



УДК 621.391:519.216

**Ю.Н. Минаев**, д-р техн. наук, **Н.Н. Гузий**, канд. техн. наук

Национальный авиационный университет  
(Украина, 03057, Киев, пр-кт космонавта Комарова, 1,  
тел. (044) 2495454, e-mail: min\_14@ukr.net),

**О.Ю. Филимонова**, канд. техн. наук, **Ю.И. Минаева**, канд. техн. наук

Киевский национальный университет строительства и архитектуры  
(Украина, 03037, Киев, Воздухофлотский пр-кт, 31,  
тел. (044) 2486427, 2425462, e-mail: filimonova@nm.ru; jumin@big-mir.net)

### **Анализ самоподобия многомерных временных рядов на основе методов интеллектуального анализа данных**

Предложены методы вычисления показателя Херста для одномерного и многомерного временных рядов (ВР) на основе главных диагоналей тензорных моделей ВР. Показано, что сложность проблемы обуславливает совместное применение нескольких математических теорий, в частности тензорный и многомерный матричный анализ. Приведены примеры применения предложенных методов.

*К л ю ч е в ы е с л о в а:* тензор, многомерный временной ряд, интеллектуальный анализ данных, 3D матрица, матричная развертка, самоподобие, параметр Херста.

Запропоновано методи обчислення показника Херста для одновимірних та багатовимірних часових рядів (ЧР) на основі головних діагоналей тензорних моделей ЧР. Показано, що складність проблеми зумовлює застосування декількох математичних моделей, зокрема тензорний і багатовимірний матричний аналіз. Наведено приклади використання запропонованих методів.

*К л ю ч о в і с л о в а:* тензор, багатовимірний часовий ряд, інтелектуальний аналіз даних, 3D матриця, матрична розвертка, самоподібність, параметр Херста.

Среди всех временных рядов (ВР) [1] многомерные (многокомпонентные, т.е. компоненты не связаны условием ортогональности) временные ряды (МВР) наименее изучены. Однако появление новых задач заставляет обращаться к поиску новых методов и моделей для их решения. Анализ МВР имеет ряд особенностей, которые отличают его от анализа числовых последовательностей. Основные из них следующие:

многомерный и многокомпонентный ВР имеют определенные отличия, так как понятие мерности предполагает ортогональный базис компонент, при

© Ю.Н. Минаев, Н.Н. Гузий, О.Ю. Филимонова, Ю.И. Минаева, 2017

этом многокомпонентность такого ограничения не предполагается, т.е. может быть неортогональный базис (в частности, косоугольный);

многомерные ВР относятся к объектам, имеющим сверхбольшие объемы, что обуславливает применение тензорных декомпозиций, которые не только аппроксимируются тензорами низких порядков (сокращение размерности), но и являются источниками скрытых знаний.

В некоторых работах [2] утверждается, что высокопорядковые тензоры (под которыми понимают мультимерные массивы) позволяют реализовать фундаментальный сдвиг парадигмы исследования к моделям, по существу полиномиальным, чья исключительность, в отличие от матричных методов, гарантирована естественными условиями.

Общая тенденция исследователей многомерных высокообъемных, нередко весьма зашумленных, данных такая: от матрицы к тензорам. Ранние методы многомерного анализа данных основаны на переформатировании тензора данных (многомерного массива) в матрицу и использовании методов, разработанных для классического  $2D$  анализа. Такой «выравнивающий» подход к данным и жесткие ограничения, присущие  $2D$  анализу, не всегда удовлетворяют требованиям анализа мультимерных сверхбольших данных, однако для большинства реальных прикладных задач, особенно для тех, где требуется качественная оценка, вполне применимы. В работе [2] показано, что тензорные декомпозиции — это не только и не столько точные матричные разложения на множители с дополнительными приложениями — мультилинейная алгебра значительно более богата структурно, чем линейная алгебра.

Современные нотации и соглашения основаны на том, что тензор порядка  $N$  обозначается  $A \in \mathbb{R}^{l_1 \times l_2 \times \dots \times l_N}$ , а данные —  $a_{i_1, i_2, \dots, i_N}$ . Тогда вектор  $N \times 1$  является тензором первого порядка, а  $N \times M$  — матрицей-тензором второго порядка. Подтензоры — части исходного тензора созданы с использованием только фиксированного подмножества индексов. Векторно-значные подтензоры называются волокнами и определяются фиксированием индексов по одному, а матрично-значные подтензоры называются кусками (слайсами) и определяются фиксированием по два. При обработке тензора часто требуется их переформатирование (восстановление формы). Восстановление тензора в матрицу называется матричным раскрытием. В работе [2] приведены основные и вспомогательные нотации, касающиеся тензоров и операций над ними. Классик тензорного анализа Г. Крон утверждает, что между тензором и матрицей такая же разница, как между линией и ее проекцией: линия одна, но проекций линии много.

Многомерные пространственные матрицы (МПМ), наиболее полное описание которых появилось в русскоязычной литературе еще в 60-е годы

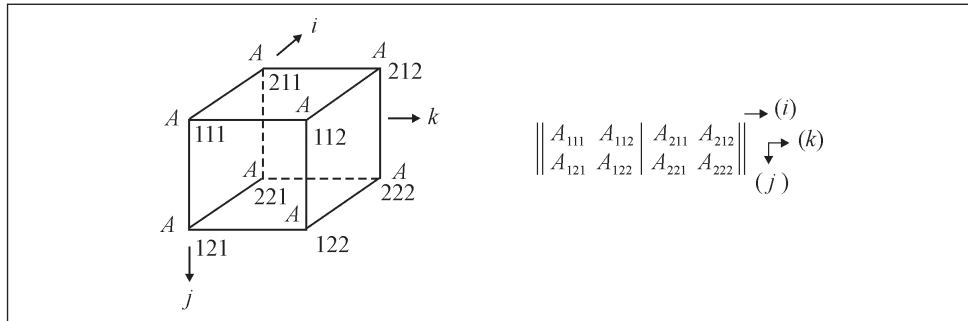


Рис. 1. Графическое изображение понятия «сечение ориентации» [3]

19-го столетия, заложили основу многомерного анализа данных. В работе [3] дано определение сечения ориентации  $i$  как совокупности элементов матрицы с фиксированным значением индекса. Все  $n$  сечений ориентации  $i$  в матрице параллельны друг другу и являются обычными двумерными матрицами  $n$ -го порядка (рис. 1).

Применение многомерных матриц (ММ) для анализа многомерных данных рассмотрено в работах [4, 5]. Так, в работе [5] доказано, что тензоры в общем случае могут быть представлены как ММ или, наоборот, некоторые операции декартова и общего тензоров произведений могут быть выполнены как многомерные матричные операции. Тем не менее, многие аспекты многомерной математической матрицы и тензорного анализа не являются взаимозаменяемыми.

**Современное состояние исследований, основные определения и математический аппарат.** В настоящее время анализ МВР направлен на решение проблемы многомерности с использованием тензорного представления данных [6], ММ [3] и поиска скрытых знаний на основе методов интеллектуального анализа данных (ИАД), реализуемых, в частности, с применением тензорных декомпозиций [7, 8]. В работах [9, 10] рассмотрено использование инвариантов тензоров для анализа МВР.

В работах [11, 12] тензор рассматривается как суперматрица в определенной координатной системе. Допустим, задан одномерный ВР  $\mathbf{X}(t) = \{x(t_i)\}$ ,  $i=1, N$ , и последовательность его фрагментов  $\{\mathbf{X}_j^{(t)}(t_i)\}$ ,  $i=1, I$ ,  $j=1, f$ . Каждый фрагмент состоит из окон  $x(t_i)$ ,  $i=1, I_x$  (см. рис. 1). С помощью процедуры матрицизации окно ВР представлено в виде тензора второго порядка (матрицей)  $m \times n$ , т.е.  $x(t_i) \rightarrow T_x = \text{reshape}(x(t_i), m, n)$ ,  $m \cdot n = I_x$ . В общем случае  $m \neq n$ , однако  $m = n$  предпочтительно с точки зрения вычислительной сложности. Рассматривается множество входных данных — векторы  $X = \{x_1, \dots, x_n\}$ , которые могут быть также 2D тен-

зорами  $\mathbf{X}(t) = \{X_{ij}\}_{i=1, j=1}^{mn}$ , в частности, они могут быть представлены 3D матрицей при использовании процедуры `reshape()`.

В [11] разработаны основные концепции обобщенных ММ и обобщенных матричных операций, где элементарная матрица является минимальной, имеющей одно или два измерения (строки и столбцы).

**Нотации и основные операции в ММ.** Обозначим  $A$   $d$ -мерную матрицу, число элементов которой равно  $m_k, k=1, 2, \dots, d$ . Тогда ММ  $A$  имеет порядок  $m_1 \times m_2 \times \dots \times m_d$ . Пусть  $a_{i_1}, a_{i_2}, \dots, a_{i_d}$  —  $\{i_1, i_2, \dots, i_d\}$ -й элемент  $d$ -мерной матрицы  $A$ . Для простоты обработки результатов допустим, что матрица  $A$  порядка  $m_1 \times m_2 \times m_3$  — это массив  $m_1 \cdot m_2 \cdot m_3$  скаляров, размещенных в  $m_1$  столбцов,  $m_2$  строк и  $m_3$  слоев. Таким образом, 3D матрица размерности  $m_1 \times m_2 \times m_3$  состоит из 2D матриц размерности  $m_1 \times m_2$ , размещенных в массиве размерностью  $m_3$ . Следовательно, матрицу  $A$  можно представлять, используя нотацию Матлаб, в форме массива матриц  $A(:, :, i_3)$  размерностью  $m_1 \times m_2$  и элементов  $A(:, :, i_3)$  в следующем виде:

$$A(:, :, i_3) = \begin{bmatrix} a_{11i_3} & a_{12i_3} & \cdots & a_{1m_2i_3} \\ a_{21i_3} & a_{22i_3} & \cdots & a_{2m_2i_3} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m_11i_3} & a_{m_12i_3} & \cdots & a_{m_1m_2i_3} \end{bmatrix}, \quad i_3 = 1, 2, \dots, m_3.$$

Базовые операции на ММ в современной нотации наиболее полно изложены в работах [3, 4, 11].

**Тензор-многомерный массив.** В работе [12] использована общепринятая в тензорном анализе нотация, где тензор  $p$ -го порядка с индексами  $p$  может быть представлен как многомерный массив данных [13—15] и записан в виде  $A = (a_{i_1, i_2, i_3}) \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ , матрица считается тензором второго порядка, а вектор — тензором первого порядка. Тензор третьего порядка может быть изображен как «куб» данных. Ориентация тензора высшего порядка не уникальна, однако она удобна, чтобы ссылаться на «слайс»-тензор как на тензор, сформированный посредством сохранения последней индексной константы. Например, если  $A = (a_{i_1, i_2, i_3})$  — тензор третьего порядка,  $k$ -й слайс в нотации Матлаб имеет вид  $A(:, :, k)$  (рис. 2) [13—16].

