

DOI: 10.15593/2499-9873/2020.4.04

УДК 51-74

**Р.В. Файзуллин, Ш. Херинг**

Ижевский государственный технический университет  
имени М.Т. Калашникова, Ижевск, Россия

## **МЕТОД АНАЛИЗА КОИНТЕГРАЦИОННЫХ ЗАВИСИМОСТЕЙ ДЛЯ ОБНАРУЖЕНИЯ НЕИСПРАВНОСТЕЙ НА ОСНОВЕ ДАННЫХ С ДАТЧИКОВ**

Сенсоры являются популярным источником информации о работе сложных динамических технических систем. Рассмотрение данных с сенсоров как многомерного временного ряда используется в том числе для описания киберфизических систем. Анализ взаимосвязей внутри многомерного временного ряда позволяет выявлять зависимости между техническими процессами. Предложен способ обнаружения неисправностей системы на основе метода анализа коинтеграционных зависимостей.

Определено, что в данных для анализа можно выявить коинтеграционные зависимости как факты взаимозависимости данных с разных датчиков. Приведены расчеты на примере системы с 52 параметрами. Если предположить, что важно изменение не только конкретного параметра системы, но также важно изменение этого параметра относительно других, то необходимо рассматривать пары параметров. Из 1326 пар данных 75 оказались коинтегрированными, т.е. они находились в каком-то долгосрочном равновесии и их соотношение могло быть описано каким-то уравнением.

Проведенный анализ показал, что предложенный метод позволяет наглядно показывать ситуации с изменениями поведения, т.е. если в коинтегрированной паре данных один из параметров начинает существенно отклоняться от ожидаемых параметров, значит, это нестандартная ситуация для киберфизической системы. На практике это означает либо смену технологий (о чем знает оператор системы), либо поломку/аварию/сбой, в том числе из-за ошибок оборудования, изменения каких-то параметров используемых ресурсов. В последнем случае такая информация (о том, что процесс стал другим) может быть использована для привлечения внимания в целом, которая в итоге может привести к необходимости ремонта оборудования или его профилактики или перенастройки и т.п.

Проведенный анализ данных о работе сложного промышленного процесса (на примере датасета смоделированного процесса Tennessee Eastman Process) показал, что предложенный метод позволяет наглядно идентифицировать сбои в работе системы (аварии). Разные пары данных, выбранные из общего числа пар данных предложенным методом, могут иметь способность идентифицировать разные ошибки в работе системы (аварии в работе промышленного процесса). Все ошибки вызывают изменение поведения одной или нескольких пар данных. Таким образом, отслеживание поведения значения случайной составляющей  $u_t$  позволяет идентифицировать случаи отклонения данных, описывающих штатную работу процесса, от долгосрочного равновесия (в терминах коинтеграции), т.е. случаи сбоя системы (аварии).

Полученные результаты обладают наглядностью и объективностью и могут быть использованы операторами процессов для идентификации сбоев в системе (аварий) или служить источником для автоматического управления процессом.

**Ключевые слова:** коинтеграция, коинтеграционное соотношение, коинтеграционная зависимость, временной ряд, многомерный временной ряд, киберфизическая система, сенсор, датчик, динамическая техническая система, Tennessee Eastman Process, идентификация ошибки, авария.

**R.V. Faizullin, S. Hering**

Kalashnikov Izhevsk State Technical University,  
Izhevsk, Russian Federation

## **COINTEGRATION ANALYSIS METHOD FOR FAULT DETECTION BASED ON SENSOR DATA**

**Abstract.** Sensors are a popular source of information about the operation of complex dynamic technical systems. Considering data from sensors as a multidimensional time series is also used to describe cyber-physical systems. The article proposes a method for detecting system malfunctions based on the method of analyzing cointegration dependencies. It is determined that in the data for analysis it is possible to reveal cointegration dependences as facts of interdependence of data from different sensors.

Calculations are given on the example of a system with 52 parameters. Out of 1,326 data pairs, 75 are cointegrated. The conducted analysis shows that the proposed method enables one to clearly illustrate situations with changes in behavior.

Having identified cointegrated pairs, we can follow them, and if cointegration has 'disappeared', that is, at some new time interval we can no longer talk about the presence of a cointegration ratio, then something has changed in the process itself. In practice, this means either a change in technology (which the operator knows about), or a breakdown/accident/failure, due to equipment errors, changes in some parameters of the resources used. In the latter case, such information (that the process has changed) can be used to attract attention in general, which may ultimately lead to the need for equipment repair or maintenance or readjustment, etc.

The analysis shows that the proposed method enables one to clearly illustrate situations with changes in behavior. As an example of using the method, we used the ready-made Tennessee Eastman Process (TEP) data set. Different pairs of data may have the ability to identify different errors. All errors cause a change in the behavior of one or several pairs of data, thus tracking the behavior of the value of random component enables identifying cases of deviation of the process from long-term equilibrium (in terms of cointegration), that is, cases of failure from the normal system operation.

The results obtained are clear and objective and can be used by process operators or by a source for automatic process control.

