

УДК 004.021

DOI: 10.25559/SITITO.14.201803.560-566

## ЭФФЕКТИВНЫЙ ПО ВРЕМЕНИ АЛГОРИТМ РАСЧЕТА ОБОБЩЕННОЙ ЭНТРОПИИ ДВУМЕРНЫХ СЛОВ МЕТОДОМ СКОЛЬЗЯЩЕГО ОКНА

И. Седакин<sup>1</sup>, М.В. Ульянов<sup>2,1</sup><sup>1</sup> Московский государственный университет имени М.В. Ломоносова, г. Москва, Россия<sup>2</sup> Институт проблем управления им. В.А. Трапезникова РАН, г. Москва, Россия

## TIME EFFICIENT ALGORITHM FOR THE GENERALIZED ENTROPY CALCULATION OF TWO-DIMENSIONS WORDS BY THE METHOD OF A SLIDING WINDOW

Ilya Sedyakin<sup>1</sup>, Michael V. Uljanov<sup>2,1</sup><sup>1</sup> Lomonosov Moscow State University, Moscow, Russia<sup>2</sup> V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

© Седакин И., Ульянов М.В., 2018

### Ключевые слова

Энтропия двумерных слов;  
самоорганизация.

### Аннотация

В статье, применительно к задаче самоорганизации стержнеобразных частиц на торе, рассматриваются методы вычисления энтропии двумерных слов конечной длины над конечным алфавитом. Наблюдение за изменением энтропии позволяет лучше понять процессы, происходящие при самоорганизации. Для вычисления энтропии используется метод скользящего окна, который заключается в нахождении уникальных конфигураций окон в матрице и подсчете числа уникальных конфигураций. Затем по полученным значениям счетчиков вычисляется энтропия. Основную временную сложность требуют алгоритмы подсчета количества окон, само значение энтропии по известным счетчикам вычисляется быстро. Потому были предложены варианты асимптотически эффективных по времени алгоритмов подсчета количества одинаковых окон, использующих методы преобразования в индекс, дерево поиска и префиксное дерево. При этом энтропия всегда рассчитывается одинаковым способом. Приводятся результаты экспериментального исследования времени работы программных реализаций этих алгоритмов: минимальное, максимальное и среднее время работы на ряде экспериментов для некоторых размеров окна. Сделан вывод об эффективности тех или иных алгоритмов в зависимости от размеров окна. Обосновывается нерепрезентативность энтропии, вычисленной прямым путем, из-за несоразмерно большого количества возможных существующих уникальных конфигураций, даже при малых размерностях алфавита и окна. Предлагается метод быстрого вычисления обобщенной энтропии двумерных слов вместо прямой энтропии.

### Keywords

Entropy of two-dimensions;  
self-organization.

### Abstract

In the article we consider methods for calculating the entropy of two-dimensional words of finite length over in finite alphabet, with reference to the problem of self-organization of rod-like particles on a torus. Monitoring the change in entropy allows a better understanding of the processes occurring during self-organization. To calculate the entropy, we use the sliding window method, which consists in finding unique configurations of windows in the matrix and counting the number of these configurations. Then, the entropy is calculated from the obtained counter values. The basic time complexity is required by the algorithms for counting the number of windows. Value of entropy is calculated quickly using known counters. Therefore, we propose variants of asymptotically time efficient algorithms for counting the number of identical win-

### Об авторах:

**Ульянов Михаил Васильевич**, доктор технических наук, профессор, ведущий научный сотрудник, Институт проблем управления им. В.А. Трапезникова РАН; (117997, Россия, г. Москва, ул. Профсоюзная, д. 65); профессор кафедры алгоритмических языков, факультет вычислительной математики и кибернетики, Московский государственный университет имени М.В. Ломоносова (119991, Россия, ГСП-1, г. Москва, Ленинские горы, д. 1), ORCID: <http://orcid.org/0000-0002-5784-9836>, [muljanov@mail.ru](mailto:muljanov@mail.ru)

**Седакин Илья**, студент, факультет вычислительной математики и кибернетики, Московский государственный университет имени М.В. Ломоносова (119991, Россия, ГСП-1, г. Москва, Ленинские горы, д. 1), ORCID: <http://orcid.org/0000-0002-9118-8517>, [ilya.sedyakin@mail.ru](mailto:ilya.sedyakin@mail.ru)



dows using the methods of conversion into an index, a search tree and a prefix tree. In this case, entropy is always calculated in the same way. The results of an experimental study of the software operating time implementations of these algorithms are presented: the minimum, maximum, and average operating time on a number of experiments for certain window sizes. The conclusions about the effectiveness of various algorithms, depending on the size of the window were made. We show the non-representativeness of the entropy calculated by the direct path is due to the disproportionately large number of possible existing unique configurations, even with small dimensions of the alphabet and the window. Finally, we propose the method of quickly generalized entropy of two-dimensional words calculating instead of direct entropy.