Высокопорядковая сингулярная декомпозиция (HOSVD) тензора  $A \in \mathbb{R}^{n_1 \times \dots \times n_d}$  включает обработку матрицы процедурой SVD на уровне своих модальных раскрытий  $A(1), \dots, A(d)$ . Результат — представление  $A$  в виде суммы  $\text{rank}-1$  тензоров. 3D массив может быть представлен вырезками (слайсами), каждый слайс — матрица (2D массив), каждая матрица — множество данных.

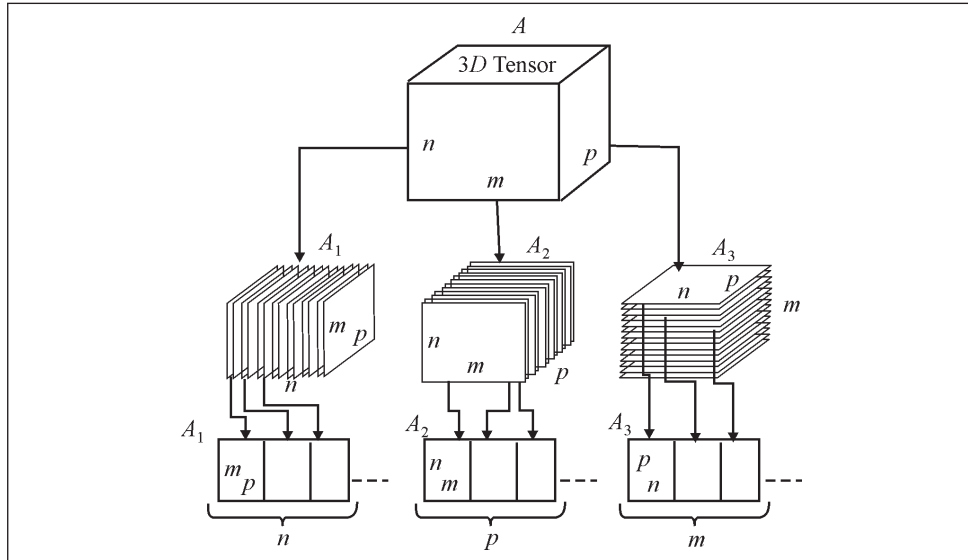


Рис. 2. Матрицизация тензора третьего порядка в матрицу из трех измерений:  $A_1 \in \mathbb{R}^{n \times (m \cdot p)}$ ,  $A_2 \in \mathbb{R}^{p \times (n \cdot m)}$ ,  $A_3 \in \mathbb{R}^{m \times (p \cdot n)}$  — первый, второй и третий способы матрицизации тензора  $A$  [12]

**Обобщенный алгоритм вычисления главной диагонали (следа) фрагмента МВР.** Для определения диагонали 3D тензора, моделирующего МВР используем два подхода: покомпонентную процедуру и учет всех компонент МВР:

1. Представление отдельной компоненты окна МВР в виде 2D тензора, рассматриваемого как фронтальный слайс —  $x(:, :, j), j = 1, J$ ; формирование 3D тензора окна  $\Xi^{(f)} = \{x(:, :, 1) | \dots | x(:, :, J), f = 1, F\}$ ,  $F$  — число тензор-окон фрагмента МВР; вычисление диагонали тензор-окна ( $\Xi^{(f)}$ ).

2. На основании таблицы измерений компонент МВР при помощи оператора reshape изначально формируется 3D тензор окна  $\Xi^{(f)} = \{x(:, :, 1) | \dots | x(:, :, J), f = 1, F\}$  и выполняется повторение подхода 1:

$$\text{reshape}(b, m, n) = (b(1:m) | b(m+1:2m) | \dots | b((n-1)m+1:mn)).$$

**Способы вычисления диагонали (следа) одномерного ВР.**

И с п о с о б ( б а з о в ы й). Диагональ (след) МВР вычисляется на основании диагонали (следа) отдельных компонент МВР. Одномерный (однокомпонентный) ВР  $\mathbf{X}^{(t)} = \{x(t_i)\}, i = 1, N$  задан в виде последовательности его фрагментов  $\{X_j^{(t)}(t_i)\}, i = 1, I, j = 1, f$ , где  $j$  — номер фрагмента; каждый фрагмент состоит из окон  $x(t_i), i = 1, I_x$  (рис. 3).

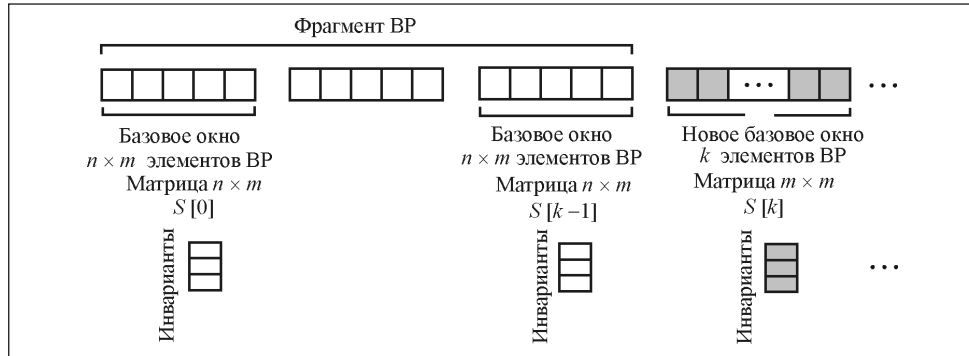


Рис. 3. Представление фрагмента ВР

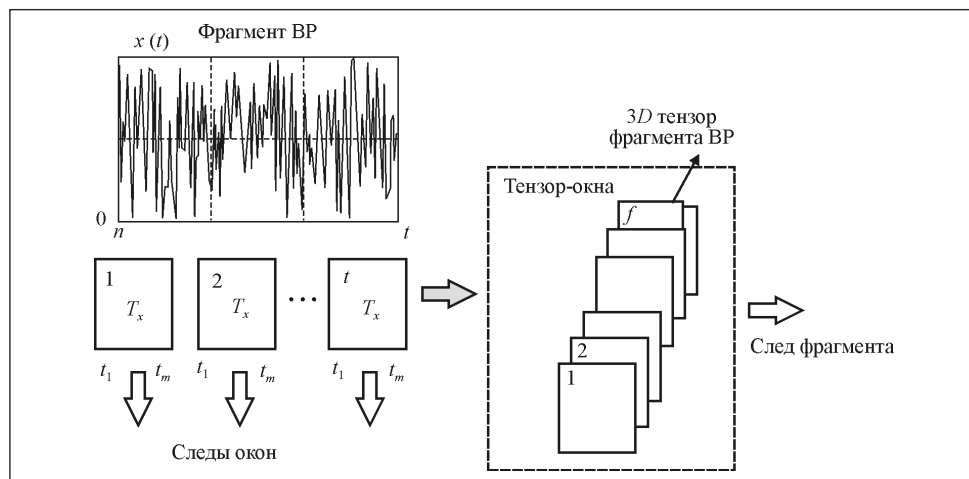


Рис. 4. Графическое представление алгоритма построения следа (диагонали) ВР для отдельного фрагмента и группы фрагментов

#### А л г о р и т м 1.

1. Выбираем размер окна  $x(t_i)$  ВР. Здесь  $i$  — порядковый номер элемента ВР в окне,  $i=1, I_x$ , где  $I_x$  — число элементов ВР в окне,  $I_x = m \cdot n$ ; в общем случае  $m \neq n$ , для прозрачности предпочтение отдаем  $m = n$ .

2. При помощи процедуры матрицизации окно ВР преобразуем в 2D тензор (матрица  $m \times n$ ), т.е.  $x(t_i) = [x(t_{i_1}), \dots, x(t_{i_x})] \rightarrow T_x = \text{reshape}(x(t_i), m, n)$ .

3. Формируем фрагмент ВР, состоящий из  $f$  окон:  $\{X_j^{(t)}(t_i)\}$ ,  $i=1, I_x$ ,  $j=1, f$ .

4. Формируем множество тензор-окон фрагмента:  $\mathbb{T}_X = \{(1)T_x, (2)T_x, \dots, (f)T_x\}$ .

5. Если  ${}^{(j)}T_x \subset \mathbb{T}_x$ , — квадратная матрица ( $m \times m$ ), то вычисляем диагональ  ${}^{(j)}T_x : z^{(j)} = \text{diag}({}^{(j)}T_x)/m$ .

П с п о с о б. Тензор-окна расположены в пространстве, как показано на рис. 3 и 4.

Рассматриваем тензор-окна, построенные для  $m \cdot n$  элементов ВР, при этом каждое из них рассматривается как фронтальный слайс, т.е.  ${}^{(1)}T_x = \mathbb{T}_x(:, :, 1)$ ,  ${}^{(2)}T_x = \mathbb{T}_x(:, :, 2)$ , ...,  ${}^{(f)}T_x = \mathbb{T}_x(:, :, f)$ . Если число тензор-окон равно трем, то тензор-фрагмент имеет вид  $\mathbb{T}_x = \langle \mathbb{T}_x(:, :, 1) | \mathbb{T}_x(:, :, 2) | \mathbb{T}_x(:, :, 3) \rangle$  (см. рис. 3, 4).

А л г о р и т м 2.

1. Формируется множество тензор-фрагментов:  $\mathbb{T}_x = \{ {}^{(1)}\mathbb{T}_x, {}^{(2)}\mathbb{T}_x, \dots, {}^{(f)}\mathbb{T}_x \}$ .

2. По правилам выполнения операций ММ вычисляется диагональ 3D матрицы  ${}^{(j)}\mathbb{T}_x : z^{(j)} = \text{diag}({}^{(1)}\mathbb{T}_x)$ .

В алгоритме 2 используются тензор-окна, построенные для  $m \cdot n$  элементов ВР. Каждое тензор-окно рассматривается как фронтальный слайс, т.е.  ${}^{(1)}T_x = \mathbb{T}_x(:, :, 1)$ ,  ${}^{(2)}T_x = \mathbb{T}_x(:, :, 2)$ , ...,  ${}^{(f)}T_x = \mathbb{T}_x(:, :, f)$ . Если, например, число тензор-окон равно трем, то тензор-фрагмент имеет вид  $\mathbb{T}_x = \langle \mathbb{T}_x(:, :, 1) | \mathbb{T}_x(:, :, 2) | \mathbb{T}_x(:, :, 3) \rangle$  (рис. 5).

Многомерный ВР, как правило, представляется в виде таблицы. Стилизованное графическое представление алгоритма для многомерного ВР, заданного таблично, показано на рис. 5.

А л г о р и т м 3.

1. Рассматривается возможность покомпонентного анализа ВР и интегрального учета всех компонент.