**Keywords:** cointegration, cointegration ratio, cointegration dependence, time series, multidimensional time series, cyber physical system, sensor, sensor, dynamic technical system, Tennessee Eastman Process, error identification, accident.

### **Введение**

Четвертая промышленная революция предзнаменует реорганизацию большинства отраслей промышленности. Далее мы будем обозначать этот процесс как Индустрия 4.0 (Industry 4.0). Как указывает Х. Вильдемманн [1], основная идея Индустрии 4.0 состоит в двух элементах: глобальной межсистемной сети людей, заводов и продуктов, а также независимой и децентрализованной организации и управления этими производственными единицами в режиме реального времени [1]. По его мнению, технико-экономическая парадигма Индустрии 4.0 основана на слиянии физического мира с виртуальным миром в киберфизическую систему [1]. Помимо интернета вещей, современные тенденции в области информационно-коммуникационных технологий, такие

как «большие данные», облачные сервисы и крупные технические достижения в области сенсорных технологий, формируют основу для расширения киберфизических систем в соответствии с Индустрией 4.0.

Интернет вещей присоединяет децентрализованные физические объекты к интернету. Эти объекты синхронизируются и автоматически сотрудничают, а также взаимодействуют с людьми через интерфейсы, образуя промышленную сеть реального времени. Это обеспечивает интеллектуальную автоматизацию. Современные технические системы являются сложными и динамичными, поэтому для описания работы такой системы необходимо отслеживать большое количество параметров, которые меняются во времени. Иными словами, сложные динамические системы могут быть описаны в виде многомерного временного ряда, где одни ряды будут описывать данные о технологических процессах, а другие об управляющем воздействии. Контроль условий технологического процесса имеет решающее значение для его нормальной работы.

Данные справедливо рассматриваются как важнейший ресурс. Для использования и управления постоянно растущими объемами данных требуются инновационные концепции интеллектуального анализа данных. Это также помогает создавать новые бизнес-модели. Популярный пример применения – «прогнозное обслуживание»: машина сообщает пользователю, когда ее нужно обслуживать, сокращая время простоя. Современный этап развития позволяет получать и использовать все больше и больше информации о технике.

Современные датчики не просто измеряют сигналы, они также анализируют окружающую среду. Из этого анализа формируются выходные сигналы для управления механизмами, что приводит к заданным механическим воздействиям. Интеллектуальные датчики позволяют не только «прогнозировать техническое обслуживание». Они влекут за собой гораздо более радикальную автоматизацию, в том числе машинное обучение, что обеспечивает дополнительную ценность при проектировании технологических процессов. Это открывает огромный потенциал для планирования безопасности, стандартов качества, гибкости и производительности.

Диагностика неисправностей в непрерывных динамических системах может быть сложной задачей, поскольку переменные в этих системах обычно характеризуются автокорреляцией, а также нестационарны-

ми параметрами, которые плохо обрабатываются методами, разработанными для стационарных систем. В динамических системах стационарные подходы расширяются для решения этих проблем, главным образом путем извлечения признаков, предназначенных для описания динамики процесса из временных рядов [2]. Современные киберфизические системы – это типичные примеры сложных динамических технических систем. Такие системы сейчас оснащены и сенсорами, и механизмами управления. Такие системы генерируют большие объемы данных, которые необходимо анализировать, и если данные не будут отвечать нашим ожиданиям, то это может быть признаком возможной поломки, аварии или необходимости вмешательства человека в управление.

Данные сенсоров в виде многомерных временных рядов являются одним из источников информации для принятия решения относительно сложного динамического объекта в концепции интернета вещей. В интернете вещей для анализа данных используют датчики. Важным является рассмотрение многомерного временного ряда (набора данных в виде временных рядов с каждого датчика и механизма управления), потому что это позволяет учитывать взаимосвязи между данными. Анализ этих данных, полученных с датчиков, является важной задачей и может найти полезную скрытую информацию, помимо самих данных [3]. Исследователи используют корреляционный анализ [3–6], анализ на наличие коинтеграции [7–10], методы главных компонент [11–13], нейронные сети [14–17] для анализа данных CPS и прогнозирования поломок. Однако даже самые сложные системы прогнозирования не дают 100%-ного результата. Способы идентификации факта того, что временной ряд изменился (произошел структурный сдвиг), например методом Перрона – Бая [18], основаны на том, что мы должны получить достаточное количество наблюдений после сдвига, чтобы оценки были достоверными.

## **1. Методология**

Рассмотрим коинтеграцию как свойство временных рядов, для чего надо также упомянуть о понятиях стационарности/нестационарности временных рядов.

Многие процессы могут быть описаны как модели авторегрессии. Действительно, временные ряды каких-либо параметров часто являются в некоем роде инерционными процессами, т.е. каждое значение зависит от предыдущего.

$$\gamma_t = \phi\gamma_{t-1} + \varepsilon_t, \quad (1)$$

где случайная составляющая  $\varepsilon_t$  имеет нулевое математическое ожидание и стабильную дисперсию  $\varepsilon_t \sim N(0, \sigma^2)$ . Тогда в зависимости от значения коэффициента  $\phi$  выделяют три типа временных рядов [19, 20].