## Введение

Вопрос устройства самоорганизующихся систем активно исследуется учеными по всему миру. В этой работе будет рассмотрена одна из таких систем. Хотя рассматриваемая система использует дискретное время и пространство, она моделирует образование структур в тонких слоях вытянутых гранул под воздействием вибрации.

Проводить глубокий анализ поведения таких систем и лучше понимать их динамику помогает использование энтропийных характеристик. Иногда графики изменения энтропии в системах наводят на интересные мысли. Расчёт энтропии для систем большой размерности может требовать значительных затрат вычислительных мощностей и времени. В данной работе будут приведены алгоритмы, позволяющие эффективно вычислять энтропию двумерных слов, будет проведен анализ и сравнение их ресурсных характеристик.

В своей работе «Образование структур в двумерных системах стержнеобразных частиц» Н.И. Лебовка, Ю.Ю. Тарасевич, В.А. Гигиберия, Н.В. Выгорницкий, А.С. Бурмистров, В.В. Лаптев рассматривают следующую задачу [1].

Пусть дана квадратная решетка размерности  $L \times L$ , состоящая из квадратных ячеек со стороной 1. Решетка является разверткой тора, то есть, она «циклическа» по горизонтали и вертикали (периодические граничные условия). Размер стороны решетки  $L$  может варьироваться от 128 до 2048. Начальное состояние системы создавалось с использованием алгоритма случайного последовательного осаждения частиц (RSA), который, условно, состоит в следующем [1]:

В начале эксперимента на этот тор размещались стержнеобразные частицы, представляющие из себя прямоугольники ширины 1 и длины  $k$ , которая варьировалась до 2 до 12. Частицы могли располагаться горизонтально и вертикально, и не могли пересекаться. Частицы размещались последовательно случайным образом до тех пор, пока не возникла ситуация, в которой на развертку тора нельзя положить более ни одной частицы. Эта ситуация называется состоянием джамминга [1].

После этого случайным образом выбиралась одна частица и предпринималась попытка сместить её на одну ячейку в произвольно выбранном направлении. Число попыток, равное полному числу частиц в рассматриваемой системе, принималось за один шаг моделирования Монте-Карло [2]. Моделирование типично продолжалось до  $10^7$  шагов.

В результате система приходила в состояние самоорганизации, связанной с образованием полосовых доменов. Частицы одинаковой направленности выстраивались в циклические диагональные полосы на торе. Это происходило только в случае, если длины стержней  $k \geq 6$ , поведение системы по времени моделирования в терминах взаимного числа контактов гори-

зонтальных и вертикальных частиц приведено на рис. 1 (источник рисунка [1]).

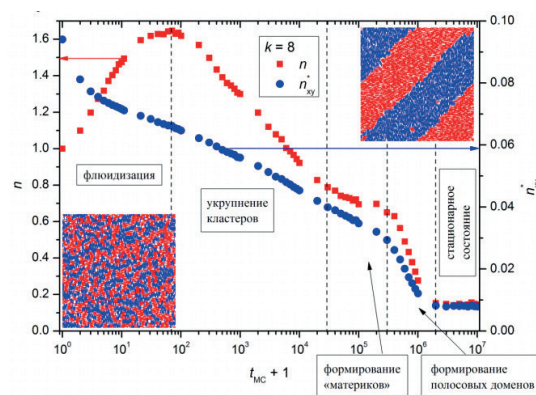


Рис. 1. Поведение системы при моделировании [1].

Fig 1. Behavior of the system in modeling [1]

## Цель исследования

Основной задачей исследования была разработка эффективных по времени алгоритмов вычисления энтропии двумерных слов. В рамках данной задачи определим понятие энтропии двумерных слов и вычислим ее значения для различных шагов моделирования. Введем следующее цифровое кодирование состояний ячейки развертки тора:

- 0 – ячейка пуста;
- 1 – через ячейку проходит вертикальная частица;
- 2 – через ячейку проходит горизонтальная частица.

Выберем на решетке окно — квадрат размера  $l \times l$  ячеек, тогда  $l^2$  цифр, расположенных в нем, создают собой двумерное слово  $\nu$ , составленное из символов алфавита мощности три. Теперь, пройдемся таким окном по всей матрице ( $m = L^2$  позиций в силу периодических граничных условий), и подсчитаем частотную встречаемость конфигураций. Пусть  $\eta_i$  — значение счетчика для  $i$ -ой конфигурации, при этом

$$\sum_{i=1}^M \eta_i = m$$

Расчет энтропии производится по аналогии с одномерным случаем, в котором полученные значения счетчиков интерпретируются как вероятности в некоторой вероятностной модели, и вычисляется энтропия соответствующего дискретного распределения [1, 2]. Теперь мы можем рассчитать значение функции информационной энтропии исследуемого двумерного слова  $\nu$  при окне  $l \times l$  в момент  $t_{MC}$  (текущий номер шага моделирования Монте-Карло) по формуле



$$H(v(t_{MC}), l, \alpha) = -\sum_{i=1}^M \left(\frac{\eta_i}{m}\right) \log_{\alpha} \left(\frac{\eta_i}{m}\right),$$

где  $\alpha$  — основание логарифма, обычно выбираемое равное двум [3, 4].