2. Формирование тензор-окон. В случае МВР — это 3D матрицы. Каждая 3D матрица в соответствии с [3, 6] представляется в виде системы 2D матриц, которая, в свою очередь, используется для вычисления инвариантов 2D матриц, моделирующих исходные 3D матрицы тензор-окон ВР.

3. Вычисление инвариантов 2D матриц, моделирующих исходные 3D матрицы тензор-окон [3, 13]. Построение диагонали, вычисление следа.

Вычисление следа (диагонали) МВР можно выполнить двумя способами:

1) сформировать тензор-окна для каждой компоненты МВР, из одноименных тензор-окон компонент, рассматриваемых как фронтальные слайсы, сформировать 3D тензор. На основании полученной совокупности 3D тензоров, образующих фрагмент, вычислить следы 3D тензоров, моделирующих одноименные многомерные тензор-окна;

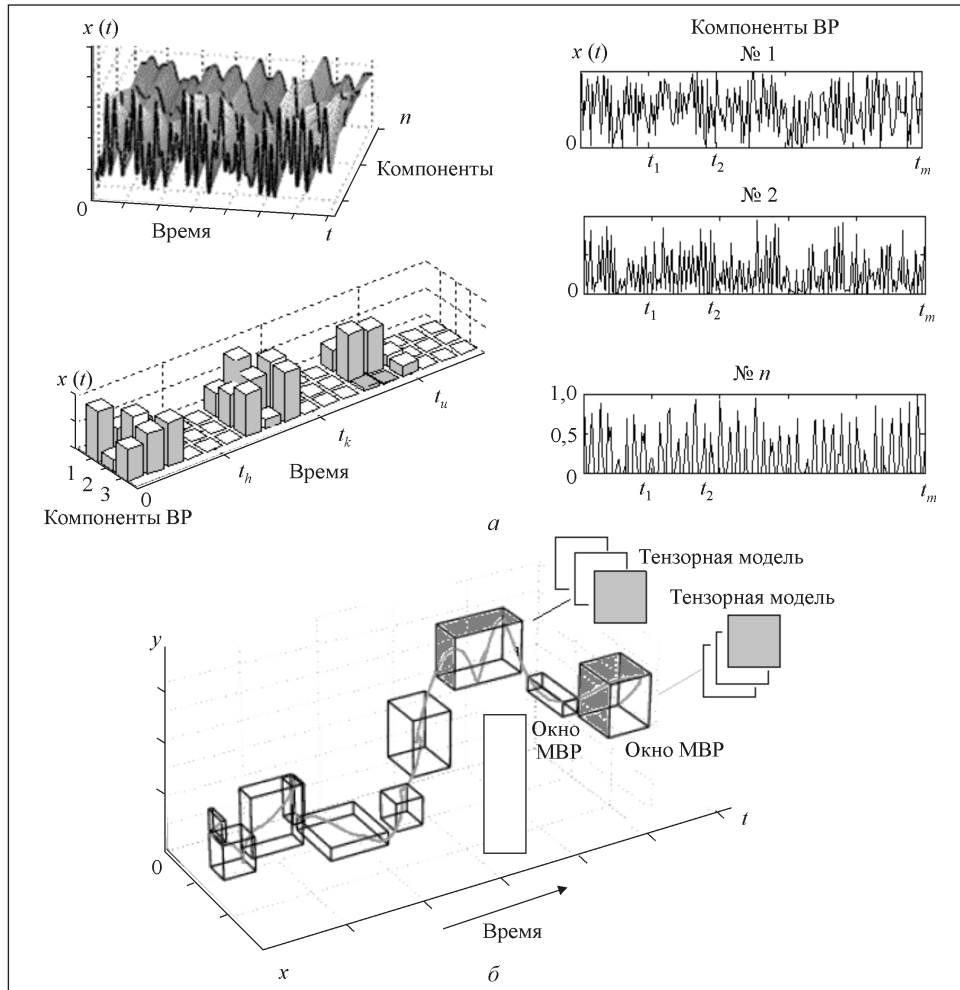


Рис. 5. Стилизованное представление многокомпонентного (а) и многомерного (б) ВР

2) таблицу МВР (где строки — компоненты ВР, столбцы — последовательные элементы одной компоненты) представить в виде тензора второго ранга с размерностью  $k \times f$  (рис. 6).

Тензор-окно, например № 1 (рассматриваемое здесь как фрагмент), преобразуем в  $3D$  тензор при помощи процедуры  $\text{reshape}(d(:, 1:f), i_1, i_2, i_3)$ ,  $i_1 \cdot i_2 \cdot i_3 = k \cdot f$ . На основании полученной совокупности  $3D$  тензоров, включающей элементы МВР ( $1 — 1 : f$ ;  $2 — 2f : 3f$ ;  $k — f(k-1) : fk$ ), образующих окно-фрагмент, вычисляем следы  $3D$  тензоров, моделирующих одноименные многомерные тензор-окна. Для полученной совокупности следов вычисляют тренд.



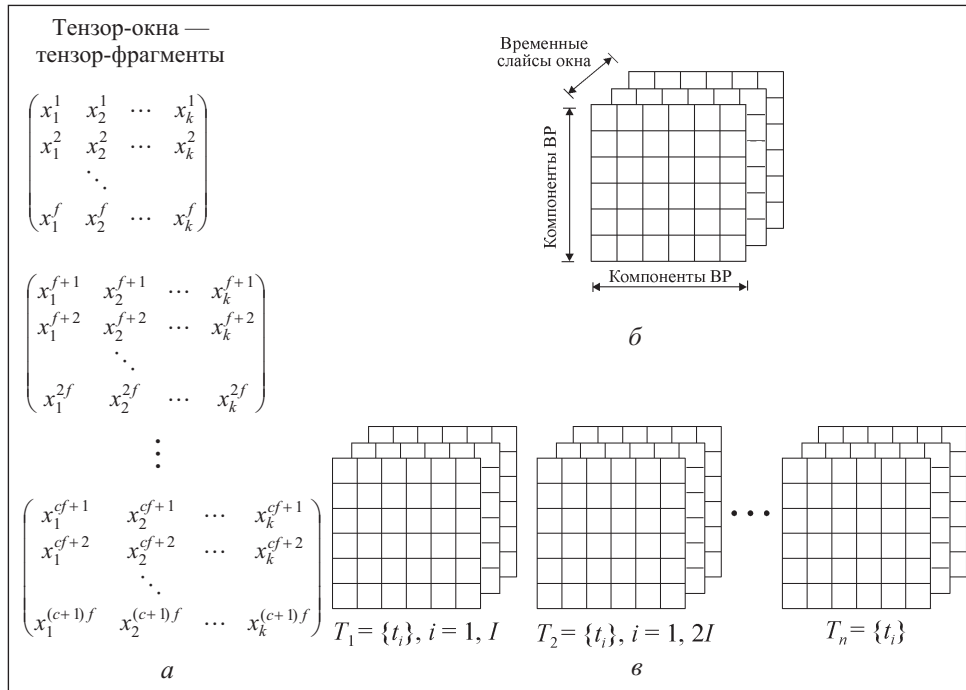


Рис. 6. Формирование тензор-окон многомерного ВР (а, б) и фрагмент (совокупность окон) МВР (в)

**Пример.** Фрагмент МВР, представленного покомпонентно

№	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1	0.6895	0.7631	0.8598	0.3423	0.9920
2	0.9231	0.9751	0.1732	0.2338	0.6338
3	0.5832	0.4397	0.7653	0.6545	0.0168
4	0.9875	0.3303	0.9185	0.2614	0.5155
5	0.5070	0.3012	0.4902	0.4496	0.4052
...	...	...	...	...	...
70	0.1298	0.7048	0.3867	0.9850	0.5638
71	0.6268	0.8461	0.2434	0.4991	0.2998
72	0.4306	0.5863	0.5993	0.4573	0.3479
73	0.9180	0.0469	0.8406	0.5741	0.7940
74	0.9780	0.8439	0.6605	0.9409	0.1266
75	0.2893	0.5868	0.6538	0.2483	0.8559
...	...	...	...	...	...

МВР (1:30, 1:5) представляется в виде 3D массива (тензора) фронтальными слайсами в виде

$$z = \text{reshape} \left( \underbrace{y(1:30,1:5)}_{2D \text{ тензор}}, \underbrace{10,5,3}_{3D \text{ тензор}} \right).$$

МВР  $\mathbf{X}(t)$ , где  $t = \{t_i\}$ ,  $i=1, I^t$ , рассматривается как совокупность окон равной длины  $\{X^{(j)}(t), j=1, J^x$ , т.е. окна  $X^{(k)}(t)$  и  $X^{(f)}(t)$  анализируются во временном интервале  $t_k + \Delta_t, t_f + \Delta_t$ , где  $t_f + \Delta_t$ , где  $\Delta_t$  — целое число,  $\Delta_t \in [1, t^\Delta]$ . С помощью процедуры  $\text{reshape}()$  оконное множество значений ВР преобразуется в тензор второго порядка (матрицу), т.е.

$$X^{(k)}(t) = \{x^k(t_k), \dots, x^k(t_k + t^\Delta)\} \xrightarrow{\text{reshape}()} \Xi_k,$$

$$\mathbf{X}(t) = \{\cup X^k(t)\}_{k=1, K} \rightarrow \{\cup \Xi_k\}_{k=1, K}.$$

Следует заметить, что для фронтальных слайсов

$$\Xi_k \in \{X_k(:, :, 1), X_k(:, :, 2), X_k(:, :, 3)\} \cdot (t) =$$

$$= \{\Xi_1 \cup \Xi_2 \cup \Xi_3 \cup \Xi_4 \cup \Xi_5\},$$

$$\Xi_1 = (x1(:, :, 1) | x1(:, :, 2) | x1(:, :, 3)), \quad \Xi_2 = (x2(:, :, 1) | x2(:, :, 2) | x2(:, :, 3)),$$

$$\Xi_3 = (x3(:, :, 1) | x3(:, :, 2) | x3(:, :, 3)), \quad \Xi_4 = (x4(:, :, 1) | x4(:, :, 2) | x4(:, :, 3)),$$

$$\Xi_5 = (x5(:, :, 1) | x5(:, :, 2) | x5(:, :, 3)).$$

Фрагмент оконного множества

	$x1(:, :, 1)$	$x1(:, :, 2)$	$x1(:, :, 3)$
$\Xi_1$	0.6895 0.1449 0.6045 0.4879	0.1672 0.5079 0.7107 0.9980	0.0589 0.6970 0.6268 0.0870
	0.9231 0.9927 0.2358 0.7426	0.5027 0.0576 0.5180 0.2818	0.7858 0.7868 0.4306 0.7917
	0.5832 0.7770 0.5819 0.4802	0.5805 0.9425 0.7230 0.6876	0.1320 0.2649 0.9180 0.3560
	0.9875 0.9354 0.8581 0.0606	0.4318 0.5271 0.8932 0.3734	0.0710 0.3102 0.9780 0.7086
	0.5070 0.1556 0.5169 0.4909	0.4173 0.8845 0.3964 0.1263	0.4616 0.8613 0.2893 0.1283
	0.7966 0.7070 0.0291 0.5231	0.3101 0.7336 0.8704 0.5776	0.4490 0.7795 0.7620 0.2279
	0.6413 0.3231 0.5137 0.8582	0.3707 0.0866 0.7570 0.4128	0.9223 0.1298 0.9877 0.3926