Можно сказать, что стационарный процесс не зависит от времени, а в нестационарном процессе есть какие-то зависимости от времени. Большая популярность исследований временных рядов в целом и стационарности временных рядов в частности пришла из эконометрики, потому что многие экономические процессы нестационарны, а большинство методов накладывают требование стационарности. В технике тоже часто встречаются временные ряды, которые демонстрируют тенденции, достаточно регулярные циклы или другие проявления нестационарного поведения. Если какой-то из показателей стабильно растет во времени, то его можно привести к стационарному виду, просто удалив линию тренда. Если это преобразование приводит к стационарному ряду, то он называется разностно-стационарным [20]. Это не единственный способ сделать ряд стационарным. Можно от исходного ряда перейти к первым разностям  $\Delta\gamma_t = \gamma_t - \gamma_{t-1}$ .

$$\Delta\gamma_t = \gamma_t - \gamma_{t-1} = \varepsilon_t. \quad (2)$$

Если для приведения к стационарному виду надо брать вторые разности, то процесс считается  $I(2)$  и т.д. В большинстве случаев достаточно первых разностей, вторые применяются заметно реже, а величина разностей выше двух почти никогда.

Определив наше понимание стационарных временных рядов и приведения нестационарных временных рядов к стационарному виду, перейдем к описанию коинтеграции.

Если линейная комбинация какого-то набора  $I(1)$  временных рядов  $Y_t = (y_{1t}, y_{2t} \dots y_{nt})^T$  обладает свойствами стационарного ряда, то их называют коинтегрированными.

$$\beta^T Y_t = \beta_1 y_{1t} + \beta_2 y_{2t} \dots + \beta_n y_{nt} \sim I(0). \quad (3)$$

Если существует хотя бы один такой вектор  $\beta$ , то набор данных считается коинтегрированным. Понятно, что если исходные ряды не-

стационарны, то, чтобы их линейная комбинация была стационарной, в нее должно входить минимум два временных ряда. Использование большого количества временных рядов в коинтеграционном соотношении осложнено взаимозависимостями между нестационарными временными рядами, что вызывает эффект мультиколлинеарности.

Понятно, что любой коинтеграционный вектор не уникален, потому что

$$k \cdot \beta^T Y_t = (\beta^*)^T Y_t \sim I. \quad (4)$$

Обычно коинтеграционное соотношение приводят к выражению зависимости одного временного ряда от других, тогда коинтеграционный вектор может быть записан следующим образом:

$$\beta = (1, -\beta_2, \dots, \beta_n)^T. \quad (5)$$

Таким образом, коинтеграционное соотношение может быть представлено как

$$\beta^T Y_t = y_{1t} - \beta_2 y_{2t} - \dots - \beta_n y_{nt} \sim I(0) \quad (6)$$

или

$$y_{1t} = \beta_2 y_{2t} + \beta_3 y_{3t} + \dots + \beta_n y_{nt} + u_t, \quad (7)$$

где  $u_t = \beta^T Y_t \sim I(0)$ .

Получается, что мы предполагаем наличие некоего уравнения долгосрочного равновесия:

$$y_{1t} = \beta_2 y_{2t} + \beta_3 y_{3t} + \dots + \beta_n y_{nt}. \quad (8)$$

Если для наличия коинтеграции необходимо, чтобы комбинация временных рядов образовывала стационарный процесс, то при наличии нескольких временных рядов таких комбинаций может быть несколько, тогда говорят о наборе коинтеграционных векторов:

$$B^T Y_t = \begin{pmatrix} \beta_1^T Y_t \\ \vdots \\ \beta_r^T Y_t \end{pmatrix} = \begin{pmatrix} u_{1t} \\ \vdots \\ u_{rt} \end{pmatrix} \sim I. \quad (9)$$

Один из результатов Грейнджера состоит в том, что в случае коинтегрированности  $I(1)$  рядов модель зависимости временных рядов

может быть представлена в форме модели коррекции ошибок (Error-Correction Model – ECM) [21]. Иными словами, существует некая функциональная зависимость

$$\Delta y_{1t} = f(y_{1t-1} - (\beta_2 y_{2t-1} + \beta_3 y_{3t-1} + \dots + \beta_n y_{nt-1})) = f(u_{t-1}). \quad (10)$$

С одной стороны, мы знаем, что случайная составляющая  $u_t$  имеет нулевое математическое ожидание, а с другой, что следующее изменение временного ряда зависит от значения случайной составляющей на предыдущем шаге. Таким образом, если процессы коинтегрированные, то они стремятся к значениям, которые заложены долгосрочным равновесием (8). Пользуясь этим, можно следить за величиной случайной составляющей  $u_t$ , и если ее анализ покажет, что в какой-то момент эта величина стала вести себя не как стационарный процесс, то можно подозревать возникновение каких-то изменений в процессах, т.е. фиксировать нарушение долгосрочного равновесия, это может означать в технике какую-то нештатную ситуацию, поломку, аварию, необходимость замены оборудования и т.д.