Подсчет энтропии методом скользящего окна — трудоемкая задача. Трудоемкость растет с увеличением размеров решетки, окна и количества оцениваемых шагов моделирования. Главную сложность составляет подсчет количества конфигураций встреченных слов. Таким образом, интерес представляет:

- исследование алгоритмов, применимых для подсчета частотной встречаемости двумерных слов и определение особенностей их реализации в рамках данной задачи;
- определение применимости выбранного способа вычисления энтропии к задаче самоорганизации стержнеобразных частиц.

## Алгоритмы вычисления энтропии и их особенности реализации

### Преобразование в индекс

Первая идея, связанная с решением поставленной задачи — некоторым образом преобразовать окно в индекс массива. Тогда поиск окон будет происходить за  $O(1)$ , и основная сложность алгоритма будет сосредоточена в преобразовании окна в значение индекса. Такой подход интуитивно кажется достаточно быстрым. При реализации данной идеи есть смысл разделить задачу на две части по размеру скользящего окна:

- если размер окна  $l = 4$ , то мы можем объявить в памяти массив размера  $3^{l^2}$ , и преобразовывать наше окно в индекс по следующему правилу: Возьмем поочередно каждую цифру  $\{0,1,2\}$  из окна как очередную цифру в троичной записи числа, обозначающего индекс. Переведем это число в десятичную систему счисления и получим наш индекс. Для того чтобы перевести окно в десятичную систему был использован следующий алгоритм: берется результат, изначально равный нулю. Затем для каждой цифры окна результат умножался на 3, после чего к нему прибавлялась эта цифра;
- если же размер окна больше чем 4, мы не сможем объявить столь большой массив. Тогда нам необходимо предложить хеш-функцию для соответствующего преобразования. Вычисление хеш-функции для квадратного массива — нетривиальная задача. Производительность такого алгоритма будет целиком зависеть от выбора хеш-функции, и этот вариант алгоритма не был реализован.

### Дерево поиска, состоящее из списков

Одна из эффективных структур поиска — сбалансированное дерево. Эта структура представляет собой бинарное дерево из элементов, которые обладают свойством сравнения, т.е. для любых двух элементов можно определить, какой из них больше. Каждый узел дерева хранит один из элементов и два поддерева

(левое и правое), которые могут быть пустыми. При этом соблюдается правило: в левом поддереве все элементы меньше элемента в узле, а в правом — больше. Стандартная библиотека шаблонов STL уже имеет хорошую реализацию такого дерева. Это шаблон `std::map`, который представляет собой сбалансированное (чёрно-красное) дерево поиска пар (элемент, значение) [5]. Эта реализация и была использована для экспериментов.

Заметим следующую особенность нашей задачи. Когда окно сдвигается, не все его элементы изменяются. К нему лишь добавляется новый ряд справа или снизу, и убирается соответствующий ряд слева или сверху. Все остальные же элементы остаются в окне, и лишь сдвигаются на одну позицию. Таким образом, лучший вариант реализации окна в этом случае — список списков.

Предположим такую реализацию: Внешний список хранит  $l$  списков строк, а каждый из  $l$  списков строк хранит  $l$  элементов, соответствующих столбцам (Рис.2). Тогда для того, чтобы сдвинуть окно влево или вниз нам будет достаточно  $O(l)$  операций, а не  $O(l^2)$ , так как список позволяет добавлять и удалять элементы за  $O(1)$ .

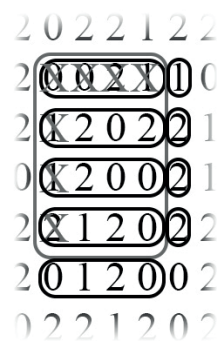


Рис. 2. Обработка ситуации сдвига скользящего окна.

Fig 2. Handling a sliding window shift situation

На сложность сравнения элементов в дереве это никак не повлияет — и массивы и списки сравниваются поэлементно. Поиск же элементов в дереве можно оценить как  $O(\log(L^2) \cdot l^2)$ .

### Дерево префиксов

Один из главных кандидатов на лучшее время работы — дерево префиксов [6]. Дерево префиксов представляет собой корневое дерево, каждое ребро которого помечено каким-то символом так, что для любого узла все ребра, соединяющие этот узел с сыновьями, помечены разными символами (Рис. 3).

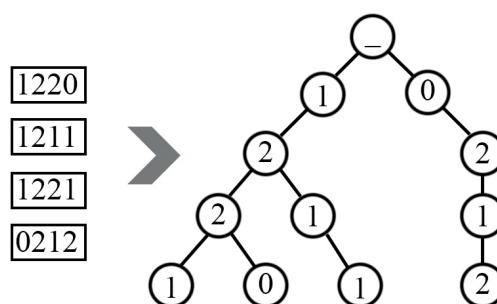


Рис. 3. Структура дерева префиксов.