$\Xi_2$	$x_2(:, :, 1)$				$x_2(:, :, 2)$				$x_2(:, :, 3)$			
	0.7631	0.8687	0.5750	0.3552	0.8153	0.1448	0.4369	0.5825	0.7913	0.7292	0.8461	0.7464
	0.9751	0.4165	0.7142	0.0276	0.7297	0.8133	0.7631	0.8232	0.0919	0.8905	0.5863	0.1650
	0.4397	0.8309	0.3918	0.6421	0.9323	0.7426	0.2520	0.2544	0.3377	0.3328	0.0469	0.0683
	0.3303	0.3578	0.8286	0.2724	0.4751	0.3310	0.5074	0.6670	0.8478	0.7381	0.8439	0.2955
	0.3012	0.4458	0.3286	0.7130	0.1035	0.9775	0.7023	0.6310	0.9491	0.7906	0.5868	0.1658
	0.1630	0.5707	0.6479	0.9785	0.1838	0.6796	0.0140	0.6321	0.2033	0.3875	0.5120	0.3029
	0.0474	0.1764	0.1897	0.6594	0.4516	0.7768	0.1880	0.1120	0.6529	0.7048	0.4784	0.7988
$\Xi_3$	$x_3(:, :, 1)$				$x_3(:, :, 2)$				$x_3(:, :, 3)$			
	0.8598	0.1052	0.1114	0.3929	0.7270	0.1376	0.5843	0.9901	0.7790	0.9195	0.2434	0.0952
	0.1732	0.0282	0.9831	0.5878	0.8924	0.3815	0.9525	0.4822	0.1829	0.6896	0.5993	0.9097
	0.7653	0.5814	0.5771	0.2004	0.5099	0.4241	0.3491	0.5890	0.7854	0.9526	0.8406	0.7774
	0.9185	0.8204	0.1377	0.9631	0.4529	0.5479	0.6778	0.9300	0.0258	0.8585	0.6605	0.1498
	0.4902	0.7496	0.3934	0.2325	0.8717	0.7656	0.0231	0.1408	0.4074	0.8220	0.6538	0.1137
	0.8085	0.1229	0.7951	0.2180	0.2750	0.0505	0.8166	0.1441	0.9064	0.4774	0.7468	0.2218
	0.1977	0.6343	0.3812	0.8407	0.8965	0.3519	0.6991	0.7191	0.3304	0.3867	0.0948	0.7138
$\Xi_4$	$x_4(:, :, 1)$				$x_4(:, :, 2)$				$x_4(:, :, 3)$			
	0.3423	0.4668	0.0443	0.7789	0.3295	0.6082	0.4407	0.6204	0.7298	0.3307	0.4991	0.3597
	0.2338	0.5328	0.9233	0.2928	0.4158	0.4363	0.2683	0.5547	0.9731	0.1988	0.4573	0.5085
	0.6545	0.6163	0.4151	0.2229	0.1904	0.4328	0.5681	0.5961	0.7805	0.0343	0.5741	0.1904
	0.2614	0.4092	0.0028	0.5484	0.8312	0.6332	0.1248	0.8871	0.2234	0.3744	0.9409	0.4164
	0.4496	0.4110	0.3366	0.8640	0.8687	0.5961	0.7382	0.9402	0.6959	0.7294	0.2483	0.9306
	0.6507	0.6786	0.2236	0.4566	0.6178	0.3361	0.8488	0.8178	0.7144	0.3671	0.6909	0.4906
	0.1822	0.9305	0.5791	0.7126	0.1161	0.2109	0.9824	0.1459	0.1494	0.9850	0.4469	0.7147
$\Xi_5$	$x_5(:, :, 1)$				$x_5(:, :, 2)$				$x_5(:, :, 3)$			
	0.3423	0.4668	0.0443	0.7789	0.3022	0.0067	0.0795	0.1351	0.5495	0.8603	0.2998	0.0820
	0.2338	0.5328	0.9233	0.2928	0.7691	0.4758	0.9958	0.0161	0.8695	0.3785	0.3479	0.5133
	0.6545	0.6163	0.4151	0.2229	0.6537	0.7989	0.9373	0.9839	0.3677	0.1947	0.7940	0.3667
	0.2614	0.4092	0.0028	0.5484	0.3155	0.4621	0.9875	0.6723	0.4132	0.6364	0.1266	0.4449
	0.4496	0.4110	0.3366	0.8640	0.6521	0.5290	0.6313	0.2648	0.4026	0.2903	0.8559	0.3312
	0.6507	0.6786	0.2236	0.4566	0.3849	0.0016	0.0949	0.1519	0.5305	0.9379	0.5748	0.0121
	0.1822	0.9305	0.5791	0.7126	0.3357	0.2380	0.9795	0.4718	0.0096	0.5638	0.1426	0.2812

Фрагмент реального МВР, состоящего из пяти тензор-окон

Окно № 1 $x_1$				Окно № 2 $x_2$				Окно № 3 $x_3$			
$x_1(:, :, 1) =$				$x_2(:, :, 1) =$				$x_3(:, :, 1) =$			
0.39	0.63	0.73	0.88	0.67	0.39	0.64	0.05	0.44	0.75	0.25	0.83
0.92	0.35	0.64	0.86	0.76	0.48	0.64	0.55	0.04	0.21	0.13	0.15
0.46	0.31	0.27	0.60	0.68	0.72	0.05	0.64	0.84	0.09	0.81	0.90
0.66	0.90	0.73	0.48	0.33	0.58	0.99	0.21	0.78	0.22	0.61	0.82
0.52	0.11	0.89	0.01	0.98	0.22	0.39	0.23	0.86	0.16	0.51	0.55
0.98	0.30	0.51	0.28	0.51	0.23	0.30	0.07	0.71	0.64	0.01	0.99
0.90	0.85	0.55	0.57	0.82	0.36	0.25	0.55	0.59	0.64	0.03	0.56

$x1(:, :, 2) =$				$x2(:, :, 2) =$				$x3(:, :, 2) =$			
0.83	0.41	0.30	0.31	0.35	0.55	0.45	0.91	1.00	0.64	0.07	0.40
0.70	0.21	0.88	0.90	0.06	0.68	0.77	0.55	0.95	0.72	0.60	0.60
0.34	0.29	0.75	0.70	0.65	0.51	0.26	0.45	0.17	0.88	0.73	0.97
0.40	0.26	0.33	0.76	0.97	0.83	0.94	0.76	0.37	0.89	0.52	0.87
0.54	0.31	0.26	0.27	0.49	0.82	0.46	0.69	0.35	0.91	0.90	0.73
0.19	0.55	0.73	0.61	0.41	0.22	0.49	0.60	0.82	0.67	0.90	0.37
0.43	0.83	0.44	0.17	0.98	0.09	0.73	0.12	0.77	0.66	0.45	0.83
$x1(:, :, 3) =$				$x2(:, :, 3) =$				$x3(:, :, 3) =$			
0.15	0.61	0.26	0.24	0.39	0.53	0.21	0.10	0.51	0.10	0.72	0.64
0.68	0.41	0.03	0.37	0.68	0.27	0.58	0.78	0.91	0.38	0.52	0.59
0.22	0.62	0.64	0.96	0.81	0.10	0.25	0.75	0.14	0.14	0.50	0.08
0.76	0.43	0.82	0.47	0.10	0.55	0.65	0.15	0.42	0.85	0.17	0.96
0.07	0.42	0.78	0.53	0.28	0.94	0.01	0.02	0.92	0.06	0.69	0.04
0.24	0.72	0.48	0.05	0.04	0.14	0.01	0.05	0.88	0.38	0.59	0.55
0.35	0.88	0.84	0.62	0.45	0.54	0.73	0.38	0.77	0.27	0.51	0.86
Окно № 4 $x4$				Окно № 5 $x5$							
$x4(:, :, 1) =$				$x5(:, :, 1) =$							
0.21	0.85	0.86	0.69	0.40	0.98	0.91	0.78				
0.01	0.41	0.46	0.79	0.44	0.71	0.16	0.54				
0.44	0.32	0.60	0.95	0.99	0.75	0.45	0.83				
0.20	0.61	0.79	0.91	0.50	0.99	0.57	0.07				
0.06	0.60	0.64	0.89	0.98	0.02	0.48	0.80				
0.91	0.94	0.70	0.61	0.35	0.20	0.11	0.62				
0.93	0.28	0.61	0.14	0.22	0.95	0.57	0.53				
$x4(:, :, 2) =$				$x5(:, :, 2) =$							
0.87	0.56	0.22	0.55	0.05	0.78	0.18	0.06				
0.43	0.41	0.03	0.42	0.91	0.45	0.06	0.12				
0.79	0.14	0.10	0.64	0.48	0.56	0.32	0.53				
0.94	0.42	0.06	0.54	0.28	0.98	0.18	0.88				
0.75	0.41	0.71	0.75	0.07	0.80	0.63	0.54				
0.73	0.42	0.53	0.03	0.79	0.65	0.50	0.47				
0.66	0.63	0.68	0.23	0.04	0.86	0.72	0.00				
$x4(:, :, 3) =$				$x5(:, :, 3) =$							
0.71	0.72	0.98	0.22	0.70	0.07	0.36	0.07				
0.37	0.27	0.82	0.57	0.61	0.63	0.68	0.23				
0.58	0.73	0.70	0.16	0.26	0.16	0.61	0.72				
0.90	0.79	0.15	0.39	0.41	0.88	0.09	0.65				
0.75	0.86	0.71	0.23	0.92	0.43	0.72	0.20				
0.02	0.09	0.53	0.76	0.75	0.90	0.78	0.42				
0.50	0.20	0.92	0.12	0.10	0.98	0.55	0.18				