Возникает необходимость тестирования временных рядов на стационарность и на коинтеграцию.

В качестве примера использования метода воспользуемся готовым массивом данных Tennessee Eastman Process (TEP), который активно используется исследователями для решения подобных задач [2, 9, 11]. Датасет используется для оценки методов управления технологическими процессами и мониторинга. Процесс испытаний основан на моделировании реального промышленного процесса, в котором компоненты, кинетика и условия эксплуатации были изменены по собственному усмотрению. Данные в обучающих и тестовых наборах включали все манипулируемые и измеряемые переменные для общей сложности  $m = 52$  переменных наблюдения [22].

Схема процесса представлена на рис. 1.

Рассмотрим некое упрощение для поиска коинтеграционных векторов. С учетом того, что если линейная комбинация пары нестационарных временных рядов образует стационарный временной ряд, то такие пары могут снова образовывать множество других коинтеграционных соотношений. Даже перебирая только пары, мы сталкиваемся с ростом количества возможных вариантов. Для вышеобозначенного

примера с 52 переменными мы получаем следующее количество возможных сочетаний:

$$C_n^k = \frac{n!}{k!(n-k)!} \quad (11)$$

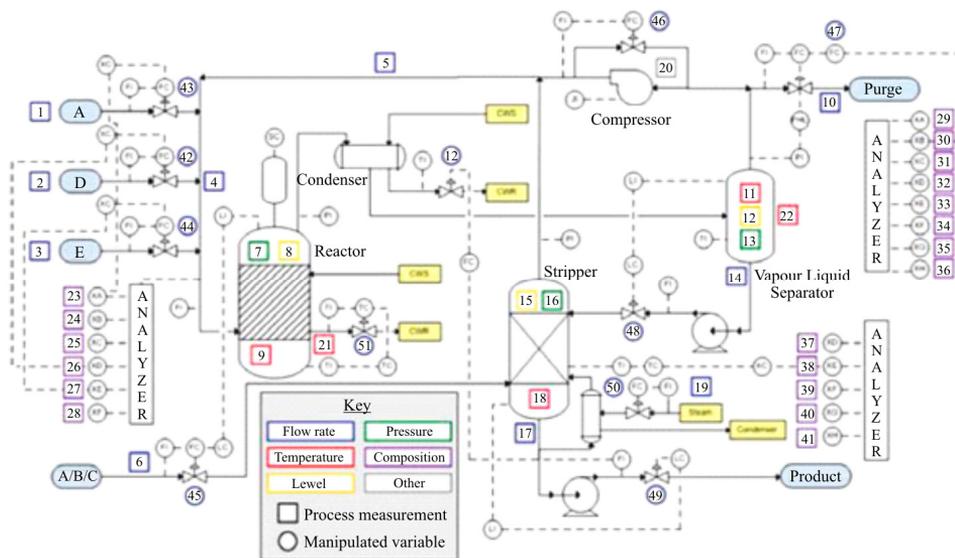


Рис. 1. Схема Tennessee Eastman Process

Иными словами, мы должны проверить  $C_{52}^2 = 1326$  пар на наличие коинтеграции. Определив коинтегрированные пары, мы можем следить за ними, и, если коинтеграция «исчезла», т.е. на каком-то новом отрезке времени мы уже не можем говорить о наличии коинтеграционного соотношения, значит, что-то изменилось в самих процессах. На практике это означает либо смену технологий (о чем знает оператор), либо поломку/аварию/сбой, в том числе из-за ошибок оборудования, изменения каких-то параметров используемых ресурсов. В последнем случае такая информация (о том, что процесс стал другим) может быть использована для привлечения внимания в целом, которая в итоге может привести к необходимости ремонта оборудования, или его профилактики, или перенастройки и т.п. Изменения в паре будем оценивать как расчет среднеквадратичного отклонения величины  $u_i$ .

## 2. Результаты

Очевидно, что оператор не способен вручную отслеживать большое количество параметров и тем более все пары. В нашем примере пришлось бы отслеживать 1326 пары параметров.

Мы проверили в наборе данных наличие коинтеграции в 1326 парах с помощью Python. Для этого использовали функцию `coint`.

Анализ временных рядов показал, что 75 пар коинтегрированы. В ходе анализа использовались методы и критерии, описанные в классических трудах по анализу временных рядов [23–26]. Рассмотрим некоторые из них на рис. 2–9. На рис. 2 (слева) мы видим график значений двух переменных в штатном режиме работы (`xmeas_33` и `xmv_5`), зависимость между которыми визуально определить сложно, однако анализ показывает, что пара переменных коинтегрирована. На рис. 2, 4, 6 и 8 (слева) мы видим достаточно схожие картины исходных значений при штатной работе, однако их изменения после аварии в системе `FaultNumber = 18`, графики стали совершенно разными (правые части рисунков).