Fig 3. Prefix tree structure



Таким образом, наши двумерные слова (которые можно так же представить, как одномерные разверткой по строкам) будут разложены по частям, и для того, чтобы прочитать любое из слов, необходимо будет пройти от корня дерева к одному из его листьев. Один из главных вопросов реализации префиксного дерева — каким образом хранить сыновей узла. Зачастую для этого используют упорядоченный массив (чтобы искать необходимого сына бинарным поиском). Было решено поступить следующим образом. Предположительно, лучший вариант реализации дерева — хранить в каждом узле по одной цифре окна. Это дает два преимущества:

- не нужно делать лишних преобразований элементов;
- можно хранить сыновей в виде простого массива размером 3, и получать доступ к ним за время  $O(1)$ .

В итоге, мы получаем следующий алгоритм: Для каждого из окон мы проходим по его элементам, одновременно с этим спускаясь по префиксному дереву. Дойдя до конца, увеличиваем счетчик того листа, в который пришли.

## Экспериментальное исследование

Тестирование алгоритмов производилось следующим образом. Программы, реализующие алгоритмы на языке C++ запускались на ноутбуке ASUS n76 с процессором Intel CORE i7 и оперативной памятью 8Gb. (Часть из этих ресурсов была затрачена на работу операционной системы Windows 10). Для чистоты эксперимента при проведении замеров из программ удалены участки кода, который, непосредственно, считал энтропию по массиву встретившихся слов. Как уже говорилось ранее, эта процедура производится очень быстро на любой структуре.

Для того чтобы максимально приблизить результаты к тем, которые были бы получены при измерении энтропии в задаче предметной области, замеры производились на результатах, полученных при моделировании этой задачи. Это 65 файлов с матрицами 256x256, соответствующие шагам временной динамики системы под номерами 1, 2, ..., 10, 20, ..., 100, 200, ..., ..., 1000000. (в ходе исследований предметной области было выявлено, что значительные изменения в матрице происходят с «логарифмическим разрывом» по времени между шагами.)

Так же, для всех программ и тестовых файлов замеры производились для разных размеров длины скользящего окна — 2, 4, 6, 8. Однако для метода преобразования в индекс замеры для  $l > 4$  не было, так как алгоритм просто не может поддерживать такие окна. Так же, если любой алгоритм показывал очень плохие результаты на определенном размере окна, для больших окон замеры не производилось. Время замерялось с помощью стандартной функции `clock()`. Для каждой программы было замерено минимальное, максимальное и среднее время выполнения на всех тестах.

## Результаты и рекомендации

Результаты экспериментов представлены в следующей таблице.

| Метод                     | Тип замера      | M = 2   | M = 4   | M = 6      | M = 8      |
|---------------------------|-----------------|---------|---------|------------|------------|
| Преобразование в индекс   | Мин. время, мс  | 1256    | 1279    | невозможно | невозможно |
|                           | Ср. время, мс   | 1315,89 | 1380,31 |            |            |
|                           | Макс. время, мс | 1469    | 1581    |            |            |
| map <list <list <int> > > | Мин. время, мс  | 370     | 1339    | нет данных | нет данных |
|                           | Ср. время, мс   | 417,622 | 1456,78 |            |            |
|                           | Макс. время, мс | 486     | 1586    |            |            |
| Дерево префиксов          | Мин. время, мс  | 1       | 14      | 120        | 305        |
|                           | Ср. время, мс   | 1,61    | 19,93   | 201,292    | 538,75     |
|                           | Макс. время, мс | 5       | 28      | 299        | 589        |

Анализируя результаты можно сделать следующие выводы:

- от преобразования в индекс ожидалась лучшая производительность на малых значениях. Однако эксперименты показали обратное. Видимо, множественные операции умножения на 3 сильно замедлили работу системы;
- для линейного поиска результаты вполне ожидаемы: хорошая производительность для окна длины 2 и степенной рост времени при увеличении размера окна;
- дерево поиска — ожидаемые результаты с умеренным ростом времени;
- дерево префиксов — наилучшие результаты для любых размеров окон.

Исходя из результатов проведенных экспериментов, можно сказать, что лучшим решением для вычисления значений энтропии для данной задачи является префиксное дерево, в котором сыновья хранятся в виде массива размерностью 3. Независимо от размеров окон, этот алгоритм показывает наилучшие результаты по времени работы.

Однако главным результатом исследований является не время работы программных реализаций алгоритмов. Эксперименты показали, что энтропия двумерных слов не показывает действительной картины происходящего в системе. Дело в том, что возможное количество различных окон ( $3^{l^2}$ ) слишком велико, из-за чего повторения встречаемости окон почти не происходит. В связи с этим значения энтропии примерно совпадают как для начальных, так и для конечных значений времен моделирования.