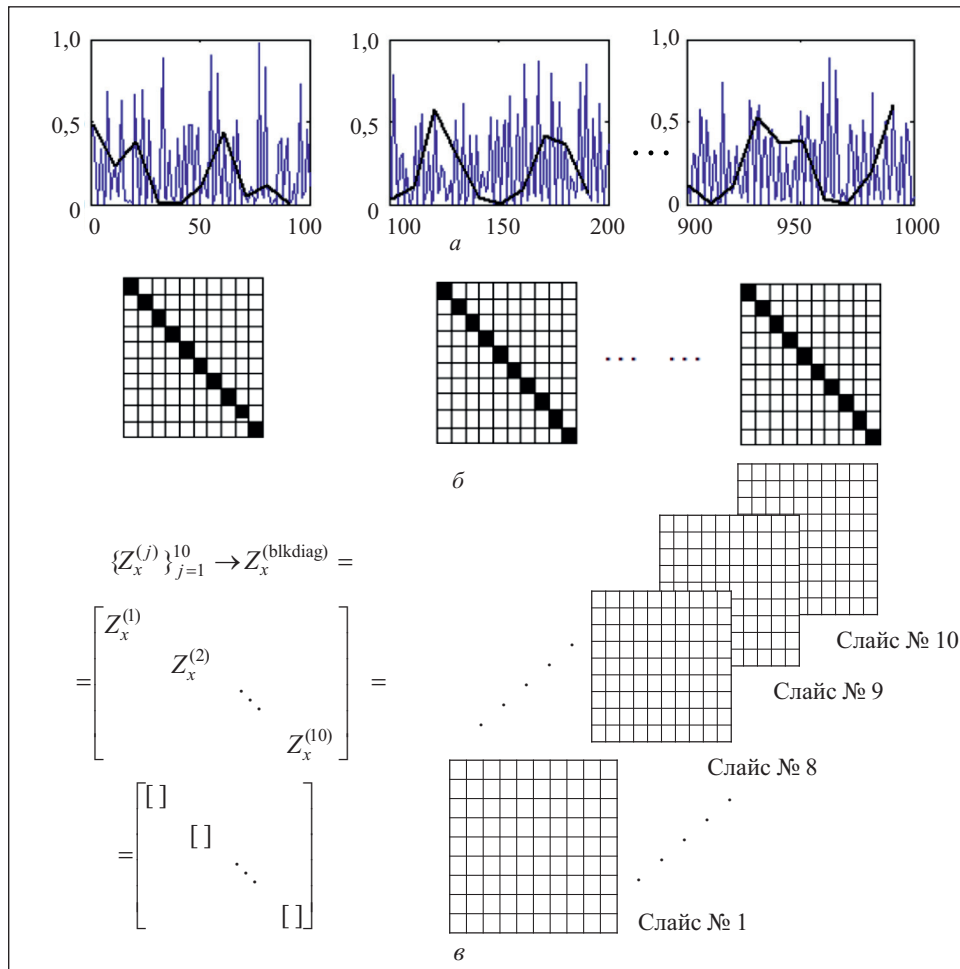


Рис. 7. Окна одномерного ВР ( $l_w = 100$  эл.) (а) и соответствующие им 2D матрицы ( $T_w$ ,  $10 \times 10$ ) — фронтальные слайсы (б) (черным цветом выделены главные диагонали матриц) и тензорная модель одномерного ВР (в)

**Самоподобие многомерных ВР.** Современные исследования трафика компьютерных систем (КС) направлены на решение следующих проблем: повышение эффективности использования КС, в частности увеличение пропускной способности;

идентификация аномалий в процессе их зарождения для возможности адекватной реакции и недопущения возможных негативных последствий.

В большинстве случаев рассматривается одномерный трафик, т.е. только компоненты, имеющие числовые форматы: число переданных байт, число пакетов, число потоков; различные источники пакетов рассматриваются в

2D пространстве, однако в общем случае пакеты, объединяемые в потоки, образуют многомерное пространство. Естественно, при этом упрощаются расчеты, но полученные результаты далеки от реальных.

Рассмотрим явление самоподобия трафика, исследуемого на уровне отдельного ВР. Как видно из матрицы фрагмента реального трафика (рис. 7), самоподобие присуще всему трафику, а не только его некоторым компонентам. Следует заметить, что до настоящего времени самоподобие трафика как многомерного ВР не рассматривалось, хотя необходимость в этом весьма остра.

Известно, что трафик КС как совокупность ВР представляет собой стохастические объекты, для которых свойство самоподобия наблюдается только после соответствующего усреднения по всем независимым реализациям объекта (процесса). Иными словами, увеличенная часть фрагмента исходного ВР и сам фрагмент самоподобны, так как имеют совпадающие (или в общем случае близкие) некоторые статистические характеристики. Понятие точное самоподобие — условно, так как справедливо только для регулярных объектов (построенных детерминированным способом).

Повторяемость (проявление самоподобия) нескольких реализаций ВР (трафика) может быть полной, но могут наблюдаться элементы случайности. В общем случае структура ВР (трафика) на малых масштабах не является идентичной структуре всего ВР, но их статистические характеристики достаточно близки, хотя полное совпадение практически отсутствует. Как следует из [19], структура ВР не определена и анализ ВР на малых масштабах недостаточен для анализа самоподобия. В ряде случаев возникают ситуации, когда необходимо исследовать весь разреженный ВР, т.е. когда элементы взяты с меньшей (чем исходная) частотой. Вычисление  $H$ -параметра для ВР  $X_t(1:10000)$  и ВР  $X_t(1:10:10000)$ ,  $X_t(1:50:10000)$ ,  $X_t(1:100:10000)$  дает практически одинаковые результаты, что позволяет сделать вывод о персистентности ВР. Если учесть, что  $X_t(1:10000) \rightarrow T^x = \text{reshape}(X, 100, 100)$ , то  $\text{diag}(T^x)$  представляет собой разреженный исходный ВР.

Известно, что свойство самоподобия с  $H$ -параметром формально имеет вид равенства по определению:  $\{X(\alpha t), t \in \mathbb{R}\} \stackrel{\Delta}{=} \{\alpha^H X(t), t \in \mathbb{R}\}$ . При этом изменение временного масштаба эквивалентно изменению пространственного масштаба состояний, а это означает, что типичные реализации самоподобного процесса визуально похожи (близки), независимо от масштаба времени, на котором они рассматриваются, т.е. наблюдается сходство статистических свойств. Однако равенство

$$Y_k^{(m)} = (1/m^H) \sum_{i=(k-1)m+1}^{km} Y_i, \quad k \in R, \quad 0 < H < 1$$

справедливо для агрегированного процесса, при котором  $m$ -агрегированный ВР  $Y^{(m)}$  представляет собой усредненный по неперекрывающимся блокам размера  $m$  (длина окна ВР) исходный ВР, т.е. каждый блок заменен его средним значением.

Среди всех параметров трафика (в общем случае их насчитывают около 100), как было указано ранее, наиболее исследованы параметры, имеющие числовые форматы: байты, пакеты, потоки пакетов, время задержки. Они образуют, в общем случае, многомерную систему, анализ которой до настоящего времени не проведен. Поэтому анализ самоподобия даже этого усеченного многомерного трафика представляет собой актуальную и достаточно сложную теоретическую и прикладную задачу, требующую применения современного математического аппарата — тензорного анализа и тензорных декомпозиций.

Покажем возможность анализа многомерного трафика на уровне плоских моделей. Заметим, что в общем случае анализ самоподобия ВР на основе  $H$ -параметра не является тривиальной задачей, так как реальные НД — конечны, что противоречит исходной концепции вычисления  $H$ -параметра. Поэтому необходимо использовать различные свойства ВР, влияющие на самоподобие. В некоторых работах [19] исследуется самоподобная структура в заданном масштабном диапазоне для заданного набора данных (при этом понятие структуры не определено). Кроме того, наличие самоподобной структуры в одной из компонент ВР (трафика КС), не гарантирует самоподобие всего (многомерного) ВР.

Показатель (параметр) Херста является параметром, характеризующим «степень» самоподобия. Он определяется для выборочного случайного набора  $X_j, j = 1, N$  на основе выборочного среднего,

$$M = \frac{1}{N} \sum_{k=1}^N X_k,$$

дисперсии,

$$S_N^2 = \frac{1}{N} \sum_{k=1}^N (X_k - M)^2,$$

интегрального отклонения,

$$D_j = \sum_{k=1}^j X_k - jM,$$

и изменчивости  $R_N = \max_{1 < j < N} D_j - \min_{1 < j < N} D_j$  с учетом эвристики вида  $R/S \approx (N/2)^H$ , сформулированной Г. Херстом для больших значений  $N$

виде  $\log(R/S) \approx H \log(N/2)$ , где величина  $H$  — показатель Херста,  $0,5 < H < 1$ . Для процессов, не обладающих свойством самоподобия,  $H = 0,5$ , для самоподобных процессов с долгосрочной зависимостью  $H \in [0,7; 0,9]$ . Параметр  $\beta$ , используемый для характеристики асимптотических свойств самоподобных процессов, имеет вид  $\beta = 2 - H$ . Параметр Херста можно определить, построив график отношения  $\log(R/S)$  в зависимости от  $\log(N/2)$  при различных значениях  $N$ , и вычислить величину  $H$  как тангенс угла наклона полученной (аппроксимированной) линии (так называемый  $R/S$ -метод). Следует заметить, что  $R/S$ -метод дает приближенное значение показателя Херста, в общем случае выводы следует делать, учитывая только два варианта:  $H < 0,5$  и  $H > 0,5$ .

В рассматриваемой задаче предполагается, что в одномерном случае самоподобие (фрактальные свойства) ВР, определенные на всем ВР, справедливы для меньшего числа элементов ВР, расположенных на главной диагонали тензора, моделирующего ВР, т.е. ⟨фрактальные свойства ВР — фрактальные свойства главной диагонали  $\mathbb{T}_X$ ⟩ (разреженного ВР),  $\mathbb{T}_X = \text{reshape}(X, m, n, k)$ ,  $X$  — ВР,  $X = \{x_i\}$ ,  $i = 1, m \cdot n \cdot k$ .

При исследовании одномерных ВР заменяют весь ВР и его окна  $2D$  матрицами, используя процедуру  $\text{reshape}(X, m, n, g)$ , где  $m \times n$  — размерность матрицы,  $m \cdot n$  — число элементов ВР  $X = \{x_i\}$ ,  $i = 1, f$ ;  $f = m \cdot n$ ,  $X \rightarrow \rightarrow X$ ,  $g$  — число  $2D$  матриц фронтальных слайсов. Многомерные (многокомпонентные) ВР  ${}^{(m)}X$  заменяют  $3D$  тензорами  $\Xi$  (многомерными массивами), которые в свою очередь, заменяют  $2D$  матрицами-слайсами (Матлаб —  $\text{reshape}({}^{(m)}X, m, n, k)$ ), используя процедуры тензорной декомпозиции:  ${}^{(m)}X \xrightarrow{\text{ТД}} \Xi$ . В общем случае  $f = m \cdot n \cdot k$ ,  $\Xi = \{\Xi(:, :, 1), \dots, \Xi(:, :, k)\}$ ,  $k$  — число слайсов.