На рис. 3, 5, 7 и 9 мы видим диаграммы размаха среднеквадратичных отклонений (СКО) случайной составляющей  $u_i$  для соответствующих пар данных для всех вариантов ошибок (1–20). Рис. 2–9 и расчеты произведены в языке R (library `tidyverse`).

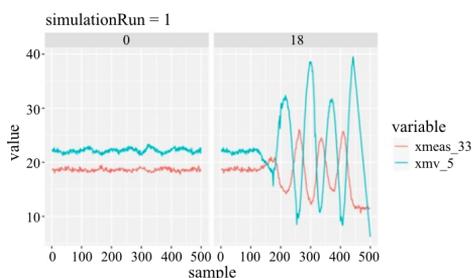


Рис. 2. График изменения переменных `xmeas_33` и `xmv_5` при штатной работе (слева) и после ошибки № 18

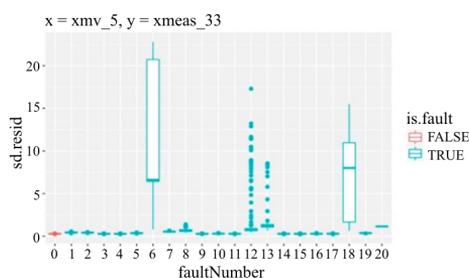


Рис. 3. Диаграммы размаха СКО случайной составляющей  $u_i$  для переменных `xmeas_33` и `xmv_5` для разных `FaultNumbers`

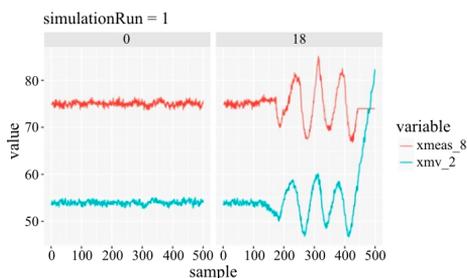


Рис. 4. График изменения переменных xmeas\_8 и xmv\_2 при штатной работе (слева) и после ошибки № 18

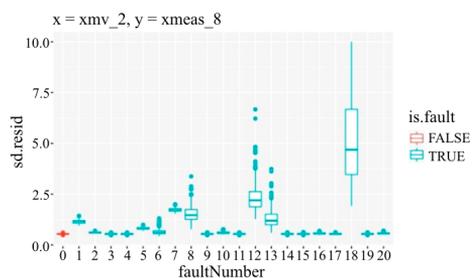


Рис. 5. Диаграммы размаха СКО случайной составляющей  $u_t$  для переменных xmeas\_8 и xmv\_2 для разных FaultNumbers

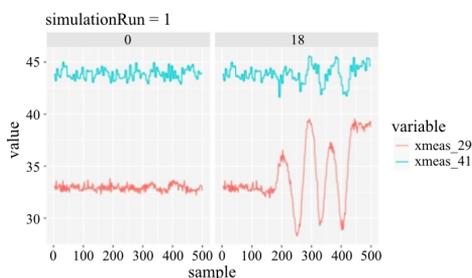


Рис. 6. График изменения переменных xmeas\_29 и xmeas\_41 при штатной работе (слева) и после ошибки № 18

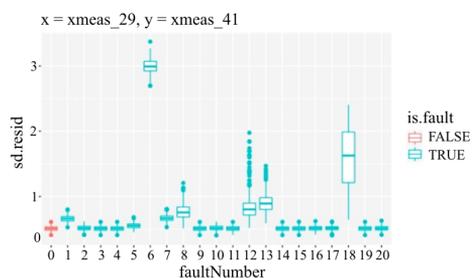


Рис. 7. Диаграммы размаха СКО случайной составляющей  $u_t$  для переменных xmeas\_29 и xmeas\_41 для разных FaultNumbers

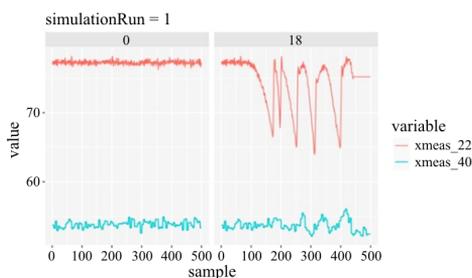


Рис. 8. График изменения переменных xmeas\_22 и xmeas\_40 при штатной работе (слева) и после ошибки № 18

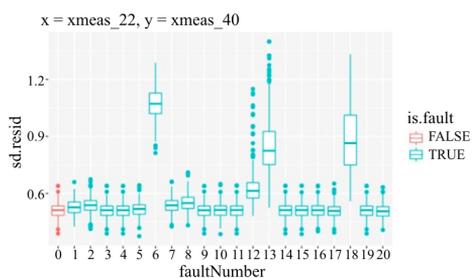


Рис. 9. Диаграммы размаха СКО случайной составляющей  $u_t$  для переменных xmeas\_22 и xmeas\_40 для разных FaultNumbers

Диаграммы размаха среднеквадратичных отклонений случайной составляющей  $u_i$  для переменных  $x_{meas\_8}$  и  $x_{mv\_2}$  для разных FaultNumbers (см. рис. 5) показывают, что для FaultNumbers 1, 7, 8, 12, 18 даже минимальные значения среднеквадратичных ошибок выше максимальных значений при штатной работе. Аналогичные результаты показывают и другие пары данных.