Было предложено использовать метод подсчета обобщенной энтропии двумерных слов. Метод предполагает подсчет уникального числа нулей, единиц и двоек в окне, не привязывая их к конкретному расположению. Таким образом, мы существенно уменьшаем число уникальных конфигураций в окне. Пусть  $CE(l)$  — число классов эквивалентности двумерных слов  $l \times l$  в алфавите  $\Sigma = \{0, 1, 2\}$ ,  $\mu^{(a,b,c)}$  — значение счетчика для соответствующего класса в решетке. Значения  $a, b, c$  — есть число символов 0, 1, 2 в текущем окне, при этом

$$\sum_{CE(l)} \mu^{(a,b,c)} = m$$



Тогда значение функции информационной энтропии исследуемого двумерного слова  $\nu$  для обобщенных конфигураций в окне  $l \times l$  в момент  $t_{MC}$  задается формулой

$$\tilde{H}(\nu(t_{MC}), l, \alpha) = - \sum_{CE(l)} \left( \frac{\mu^{(a,b,c)}}{m} \right) \log_{\alpha} \left( \frac{\mu^{(a,b,c)}}{m} \right)$$

Для реализации подсчета обобщенной энтропии обратим внимание на то, что информация о количестве нулей в окне при его фиксированном размере  $l$  избыточна, поскольку  $a + b + c = l^2$ . Будем хранить количество встретившихся окон в квадратном массиве  $A$  размерностью  $(l+1) \times (l+1)$ , индексы которого определяются по числу символов «1» и «2» в скользящем окне. Т.е. для фиксированной позиции скользящего окна в решетке после подсчета числа символов «1» —  $\ll Eqn0044$ . eps>> и «2» —  $c$  значение элемента массива  $A[b, c]$  увеличивается на единицу. После полного прохода по решетке верхнедиагональная часть массива  $A$  содержит необходимые нам значения счетчиков  $\mu^{(a,b,c)}$ . Результаты расчета при длине частицы 8 и размере скользящего окна от 3 до 32 приведены на рис. 4.

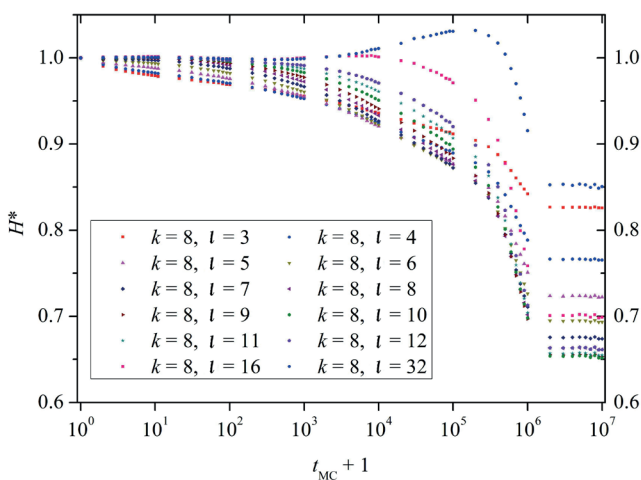


Рис. 4. Результаты расчета обобщенной энтропии [1].

Fig 4. The results of the calculation of generalized entropy [1]

## Заключение

В данной статье была рассмотрена задача вычисления энтропии двумерных слов методом скользящего окна на квадратных решетках. При этом алфавит слов состоит из трех символов. Были приведены возможные алгоритмические решения и произведены замеры времени работы программных реализаций этих алгоритмов. В результате был определен самый действующий метод, использующий префиксное дерево. Показана неэффективность метода прямого подсчета энтропии для данной задачи и предложен метод подсчета обобщенной энтропии и соответствующий алгоритм быстрой реализации этого метода.

