управления и интеллектуальных технологий НИУ «МЭИ», e-mail: loktyushov_vl@mail.ru

Information about authors:

Filaretov Gennadiy F. — Dr.Sci. (Techn.), Professor of Control and Intelligent Technologies Dept., NRU MPEI, e-mail: gefefi@yandex.ru
Larin Andrey A.— Ph.D.-student of Control and Intelligent Technologies Dept., NRU MPEI, e-mail: andrey.larin.a@yandex.ru
Loktyushov Vladimir A. — Student of Control and Intelligent Technologies Dept., NRU MPEI, e-mail: loktyushov_vl@mail.ru

Конфликт интересов: авторы заявляют об отсутствии конфликта интересов
Conflict of interests: the authors declare no conflict of interest

Статья поступила в редакцию: *10.03.2022*
The article received to the editor: *10.03.2022*