### **Выводы**

Проведенный анализ данных о работе сложного промышленного процесса (на примере датасета смоделированного процесса Tennessee Eastman Process) показал, что предложенный метод позволяет наглядно идентифицировать сбои в работе системы (аварии). Разные пары данных, выбранные из общего числа пар данных предложенным методом, могут иметь способность идентифицировать разные ошибки в работе системы (аварии в работе промышленного процесса). Все ошибки вызывают изменение поведения одной или нескольких пар данных. Таким образом, отслеживание поведения значения случайной составляющей  $u_i$  позволяет идентифицировать случаи отклонения данных, описывающих штатную работу процесса, от долгосрочного равновесия (в терминах коинтеграции), т.е. случаи сбоя системы (аварии).

Полученные результаты обладают наглядностью и объективностью и могут быть использованы операторами для идентификации сбоев в системе (аварий) или служить источником для автоматического управления процессами.

### **Благодарности**

*Работа выполнена при финансовой поддержке Ижевского государственного технического университета имени М.Т. Калашникова в рамках гранта № ФРВ / 20-55-25.*

### **Список литературы**

1. Wildemann H. Produktivität durch Industrie 4.0. – TCW Transfer-Centrum GmbH & Company KG, 2018. – 282 p.
2. Aldrich C 2019 Process Fault Diagnosis for Continuous Dynamic Systems Over Multivariate Time Series Time Series Analysis – Data, Methods, and Applications. – URL: <https://www.intechopen.com/books/time-series-analysis>

data-methods-and-applications/process-fault-diagnosis-for-continuous-dynamic-systems-over-multivariate-time-series (accessed 24 April 2020).

3. Mining data correlation from multi-faceted sensor data in the Internet of things / C. Dong, Q. Xiuquan, J. Gelernter, L. Xiaofeng, M. Luoming // *China Communications*. – 2011. – Vol. 8, no. 1. – P. 132–138.

4. Analyzing invariants in cyber-physical systems using latent factor regression / M. Momtazpour, J. Zhang, S. Rahman, R. Sharma, N. Ramakrishnan // *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015*. – Sydney, Australia, 2015. – P. 2009–2018.

5. Detection of cyber-physical faults and intrusions from physical correlations / A.Y. Lokhov, N. Lemons, T.C. McAndrew, A. Hagberg, S. Backhaus // *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW), 12–15 December 2016, Barcelona, Spain*. – Barcelona, Spain: IEEE, 2016. – P. 303–310.

6. Lavrova D., Pechenkin A. Applying correlation and regression analysis to detect security incidents in the internet of things // *International Journal of Communication Networks and Information Security*. – 2015. – Vol. 7, no. 3. – P. 131–137.

7. de Souza Goncalves E., Rosa P.F.F. Sensor fusion with cointegration analysis for IMU in a simulated fixed-wing UAV // *2017 International Conference on Military Technologies (ICMT), 31 May–2 June 2017, Brno, Czech Republic*. – Brno, Czech Republic: IEEE, 2017. – P. 493–499.

8. Tool wear monitoring based on cointegration modelling of multisensory information / G. Wang [et al.] // *International Journal of Computer Integrated Manufacturing*. – 2014. – Vol. 27, no. 5. – P. 479–487.

9. Li G., Qin S.J., Yuan T. Nonstationarity and cointegration tests for fault detection of dynamic processes // *IFAC Proceedings Volumes*. – 2014. – Vol. 47, no. 3. – P. 10616–10621.

10. Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data / P.B. Dao, W.J. Staszewski, T. Barszcz, T. Uhl // *Renewable Energy*. – 2017. – Vol. 116 (B). – P. 107–122.

11. Chiang L.H., Russell E.L., Braatz R.D. Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis // *Chemometrics and Intelligent Laboratory Systems*. – 2000. – Vol. 50, no. 2. – P. 243–252.

12. Identification of faulty sensors using principal component analysis / R. Dunia, S.J. Qin, T.F. Edgar, T.J. McAvoy // *AIChE Journal*. – 1996. – Vol. 42, iss. 10. – P. 2797–2812.

13. Yu T., Wang X., Shami A. Recursive principal component analysis-based data outlier detection and sensor data aggregation in IoT systems // *IEEE Internet of Things Journal*. – 2017. – Vol. 4, no. 6. – P. 2207–2216.

14. Zhang D., Lindholm G., Ratnaweera H. Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring // *Journal of Hydrology*. – 2018. – Vol. 556. – P. 409–418.

15. Neural network approach to forecast the state of the internet of things elements / I. Saenko, I. Kotenko, F. Skorik, S. Bushuev // 2015 XVIII International Conference on Soft Computing and Measurements (SCM), 19–21 May 2015, St.-Petersburg, Russia. – St.-Petersburg: IEEE, 2015. – P. 133–135.