Вместо исходных объектов  $2D$  матрицы  $X$  и  $3D$  тензора  $\Xi$  предложено использовать их диагонали  $\text{diag}(X)$  или  $\text{diag}(\Xi)$ , или соответствующие диагональные матрицы,  $X^{(d)}$  и  ${}^{(m)}X^{(d)}$ , которые в развернутом (векторизованном) виде представляют разреженные ВР и число элементов которых значительно меньше, чем в исходных ВР. Возникает вопрос — насколько совпадают свойства  $X^{(d)}$  и  $X$ ,  ${}^{(m)}X^{(d)}$  и  ${}^{(m)}X$ , а именно, показатели самоподобия Херста. В трафике КС наиболее исследуемыми являются ВР для байтов, пакетов, потоков, которые исследуются независимо. Однако в ряде случаев возникает необходимость исследования этих рядов совместно, особенно при идентификации аномалий в этих ВР [17—19].

**Результаты экспериментальных исследований показателя Херста для исходных и разреженных ВР, построенных на главных диагоналях тензорных моделей ВР. I. Исходный ВР  $x = 3 * \text{rand}(1, 9000)$**



моделируется 3D тензором  $\Xi_t = \text{reshape}(x(1 : 9000), p, p, n)$ ,  $p = 10, n = 90$ , где  $n$  — число фронтальных слайсов. На рис. 7 представлены фрагменты ВР (отдельные слайсы), их диагонали и показан способ образования блочно-диагонального тензора, моделирующего ВР.

Два фронтальных слайса из 90:  $\Xi_t = \{x_t(:, :, 1), x_t(:, :, 2), \dots, x_t(:, :, 90)\}$ :

$x_t(:, :, 1) =$									
2.0429	1.7937	1.4738	0.8544	0.5365	0.0823	1.9552	2.9562	0.9656	2.8898
1.0493	2.8814	1.8151	0.6869	2.1259	0.9915	1.6595	1.3845	1.6691	0.9737
0.0991	1.7534	1.9429	0.3242	2.3369	2.1610	1.7793	0.9593	1.7642	2.1801
1.8012	2.3418	2.7258	2.1078	2.1833	1.0913	2.9665	2.7456	1.7497	0.4115
0.6737	2.7196	0.1360	0.2096	1.4379	0.1914	2.9866	1.7015	0.9177	2.7755
1.4717	1.3154	1.0183	1.1575	0.9331	0.2434	1.5832	1.5136	0.7616	2.7301
1.0636	0.6223	1.7299	1.3302	1.5616	2.1670	2.7228	2.2539	2.0572	1.4531
1.2147	2.6664	1.9870	1.1148	0.0277	0.0413	2.6989	1.9698	2.2360	2.8548
0.9229	0.6194	2.4955	1.5858	1.3910	2.9384	2.9853	1.1609	2.7590	0.7385
2.0658	0.9236	2.4348	1.3434	1.7012	2.3671	0.3643	1.8287	1.7117	0.4731
$x_t(:, :, 90) =$									
0.1850	2.6130	0.2860	1.1432	1.7478	1.9967	0.8816	2.1003	0.7812	0.1776
0.7973	0.3831	2.7210	1.1523	1.9397	0.5350	1.8029	0.5548	1.2125	2.7766
0.0368	2.2839	0.3657	0.6249	0.4780	1.2176	1.2235	2.4681	2.6059	2.3435
1.0786	1.6008	0.5521	1.2895	1.4922	1.6961	2.3135	2.4030	0.4630	1.3651
1.2132	1.2232	2.2903	2.2355	0.1979	0.0194	2.9464	2.0801	1.0949	2.4172
2.9970	0.4858	0.9821	2.1028	1.9272	1.5667	1.6849	2.3483	2.1977	0.5531
0.3924	1.7212	0.7470	2.3143	1.2004	1.9280	2.4283	0.4330	2.1634	0.2594
1.0189	0.5221	0.6410	1.6526	0.9138	1.4780	0.6375	1.8728	0.4605	2.3940
2.2635	2.1185	2.7841	0.0054	2.4881	2.2771	0.1057	2.2432	0.7733	1.4869
1.3702	2.3583	0.6999	0.6542	2.3059	0.4062	0.9973	2.4208	2.2964	0.8173

Таким образом, главная диагональ  $\Xi_t$  имеет 900 элементов исходного ВР, соответственно контрольный ВР принят в виде  $x_k = 3 \cdot \text{rand}(1, 900)$ . На рис. 8 показан исходный одномерный ВР и главная диагональ тензора  $\Xi_t$ .

Приведенные на рис. 9 показатели Херста\* для исходного, разреженного (диагональ тензора) и контрольного ВР, свидетельствуют о том, что значения  $H_x$  и  $H_{\text{out}}$  практически совпадают. Следовательно, главная диагональ тензора, моделирующего ВР, адекватно представляет свойство самоподобия ВР.

II. Многомерный (многокомпонентный) ВР был исследован по предложенной методике, визуализация которой представлена на рис. 9—11.

\* Использована подпрограмма Матлаб: function [Hal, He, Ht, pval95] = hurst(x, d, fontsize), R. Weron (2002) Estimating long range dependence: finite sample properties and confidence intervals, Physica A 312, 285-299.

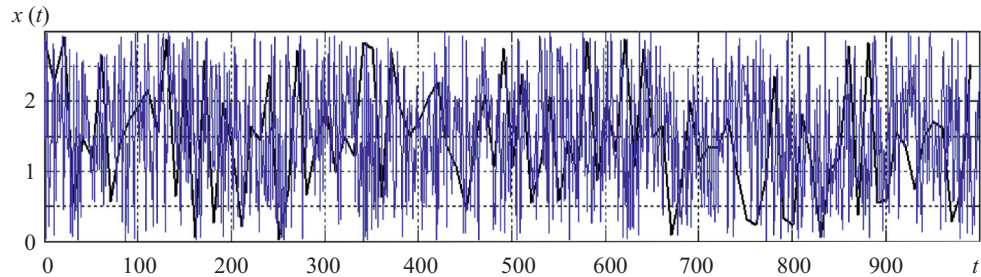


Рис. 8. Исходный одномерный ВР (синий цвет) и диагональ 3D тензора, моделирующего ВР (черный цвет)

Фрагмент исходного пятикомпонентного ВР  $X = \{x_1 \ x_2 \ x_3 \ x_4 \ x_5\}$ :

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
1.5990	1.6598	0.8308	0.4570	1.3824
0.4203	1.6100	0.7503	1.6431	1.3391
0.1713	0.1143	0.2767	0.0017	0.2768
0.7356	0.9492	0.8279	0.8151	1.2478
1.7094	0.7392	1.3004	0.4132	1.2852
0.4259	0.4738	0.2846	0.4355	0.5517
0.5998	0.6923	1.0897	0.6891	0.7204
0.0366	1.5976	1.6160	1.6901	1.8034
0.6770	0.3243	0.1815	0.6466	0.5457
0.4839	1.0464	0.0995	0.5892	0.9136
1.2309	0.0602	0.9052	0.0081	0.4049
0.8499	1.0235	0.8222	0.8224	0.8622
0.7746	0.6003	0.2211	0.2078	0.6028
1.4625	1.2810	1.8688	1.3184	1.5107
0.2292	0.5809	0.6032	0.1416	0.1363
0.2336	0.1006	0.2818	0.0323	0.3773
1.7987	1.6060	1.3992	0.3343	0.8575
1.3772	1.4571	1.3714	0.3296	0.8399
0.1230	0.0405	0.1660	0.0755	0.1001
1.6317	0.4583	0.8158	0.9447	0.6002
...	...	...	...	...

Протокол вычисления коэффициента Херста:  
R/S-AL using 1 divisors (50, ..., 50) for a sample of 100 values  
Corrected theoretical Hurst exponent 0.5  
Corrected empirical Hurst exponent 0.5887  
Theoretical Hurst exponent 0.5229  
Empirical Hurst exponent 0.5581

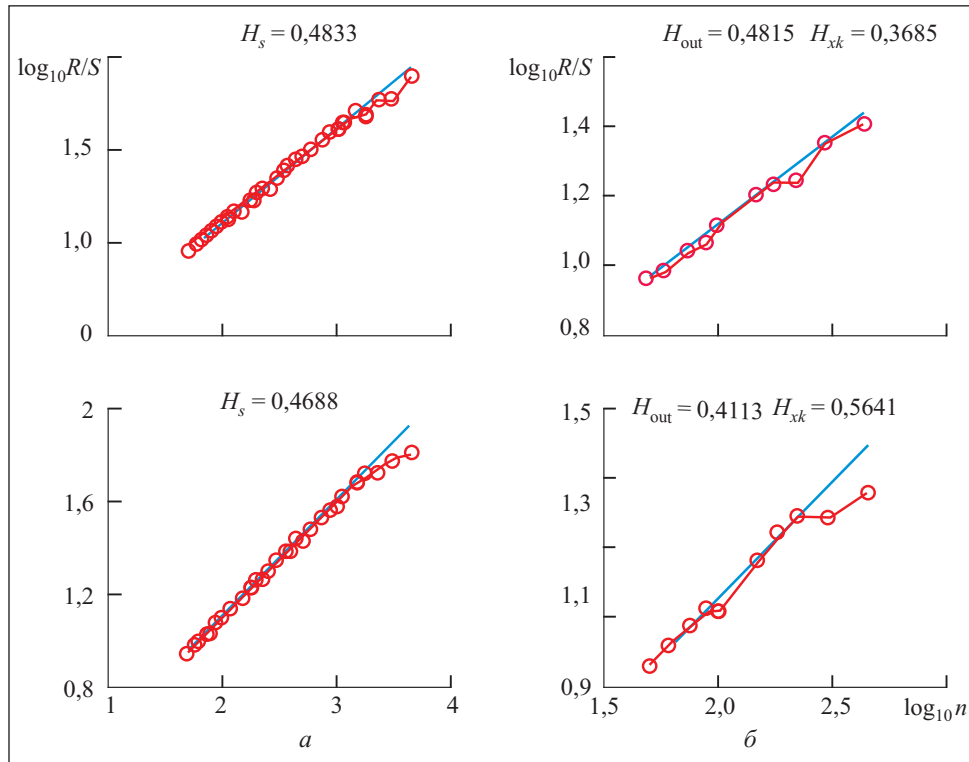


Рис. 9. Показатели Херста:  $a$  — исходный ВР,  $b$  — диагональ тензор-окна: синий цвет — теоретическая величина  $R/S$ , красный цвет — эмпирическая величина  $R/S$

Представление 200 элементов пятикомпонентного ВР фронтальными слайсами в виде двумерной матрицы  $10 \times 10 \rightarrow zt = \text{reshape}(zm(1 : 200, :, :, 1), p, p, n)$ :

$zt(:, :, 1) =$									
1.5990	1.2309	0.0969	0.0123	0.2659	0.2593	0.9595	1.3830	1.0013	0.0289
0.4203	0.8499	0.0062	0.8236	0.0360	1.3462	1.3302	0.1570	0.5994	0.0183
0.1713	0.7746	1.3762	0.8901	1.1333	0.2397	0.2750	1.1349	1.0121	1.6953
0.7356	1.4625	0.0179	0.9860	0.0134	0.6054	1.1867	1.1193	1.2906	0.0977
1.7094	0.2292	0.0368	0.3991	1.4156	0.3017	1.3526	0.2873	0.0609	0.4082
0.4259	0.2336	0.3092	0.8304	0.9069	0.7280	0.0351	0.7957	1.8095	1.3012
0.5998	1.7987	0.3801	1.0892	0.1753	0.3301	0.5851	1.0926	0.4462	0.2159
0.0366	1.3772	0.3271	0.3113	0.6590	1.7078	0.5203	0.4573	0.0179	0.5381
0.6770	0.1230	0.3613	0.3906	0.5811	1.0871	0.0783	0.6169	1.5064	0.1295
0.4839	1.6317	0.3929	1.0016	0.0995	0.3618	0.8266	1.2350	1.3183	1.0009
...	...	...	...	...	...	...	...	...	...