## Список использованных источников

- [1] Лебовка Н.И., Тарасевич Ю.Ю., Гигиберия В.А., Выгорницкий Н.В., Бурмистров А.С., Лантев В.В. Образование структур в двумерных системах стержнеобразных частиц / Под ред. М.Г. Садовского, отв. за вып. М.Ю. Сенашова // Моделирование неравновесных систем – 2017. Материалы XX Всероссийского семинара, Красноярск, 06-08 октября 2017. Красноярск, ИВМ СО РАН, 2017. С. 80-83. URL: <https://elibrary.ru/item.asp?id=30073734> (дата обращения: 24.05.2018).
- [2] Михайлов Г.А., Войтишек А.В. Численное статистическое моделирование. Методы Монте-Карло. М.: Академия, 2006. 368 с.
- [3] Uljanov M., Smetanin Y. Entropy Function of Finite Words // Proceedings of 2017 IEEE International Workshop on Engineering Technologies and Computer Science (EnT). Moscow, 2017. Pp. 8-11. DOI: 10.1109/EnT.2017.7
- [4] Сметанин Ю.Г., Ульянов М.В., Пестова А.С. Энтропийный подход к построению меры символьного разнообразия слов и его применение к кластеризации геномов растений // Математическая биология и биоинформатика. 2016. Т. 11, № 1. С. 114–126. DOI: 10.17537/2016.11.114
- [5] Stroustrup B., Sutter H. C++ Core Guidelines [Электронный ресурс] // The world's leading software development platform - GitHub. URL: <https://github.com/isocpp/CppCoreGuidelines/blob/master/CppCoreGuidelines.md> (дата обращения: 24.05.2018).
- [6] Кнут Д.Э. Искусство программирования. Том 3. Сортировка и поиск / Под общ. ред. Ю. В. Козаченко; [пер. с англ. В.Т. Тертышного, И.В. Красикова]. Москва: Вильямс, 2007. Т. 3. 832 с.
- [7] Керниган Б., Ритчи Д. Язык программирования С / пер. с англ. и ред. В.Л. Бродового. Москва: Вильямс, 2012. 289 с.
- [8] Сметанин Ю.Г., Ульянов М.В. Мера символьного разнообразия: подход комбинаторики слов к определению обобщенных характеристик временных рядов // Бизнес-информатика. 2014. № 3(29). С. 40-48. URL: <https://elibrary.ru/item.asp?id=22913101> (дата обращения: 24.05.2018).
- [9] Ulyanov M.V. Cluster space of time series: generalized characteristics and symbolic dynamics // Proceedings of 2014 IEEE International Conference on Computer Technologies in Physical and Engineering Applications (ICCTPEA). St. Petersburg, 2014. Pp. 193-194. DOI: 10.1109/ICCTPEA.2014.6893355
- [10] Сметанин Ю.Г., Ульянов М.В. Энтропийные характеристики разнообразия в символьном представлении временных рядов // Современные информационные технологии и ИТ-образование. 2014. № 10. С. 426-436. URL: <https://elibrary.ru/item.asp?id=23020662> (дата обращения: 24.05.2018).
- [11] Ульянов М.В., Сметанин Ю.Г. Об одном характеристическом функционале слов над конечным алфавитом // Информационные технологии. 2017. Т. 23, № 5. С. 333-341. URL: <https://elibrary.ru/item.asp?id=29213868> (дата обращения: 24.05.2018).
- [12] Сметанин Ю.Г., Ульянов М.В., Пестова А.С. О числе возможных реконструкций слов по подсловам при окне переменного сдвига // Информационные технологии. 2018. Т. 24, № 4. С. 233–238. DOI: 10.17587/it.24.233-238



- [13] Tarasevich Y.Y., Lebovka N.I., Laptev V.V. Percolation of linear  $k$ -mers on a square lattice: From isotropic through partially ordered to completely aligned states // *Physical Review E*. 2012. Vol. 86, issue 6. P. 061116. DOI: 10.1103/PhysRevE.86.061116
- [14] Tarasevich Yu.Yu., Laptev V.V., Burmistrov A., Shinyaeva T. Influence of anisotropy on percolation and jamming of linear  $k$ -mers on square lattice with defects // *Journal of Physics: Conference Series*. 2015. Vol. 633, no. 1. P. 012064. DOI: 10.1088/1742-6596/633/1/012064
- [15] Tarasevich Yu.Yu., Burmistrov A.S., Shinyaeva T.S., Laptev V.V., Vygornitskii N.V., Lebovka N.I. Percolation and jamming of linear  $k$ -mers on a square lattice with defects: Effect of anisotropy // *Physical Review E*. 2015. Vol. 92, issue 6. P. 062142. DOI: 10.1103/PhysRevE.92.062142
- [16] Tarasevich Yu.Yu., Laptev V.V., Vygornitskii N.V., Lebovka N.I. Impact of defects on percolation in random sequential adsorption of linear  $k$ -mers on square lattices // *Physical Review E*. 2015. Vol. 91, issue 1. P. 012109. DOI: 10.1103/PhysRevE.91.012109
- [17] Lebovka N.I., Tarasevich Yu.Yu., Dubinin D.O., Laptev V.V., Vygornitskii N.V. Jamming and percolation in generalized models of random sequential adsorption of linear  $k$ -mers on a square lattice // *Physical Review E*. 2015. Vol. 92, issue 6. P. 062116. DOI: 10.1103/PhysRevE.92.062116
- [18] Tarasevich Yu.Yu., Dubinin D.O., Laptev V.V., Lebovka N.I. Impact of defects on electrical connectivity of monolayer of ideally aligned rods // *Journal of Physics: Conference Series*. 2016. Vol. 681, no. 1. P. 012038. DOI: 10.1088/1742-6596/681/1/012038
- [19] Tarasevich Yu.Yu., Goltseva V.A., Laptev V.V., Lebovka N.I. Electrical conductivity of a monolayer produced by random sequential adsorption of linear  $k$ -mers onto a square lattice // *Physical Review E*. 2016. Vol. 94, issue 4. P. 042112. DOI: 10.1103/PhysRevE.94.042112
- [20] Tarasevich Yu.Yu., Laptev V.V., Goltseva V.A., Lebovka N.I. Influence of defects on the effective electrical conductivity of a monolayer produced by random sequential adsorption of linear  $k$ -mers onto a square lattice // *Physica A: Statistical Mechanics and its Applications*. 2017. Vol. 477. Pp. 195–203. DOI: 10.1016/j.physa.2017.02.084
- [21] Budinski-Petković Lj., Lončarević I., Dujak D., Karač A., Šćepanović J. R., Jakšić Z.M., Vrhovac S.B. Particle morphology effects in random sequential adsorption // *Physical Review E*. 2017. Vol. 95, issue 2. P. 022114. DOI: 10.1103/PhysRevE.95.022114
- [22] Lebovka N.I., Vygornitskii N.V., Gigiberiya V.A., Tarasevich Yu.Yu. Monte Carlo simulation of evaporation-driven self-assembly in suspensions of colloidal rods // *Physical Review E*. 2016. Vol. 94, issue 6. P. 062803. DOI: 10.1103/PhysRevE.94.062803
- [23] Lebovka N.I., Tarasevich Yu.Yu., Gigiberiya V.A., Vygornitskii N.V. Diffusion-driven self-assembly of rodlike particles: Monte Carlo simulation on a square lattice // *Physical Review E*. 2017. Vol. 95, issue 5. P. 052130. DOI: 10.1103/PhysRevE.95.052130
- [24] Tarasevich Yu.Yu., Laptev V.V., Burmistrov A.S., Lebovka N.I. Pattern formation in a two-dimensional two-species diffusion model with anisotropic nonlinear diffusivities: a lattice approach // *Journal of Statistical Mechanics: Theory and Experiment*. 2017. Vol. 2017, no. 9. P. 093203. DOI: 10.1088/1742-5468/aa82bf
- [25] Evans J.W. Random and cooperative sequential adsorption // *Reviews of Modern Physics*. 1993. Vol. 54, issue 4. Pp. 1281–1329. DOI: 10.1103/RevModPhys.65.1281