16. Machine learning for Internet of Things data analysis: A survey / M.S. Mahdavejad, M. Rezvan, M. Barekatian, P. Adibi, P. Barnaghi, A.P. Sheth // Digital Communications and Networks. – 2018. – Vol. 4, no. 3. – P. 161–175. DOI: 10.1016/j.dcan.2017.10.002

17. Anomaly detection in cyber physical systems using recurrent neural networks / J. Goh, S. Adepu, M. Tan, Z.Sh. Lee // 2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE), 14 January 2017, Singapore. – Singapore: IEEE, 2017. – P. 140–145. DOI: 10.1109/HASE.2017.36

18. Bai J., Perron P. Estimating and testing linear models with multiple structural changes // *Econometrica*. –1998. – Vol. 66, no. 1. – P. 47–78.

19. Tsay R.S. Analysis of financial time series. – Second edition. – Hoboken: John Wiley & Sons, 2005. – 576 p.

20. Zivot E., Wang J. Modeling financial time series with S-Plus®. – New York: Springer-Verlag New York, 2006. – 1020 p.

21. Granger C.W.J. Essays in econometrics: collected papers of Clive W.J. Granger. – New York: Cambridge University Press, 2001. – 523 p. (Vol. 32. Econometric Society monographs).

22. Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation / C.A. Rieth, B.D. Amsel, R. Tran, M.B. Cook // Harvard Dataverse. – 2017. – Vol. 1. – URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/6C3JR1> (accessed 24 April 2020).

23. Hamilton J.D. Time Series Analysis. – Princeton: Princeton University Press, 1994. – 816 p.

24. Johansen S. Statistical analysis of cointegration vectors // *Journal of Economic Dynamics and Control*. – 1988. – Vol. 12, iss. 2–3. – P. 231–254.

25. Engle R.F., Granger C.W.J. Co-integration and error correction: representation, estimation, and testing // *Econometrica: journal of the Econometric Society*. – 1987. – Vol. 55, iss. 2. – P. 251–276.

26. Dickey D.A., Fuller W.A. Distribution of the estimators for autoregressive time series with a unit root // *Journal of the American Statistical Association*. – 1979. – Vol. 74, no. 366. – P. 427–431.

## References

1. Wildemann H. *Produktivität durch Industrie 4.0*. TCW Transfer-Centrum GmbH & Company KG, 2018. 282 p.

2. Aldrich C. 2019 Process Fault Diagnosis for Continuous Dynamic Systems Over Multivariate Time Series *Time Series Analysis Data, Methods, and*

*Applications*, available at: <https://www.intechopen.com/books/time-series-analysis-data-methods-and-applications/process-fault-diagnosis-for-continuous-dynamic-systems-over-multivariate-time-series> (accessed 24 April 2020).

3. Dong C., Xiuquan Q., Gelernter J., Xiaofeng L., Luoming M. Mining data correlation from multi-faceted sensor data in the Internet of Things. *China Communications*, 2011, vol. 8, no. 1, pp. 132-138.

4. Momtazpour M., Zhang J., Rahman S., Sharma R., Ramakrishnan N. Analyzing invariants in cyber-physical systems using latent factor regression. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2015, pp. 2009-2018.

5. Lokhov A.Y., Lemons N., McAndrew T.C., Hagberg A., Backhaus S. Detection of cyber-physical faults and intrusions from physical correlations. *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*, IEEE, 2016, pp. 303-310.

6. Lavrova D., Pechenkin A. Applying correlation and regression analysis to detect security incidents in the internet of things. *International Journal of Communication Networks and Information Security*, 2015, vol. 7, no. 3, pp. 131-137.

7. de Souza Goncalves E., Rosa P.F.F. Sensor fusion with cointegration analysis for IMU in a simulated fixed-wing UAV. *2017 International Conference on Military Technologies (ICMT), 31 May-2 June 2017, Brno, Czech Republic*, IEEE, 2017, pp. 493-499.

8. Wang G. et al. Tool wear monitoring based on cointegration modelling of multisensory information. *International Journal of Computer Integrated Manufacturing*, 2014, vol. 27, no. 5, pp. 479-487.

9. Li G., Qin S.J., Yuan T. Nonstationarity and cointegration tests for fault detection of dynamic processes. *IFAC Proceedings Volumes*, 2014, vol. 47, no. 3, pp. 10616-10621.

10. Dao P.B., Staszewski W.J., Barszcz T., Uhl T. Condition monitoring and fault detection in wind turbines based on cointegration analysis of SCADA data. *Renewable Energy*, 2017, vol. 116 (B), pp. 107-122.

11. Chiang L.H., Russell E.L., Braatz R.D. Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis. *Chemometrics and intelligent laboratory systems*, 2000, vol. 50, no. 2, pp. 243-252.

12. Dunia R. et al. Identification of faulty sensors using principal component analysis. *AIChE Journal*. 1996, vol. 42, no. 10, pp. 2797-2812.