$zt(:, :, 10) =$

0.6642	0.5891	0.5596	0.7674	0.4898	0.1831	0.1929	1.6349	0.5907	0.4690
1.4026	1.5510	0.5247	0.0993	0.7668	0.0648	1.0127	0.3540	0.2614	0.6976
0.6185	0.1694	0.6284	0.3684	1.6953	0.4067	0.4418	0.1775	0.8141	0.4461
0.0082	0.2079	0.5931	0.3881	0.2782	0.0826	0.1265	0.1448	0.7785	0.3141
1.8018	0.9169	0.1417	0.0840	0.9201	1.7034	0.8466	0.5589	0.1238	0.0730
0.2242	0.3001	0.6576	1.2033	1.0016	0.3509	0.1492	0.1310	0.1580	0.1910
0.1671	1.2415	1.0170	1.1386	0.8652	0.1478	0.4764	0.2080	0.3235	1.5270
1.7719	0.9030	1.3574	0.0419	0.2514	0.5729	0.6140	1.7090	0.9203	0.0679
1.6188	0.3084	0.0439	0.9912	0.6503	1.5417	0.3100	0.0827	0.1775	1.0416
0.0192	0.7302	1.0989	0.4601	0.2810	0.0527	0.3524	0.9631	0.1850	0.3382

Протокол вычисления:

R/S-AL using 4 divisors (50, ..., 250) for a sample of 500 values

Corrected theoretical Hurst exponent 0.5

Corrected empirical Hurst exponent 0.4704

Theoretical Hurst exponent 0.5475

Empirical Hurst exponent 0.5189

Представление 1000 элементов пятикомпонентного ВР фронтальными слайсами в виде двумерной матрицы  $10 \times 10 \rightarrow zt = \text{reshape}(zm(1 : 200 * 5, :, :, 1), p, p, n * 5)$ . Фронтальные слайсы:

$zt(:, :, 1) =$

1.5990	1.2309	0.0969	0.0123	0.2659	0.2593	0.9595	1.3830	1.0013	0.0289
0.4203	0.8499	0.0062	0.8236	0.0360	1.3462	1.3302	0.1570	0.5994	0.0183
0.1713	0.7746	1.3762	0.8901	1.1333	0.2397	0.2750	1.1349	1.0121	1.6953
0.7356	1.4625	0.0179	0.9860	0.0134	0.6054	1.1867	1.1193	1.2906	0.0977
1.7094	0.2292	0.0368	0.3991	1.4156	0.3017	1.3526	0.2873	0.0609	0.4082
0.4259	0.2336	0.3092	0.8304	0.9069	0.7280	0.0351	0.7957	1.8095	1.3012
0.5998	1.7987	0.3801	1.0892	0.1753	0.3301	0.5851	1.0926	0.4462	0.2159
0.0366	1.3772	0.3271	0.3113	0.6590	1.7078	0.5203	0.4573	0.0179	0.5381
0.6770	0.1230	0.3613	0.3906	0.5811	1.0871	0.0783	0.6169	1.5064	0.1295
0.4839	1.6317	0.3929	1.0016	0.0995	0.3618	0.8266	1.2350	1.3183	1.0009

$zt(:, :, 50) =$

1.0446	0.0963	0.0027	0.6332	1.6667	1.2580	0.1646	0.2584	0.3211	1.9667
0.4805	1.4620	1.4778	0.6205	0.3388	0.1273	0.4174	0.0137	0.3234	1.1782
0.2883	1.0090	0.3140	0.1011	0.9473	1.3951	0.1385	0.2650	0.0204	0.2184
0.2925	0.2695	0.6966	1.1050	0.2146	0.5415	0.1007	0.0006	1.0225	0.7236
0.2350	0.6788	0.9516	0.4426	0.6459	0.0636	0.6375	0.5394	0.9548	1.4057
0.3503	1.4633	0.8935	0.3601	0.1530	1.4870	1.7269	0.1438	0.1652	0.0971
0.0503	0.2086	0.0686	0.9874	0.6878	0.1004	1.1109	0.0282	0.8190	0.4337

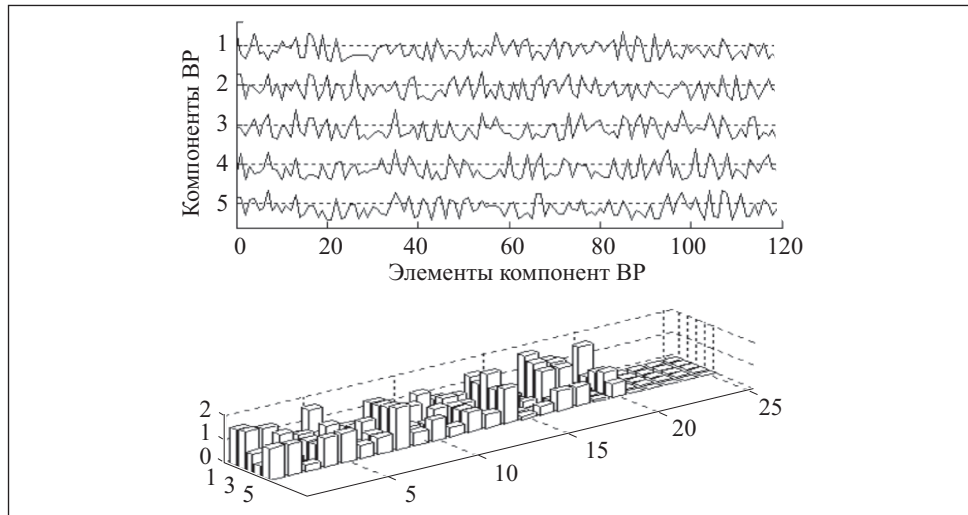


Рис. 10. Фрагмент исходного многокомпонентного ВР, представленный в двух формах

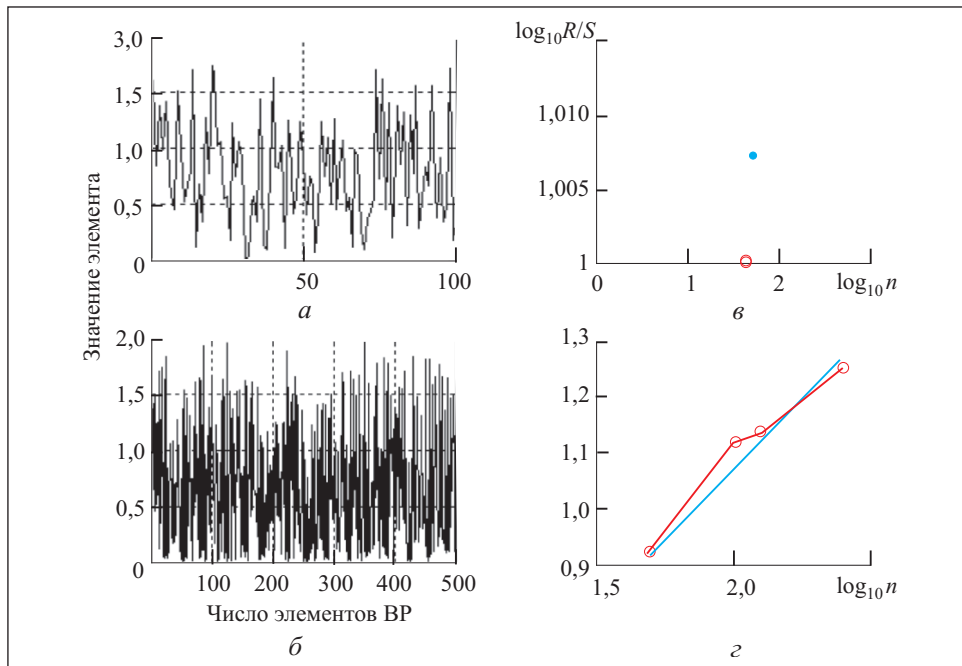


Рис. 11. Фрагменты МВР  $X(200 \times 5)$  (а),  $X(1000 \times 5)$  (б) и коэффициент Херста ( $\nu, z$ ):  $a$ —диагональ тензора, полученного в результате представления фрагмента МВР ( $200 \times 5$ ) фронтальными слайсами (матрицами  $10 \times 10$ , число слайсов – 10);  $b$ — диагональ тензора, полученного в результате представления фрагмента МВР ( $1000 \times 5$ ) фронтальными слайсами (матрицами  $10 \times 10$ , число слайсов — 50); ( $\nu, z$ ) — теоретический (синий цвет) и экспериментальный (красный цвет) результаты

0.0463 0.0227 1.8094 1.6889 1.2068 0.3292 0.1390 1.1350 1.1100 0.8406  
0.0728 0.5640 0.4546 0.6479 0.3883 0.7321 1.8260 0.0380 0.0963 0.0446  
1.7169 0.8697 0.1296 0.2412 0.7810 1.3312 0.2653 0.0957 0.4622 0.6082

Таким образом, многомерный ВР, рассматриваемый как совокупность окон, может быть адекватно представлен совокупностью тензор-окон: отдельное окно может быть одномерным (сформированным для отдельной компоненты МВР), либо многомерным (сформированным с учетом всех компонент МВР). В случае I окно преобразуется либо в  $2D$  тензор с матрицей  $t_{x_k} = \text{reshape}(x_k, m, n)$ ,  $m \cdot n = f$ , либо в  $3D$  тензор с матрицей  $t_{x_k} = \text{reshape}(x_k, m, n, n_1)$ ,  $m \cdot n \cdot n_1 = f$ . Инварианты вычисляются по стандартным (для плоских матриц) формулам либо по формулам, справедливым только для пространственных матриц. В случае II окно преобразуется в  $3D$  тензор с матрицей  $t_{x_k} = \text{reshape}(x_k, m, n, n_1)$ ,  $m \cdot n \cdot n_1 = f$ . Инварианты вычисляются по формулам, справедливым только для пространственных матриц.

## Выводы

Показатель Херста для многомерного ВР является в значительной степени формальным, так как определить похожесть, даже явную, как в одномерных ВР, в данном случае весьма сложно по причине субъективности понятия похожесть. Однако интеллектуальный анализ данных, в данном случае ВР, позволил определить скрытые знания, которые могут быть использованы при решении прикладных задач: идентификация аномалий реального (многомерного) трафика, повышение пропускной способности и других.