Поступила 24.05.2018; принята в печать 10.09.2018;  
опубликована онлайн 30.09.2018.

## References

- [1] Lebovka N.I., Tarasevich Yu.Yu., Gigiberiya V.A., Vygornitskii N.V., Burmistrov A.S., Laptev V.V. Formation of structures in two-dimensional systems of rod-shaped particles. *Proceedings of the All-Russian Scientific seminar Nonequilibrium Systems Modeling – 2017*. ICM SB RAS, Krasnoyarsk, 2017. Pp. 80–83. Available at: <https://elibrary.ru/item.asp?id=30073734> (accessed 24.05.2018). (In Russian)
- [2] Mikhailov G.A., Voitishk A.V. 2006. Numerical statistical modeling (Monte Carlo methods) [Chislennoe statisticheskoe modelirovanie. Metody Monte-Karlo]. Academia, Moscow, 2006. 368 p. (In Russian)
- [3] Uljanov M., Smetanin Y. Entropy Function of Finite Words. *Proceedings of 2017 IEEE International Workshop on Engineering Technologies and Computer Science (EnT)*. Moscow, 2017. Pp. 8–11. DOI: 10.1109/EnT.2017.7
- [4] Smetanin Yu.G., Ulyanov M.V., Pestova A.S. The entropic approach to constructing a measure of the symbolic variety of words and its application to the clustering of plant genomes. *Matematicheskaya Biologiya i Bioinformatika = Mathematical Biology and Bioinformatics*. 2016; 11(1):114–126. (In Russian) DOI: 10.17537/2016.11.114
- [5] Stroustrup B., Sutter H. C++ Core Guidelines. The world's leading software development platform - GitHub. Available at: <https://github.com/isocpp/CppCoreGuidelines/blob/master/CppCoreGuidelines.md> (accessed 24.05.2018).
- [6] Knuth D.E. The Art of Computer Programming, Volume 3: (2nd Ed.) Sorting and Searching. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1998. 800 p.
- [7] Kernighan B.W., Ritchie D.M. C Programming Language, 2nd Ed., Prentice Hall, 1998. 272 p.
- [8] Smetanin Yu.G., Ulyanov M.V. Measure of symbolical diversity: Combinatorics on words as an approach to identify generalized characteristics of time series. *Business Informatics*. 2014; 3(29):40–48. Available at: <https://elibrary.ru/item.asp?id=22913101> (accessed 24.05.2018). (In Russian)
- [9] Ulyanov M.V. Cluster space of time series: generalized characteristics and symbolic dynamics. *Proceedings of 2014 IEEE International Conference on Computer Technologies in Physical and Engineering Applications (ICCTPEA)*. St. Petersburg, 2014. Pp. 193–194. DOI: 10.1109/ICCTPEA.2014.6893355
- [10] Smetanin Yu.G., Ulyanov M.V. Entropy characteristics of diversity in the symbols presentation of time series. *Modern Information Technologies and IT-Education*. 2014; 10:426–436. Available at: <https://elibrary.ru/item.asp?id=23020662> (accessed 24.05.2018). (In Russian)
- [11] Uljanov M., Smetanin Y. On a Characteristic Functional for Words over a Finite Alphabet. *Informacionnye tekhnologii = Information technologies*. 2017; 23(5):333–341. Available