13. Yu T., Wang X., Shami A. Recursive principal component analysis-based data outlier detection and sensor data aggregation in IoT systems. *IEEE Internet of Things Journal*. 2017, vol. 4, no. 6, pp. 2207-2216.

14. Zhang D., Lindholm G., Ratnaweera H. Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring. *Journal of hydrology*, 2018, vol. 556, pp. 409-418.
15. Saenko I., Kotenko I., Skorik F., Bushuev S. Neural network approach to forecast the state of the internet of things elements. *2015 XVIII International Conference on Soft Computing and Measurements (SCM)*, IEEE, 2015, pp. 133-135.
16. Mahdaveinejad M.S., Rezvan M., Barekatin M., Adibi P., Barnaghi P., Sheth A.P. Machine learning for Internet of Things data analysis: A survey. *Digital Communications and Networks*, 2018, vol. 4, no. 3, pp. 161-175. DOI: 10.1016/j.dcan.2017.10.002
17. Goh J., Adepu S., Tan M., Lee Z. Sh. Anomaly detection in cyber physical systems using recurrent neural networks. *2017 IEEE 18th International Symposium on High Assurance Systems Engineering (HASE)*, IEEE, 2017. pp. 140-145. DOI: 10.1109/HASE.2017.36
18. Bai J., Perron P. Estimating and testing linear models with multiple structural changes. *Econometrica*, 1998, vol. 66, no. 1, pp. 47-78.
19. Tsay R.S. Analysis of financial time series. Second edition. Hoboken, John Wiley & Sons, 2005, 576 p.
20. Zivot E., Wang J. Modeling financial time series with S-Plus®. New York, Springer-Verlag New York, 2006, 1020 p.
21. Granger C.W.J. Essays in econometrics: collected papers of Clive W.J. Granger. New York, Cambridge University Press, 2001, 523 p. (Vol. 32. Econometric Society monographs).
22. Rieth C.A., Amsel B.D., Tran R., Cook M.B. Additional Tennessee Eastman Process Simulation Data for Anomaly Detection Evaluation. *Harvard Dataverse*. 2017, vol. 1, available at: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/6C3JR1> (accessed 24 April 2020).
23. Hamilton J.D. Time Series Analysis, Princeton, Princeton University Press, 1994, 816 p.
24. Johansen S. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 1988, vol. 12, iss. 2-3, pp. 231-254.
25. Engle R.F., Granger C.W.J. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 1987, vol. 55, iss. 2, pp. 251-276.
26. Dickey D.A., Fuller W.A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 1979, vol. 74, no. 366, pp. 427-431.

Статья получена: 22.06.2020

Статья принята: 16.11.2020

### Сведения об авторах

**Файзуллин Ринат Васильевич** (Ижевск, Россия) – кандидат экономических наук, доцент кафедры «Менеджмент», Ижевский государственный технический университет имени М.Т. Калашникова (426069, Ижевск, ул. Студенческая, 7, e-mail: rf85@mail.ru).

**Штефан Херинг** (Ижевск, Россия) – аспирант кафедры «Приборы и методы измерений, контроля, диагностики», Ижевский государственный технический университет имени М.Т. Калашникова, (426069, Ижевск, ул. Студенческая, 7, e-mail: pmkk@istu.ru).

### About the authors

**Rinat V. Faizullin** (Izhevsk, Russian Federation) – Ph.D. in Economics, Associate Professor, Department of Management, Kalashnikov Izhevsk State Technical University (7, Studentcheskaya st., Izhevsk, 426069, e-mail: rf85@mail.ru).

**Stefan Hering** (Izhevsk, Russian Federation) – Ph.D. Student, Department of Instruments and Methods of Measurements, Control, Diagnostics, Kalashnikov Izhevsk State Technical University (7, Studentcheskaya st., Izhevsk, 426069, e-mail: pmkk@istu.ru).

### Библиографическое описание статьи согласно ГОСТ Р 7.0.100–2018:

**Файзуллин, Р.В.** Метод анализа коинтеграционных зависимостей для обнаружения неисправностей на основе данных с датчиков / Р.В. Файзуллин, Ш. Херинг. – DOI 10.15593/2499-9873/2020.4.04. – Текст: непосредственный // Прикладная математика и вопросы управления = Applied Mathematics and Control Sciences. – 2020. – № 4. – С. 49–64.

### Цитирование статьи в изданиях РИНЦ:

Файзуллин Р.В., Херинг Ш. Метод анализа коинтеграционных зависимостей для обнаружения неисправностей на основе данных с датчиков // Прикладная математика и вопросы управления. – 2020. – № 4. – С. 49–64. DOI: 10.15593/2499-9873/2020.4.04

### Цитирование статьи в references и международных изданиях:

#### Cite this article as:

Faizullin R.V., Hering S. Cointegration analysis method for fault detection based on sensor data. *Applied Mathematics and Control Sciences*, 2020, no. 4, pp. 49-64. DOI: 10.15593/2499-9873/2020.4.04 (in Russian)