## СПИСОК ЛИТЕРАТУРЫ

1. *Time series: Advanced methods* IIa. Multivariate time series. [Электрон. ресурс]: [www.ucl.ac.uk/jdi/events/int-CIA-conf/ICIAC11\\_Sli-des/ICIAC11\\_1E\\_LTompson](http://www.ucl.ac.uk/jdi/events/int-CIA-conf/ICIAC11_Sli-des/ICIAC11_1E_LTompson).
2. *Cichocki A.* Tensor decompositions for Signal Processing Applications. From Two-way to Multiway Component Analysis. Mandic D., Phan A-H., Caiafa C. [Электрон.ресурс]: <http://www.commsp.ee.ic.ac.uk/~mandic/SPM-Cichocki-Mandic-DeLathauwer.pdf>
3. *Соколов Н.П.* Пространственные матрицы и их приложения. М: Гос. изд-во физ.-мат. лит-ры, 1960, 299 с.
4. *Claude Z.B.* Introduction to the General Multidimensional Matrix in Mathematics. [Электрон. ресурс]: [www.ijera.com/pages/v3no6.html](http://www.ijera.com/pages/v3no6.html)
5. *Solo A.* Multidimensional Matrix Mathematics: Notation, Representation, and Simplification, Parts: 1-6. Proceedings of the World Congress on Engineering (3). [Электрон. ресурс]: [www.ijera.com/papers/Vol.3\\_issue6/U36123129.pdf](http://www.ijera.com/papers/Vol.3_issue6/U36123129.pdf)
6. *De Lathauwer L., Moor B.* From matrix to tensor: Multilinear algebra and signal processing 4-h IMA Int. Conf. on Mathematics in Signal Processing, Oxford, United Kingdom, 1998: Selected papers presented at pp. 1-15. J. McWhirter (Ed.), Mathematics in Signal Processing IV, Oxford University Press, 1998.

7. Skillicorn D. Understanding complex datasets : data mining with matrix decompositions. Chapman & Hall/ CRC Taylor & Francis Group 6000 Broken Sound Parkway NW. Suite 300 Boca Raton, FL 33487, 2742, 257 p.
8. Cichocki A. Tensor Decompositions: A New Concept in Brain Data Analysis? [Электрон. ресурс]: arXiv:1305.0395v1 [cs.NA] 2 May 2013, 19 p.
9. Lim L.-H. Singular values and eigenvalues of tensors: A variational approach. 1-st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), December 13-15, 2005: Proceedings of the Workshop, 2005, p. 129-132.
10. Liqun Qi. Eigenvalues and invariants of tensors J. Math. Anal. Appl. 325 (2007), 1363-1377. [Электрон. ресурс]: [www.elsevier.com/locate/jmaa](http://www.elsevier.com/locate/jmaa).
11. Kolda T.G., Bader B.W. Tensor Decompositions and Applications / SIAM Review, 2009, Vol. 51, N 3, P. 455-500.
12. Kamalja K.K., Khangar N.V. Singular Value Decomposition for Multidimensional Matrices. Int. Journal of Engineering Research and Applications, 2013, Vol. 3, Issue 6, p. 123-129.
13. Bader B.W., Kolda T.G. Tensor Decompositions, the MATLAB Tensor Toolbox, and Applications to Data Analysis. [Электрон. ресурс]: [www.sandia.gov/~tgkolda/TensorToolbox](http://www.sandia.gov/~tgkolda/TensorToolbox).
14. Bader B.W., Kolda T.G. Multilinear operators for higher-order decompositions: Technical Report SAND 2006-2081 / Sandia National Laboratories, April 2006. [Электрон. ресурс]: [pubs/pubfiles/SAND2007-6702.pdf](http://pubs/pubfiles/SAND2007-6702.pdf).
15. Stegeman A. The Parafac Model for Multi-way Data Analysis. [Электрон. ресурс]: <http://www.ppsw.rug.nl/~stegeman>.
16. Tensor Toolbox is Software for Working with Multidimensional Arrays. [Электрон. ресурс]: <http://csmr.ca.sandia.gov/~tgkolda/>
17. Kindlmann G. Tensor Invariants and their Gradients. [Электрон. ресурс]: [USA.gk@cs.utah.edu](mailto:USA.gk@cs.utah.edu).
18. Божокин С.В., Паршин Д.А. Фракталы и мультифракталы. Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001, 128 с.
19. Шелухин О.И. Мультифракталы. Инфокоммуникационные приложения. М.: Горячая линия – Телеком, 2011, 576 с.

Поступила 28.04.16

## REFERENCES

1. Time series: Advanced methods IIa. Multivariate time series, available at: [www.ucl.ac.uk/jdi/events/int-CIA-conf/ICIAC11\\_Sli-des/ICIAC11\\_1E\\_LTompson](http://www.ucl.ac.uk/jdi/events/int-CIA-conf/ICIAC11_Sli-des/ICIAC11_1E_LTompson).
2. Cichocki, A., Mandic, D., Phan, A.-H., Caiafa, C., Tensor decompositions for signal processing applications. From two-way to multiway component analysis, available at: <http://www.commsp.ee.ic.ac.uk/~mandic/SPM-Cichocki-Mandic-DeLathauwer.pdf>
3. Sokolov, N.P. (1960), *Prostranstvennyye matritsy i ikh prilozheniya* [Space matrices and their application], Gosudarstvennoe izdatelstvo fiz.-mat. literatury, Moscow, USSR.
4. Claude, Z.B., Introduction to the general multidimensional matrix in mathematics, available at: [www.ijera.com/pages/v3no6.html](http://www.ijera.com/pages/v3no6.html).
5. Solo, A., Multidimensional matrix mathematics: Notation, representation, and simplification. Parts: 1-6, *Proceedings of the World Congress on Engineering* (3), available at: [www.ijera.com/papers/Vol.3\\_issue6/U36123129.pdf](http://www.ijera.com/papers/Vol.3_issue6/U36123129.pdf).
6. De Lathauwer, L. and Moor, B. (1998), From matrix to tensor: Multilinear algebra and signal processing, *Proceedings of the 4-h IMA Int. Conf. on Mathematics in Signal Processing*, Oxford, United Kingdom, Selected papers presented at pp. 1-15, J. McWhirter (Ed.), Mathematics in Signal Processing IV, Oxford University Press.

7. Skillicorn, D., Understanding complex datasets : data mining with matrix decompositions. Chapman & Hall/ CRC Taylor & Francis Group 6000 Broken Sound Parkway NW. Suite 300 Boca Raton, FL 33487, 2742.
8. Cichocki, A. (2013), Tensor decompositions: A new concept in brain data analysis?, available at: arXiv:1305.0395v1 [cs.NA] May 2, 2013.
9. Lim, L.-H. (2005), Singular values and eigenvalues of tensors: A variational approach, *Proceedings of the 1-st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, December 13-15, 2005, pp. 129-132.
10. Liqun, Qi. (2007), Eigenvalues and invariants of tensors, *J. Math. Anal. Appl.*, Vol. 325, pp. 1363-1377, available at: [www.elsevier.com/locate/jmaa](http://www.elsevier.com/locate/jmaa).
11. Kolda, T.G. and Bader, B.W. (2009), Tensor decompositions and applications, *SIAM Review*, Vol. 51, no. 3, pp. 455-500.
12. Kamalja, K.K. and Khangar, N.V. (2013), Singular value decomposition for multidimensional matrices, *Int. Journal of Engineering Research and Applications*, Vol. 3, Iss. 6, pp. 123-129.
13. Bader, B.W. and Kolda, T.G. Tensor decompositions, the MATLAB tensor toolbox, and applications to data analysis, available at: [www.sandia.gov/~tgkolda/TensorToolbox](http://www.sandia.gov/~tgkolda/TensorToolbox).
14. Bader, B.W. and Kolda, T.G. (2006), Multilinear operators for higher-order decompositions: Technical report SAND 2006-2081, Sandia National Laboratories, available at: [pubs/pubfiles/SAND2007-6702.pdf](http://pubs/pubfiles/SAND2007-6702.pdf).
15. Stegeman, A., The Parafac model for multiway data analysis, available at: <http://www.ppsw.rug.nl/~stegeman>.
16. Tensor toolbox is software for working with multidimensional arrays, available at: <http://csmr.ca.sandia.gov/~tgkolda>.
17. Kindlmann, G., Tensor invariants and their gradients, available at: [USA.gk@cs.utah.edu](mailto:USA.gk@cs.utah.edu).
18. Bozhokin, S.V. and Parshin, D.A. (2001), *Fraktaly i multifraktaly* [Fractals and multi fractals], NITs Reguljarnaya i khaoticheskaya dinamika, Izhevsk, Russia.
19. Shelukhin, O.I. (2011), *Multifraktaly. Infokomunikatsionnyie prilozheniya* [Multifractals. Infocommunication applications], Goryachaya liniya – Telekom, Moscow, Russia.

Received 28.04.16

*Yu.N. Minaev, N.N. Guziy, O.Yu Filimonova., J.I. Minaeva*

#### ANALYSIS OF SELF-SIMILARITY OF MULTIVARIATE TIME SERIES (TS) ON THE BASIS OF THE METHODS OF INTELLECTUAL ANALYSIS OF THE DATA

Calculation methods have been proposed for the Hurst factor for univariate and multivariate TS on the basis of the main diagonals of TS tensor models. It is shown that the problem complexity determines the joint use of several mathematical theories, in particular, the tensor and multivariate matrix analysis. The examples of using the proposed methods are presented.

*Keywords: tensor, multivariate time series, intellectual analysis of the data, 3D matrices, diagonal 3D matrices, matrix development, self-similarity, Hurst parameter.*

*МИНАЕВ Юрий Николаевич, д-р техн. наук, профессор кафедры компьютерных систем и сетей Национального авиационного университета Украины. В 1959 г. окончил Харьковский политехнический ин-т. Область научных исследований — интеллектуальный анализ данных, применение интеллектуальных технологий в системах принятия решений.*



*ГУЗИЙ Николай Николаевич, канд. техн. наук, профессор, зам. зав. кафедры компьютерных систем и сетей Национального авиационного университета Украины. В 1972 г. окончил Киевский ин-т гражданской авиации. Область научных исследований — защита информации в компьютерных системах, системы искусственного интеллекта.*

*ФИЛИМОНОВА Оксана Юрьевна, канд. техн. наук, доцент Киевского национального университета строительства и архитектуры. В 1989 г. окончила Киевский инженерно-строительный ин-т. Область научных исследований — интеллектуальный анализ данных.*

*МИНАЕВА Юлия Ивановна, канд. техн. наук, и.о. доцента кафедры основ информатики Киевского национального университета строительства и архитектуры, который окончила в 2008 г. Область научных исследований — интеллектуальный анализ данных.*