- at: <https://elibrary.ru/item.asp?id=29213868> (accessed 24.05.2018). (In Russian)
- [12] Smetanin Yu.G., Ulyanov M.V. On the Number of Possible Reconstructions of Words Using Subwords with Windows of Different Shift. *Informacionnye tekhnologii = Information technologies*. 2018; 24(4):233–238. (In Russian) DOI: 10.17587/it.24.233-238
- [13] Tarasevich Y.Y., Lebovka N.I., Laptev V.V. Percolation of linear  $k$ -mers on a square lattice: From isotropic through partially ordered to completely aligned states. *Physical Review E*. 2012; 86(6):061116. DOI: 10.1103/PhysRevE.86.061116
- [14] Tarasevich Yu.Yu., Laptev V.V., Burmistrov A., Shinyaeva T. Influence of anisotropy on percolation and jamming of linear  $k$ -mers on square lattice with defects. *Journal of Physics: Conference Series*. 2015; 633(1):012064. DOI: 10.1088/1742-6596/633/1/012064
- [15] Tarasevich Yu.Yu., Burmistrov A.S., Shinyaeva T.S., Laptev V.V., Vygornitskii N.V., Lebovka N.I. Percolation and jamming of linear  $k$ -mers on a square lattice with defects: Effect of anisotropy. *Physical Review E*. 2015; 92(6):062142. DOI: 10.1103/PhysRevE.92.062142
- [16] Tarasevich Yu.Yu., Laptev V.V., Vygornitskii N.V., Lebovka N.I. Impact of defects on percolation in random sequential adsorption of linear  $k$ -mers on square lattices. *Physical Review E*. 2015; 91(1):012109. DOI: 10.1103/PhysRevE.91.012109
- [17] Lebovka N.I., Tarasevich Yu.Yu., Dubinin D.O., Laptev V.V., Vygornitskii N.V. Jamming and percolation in generalized models of random sequential adsorption of linear  $k$ -mers on a square lattice. *Physical Review E*. 2015; 92(6):062116. DOI: 10.1103/PhysRevE.92.062116
- [18] Tarasevich Yu.Yu., Dubinin D.O., Laptev V.V., Lebovka N.I. Impact of defects on electrical connectivity of monolayer of ideally aligned rods. *Journal of Physics: Conference Series*. 2016; 681(1):012038. DOI: 10.1088/1742-6596/681/1/012038
- [19] Tarasevich Yu.Yu., Goltseva V.A., Laptev V.V., Lebovka N.I. Electrical conductivity of a monolayer produced by random sequential adsorption of linear  $k$ -mers onto a square lattice. *Physical Review E*. 2016; 94(4):042112. DOI: 10.1103/PhysRevE.94.042112
- [20] Tarasevich Yu.Yu., Laptev V.V., Goltseva V.A., Lebovka N.I. Influence of defects on the effective electrical conductivity of a monolayer produced by random sequential adsorption of linear  $k$ -mers onto a square lattice. *Physica A: Statistical Mechanics and its Applications*. 2017; 477:195–203. DOI: 10.1016/j.physa.2017.02.084
- [21] Budinski-Petković Lj., Lončarević I., Dujak D., Karač A., Šćepanović J. R., Jakšić Z.M., Vrhovac S.B. Particle morphology effects in random sequential adsorption. *Physical Review E*. 2017; 95(2):022114. DOI: 10.1103/PhysRevE.95.022114
- [22] Lebovka N.I., Vygornitskii N.V., Gigiberiya V.A., Tarasevich Yu.Yu. Monte Carlo simulation of evaporation-driven self-assembly in suspensions of colloidal rods. *Physical Review E*. 2016; 94(6):062803. DOI: 10.1103/PhysRevE.94.062803
- [23] Lebovka N.I., Tarasevich Yu.Yu., Gigiberiya V.A., Vygornitskii N.V. Diffusion-driven self-assembly of rodlike particles: Monte Carlo simulation on a square lattice. *Physical Review E*. 2017; 95(5):052130. DOI: 10.1103/PhysRevE.95.052130
- [24] Tarasevich Yu.Yu., Laptev V.V., Burmistrov A.S., Lebovka N.I. Pattern formation in a two-dimensional two-species diffusion model with anisotropic nonlinear diffusivities: a lattice approach. *Journal of Statistical Mechanics: Theory and Experiment*. 2017; 2017(9):093203. DOI: 10.1088/1742-5468/aa82bf
- [25] Evans J.W. Random and cooperative sequential adsorption. *Reviews of Modern Physics*. 1993; 54(4):1281–1329. DOI: 10.1103/RevModPhys.65.1281

Поступила 24.05.2018; принята в печать 10.09.2018;  
опубликована онлайн 30.09.2018.

#### About the authors:

**Michael V. Uljanov**, Doctor of Technical Sciences, Professor, leading scientist researcher, V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences (65 Profsoyuznaya Str., Moscow 117997, Russia); Professor, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1, Leninskie gory, GSP-1, Moscow 119991, Russia), ORCID: <http://orcid.org/0000-0002-5784-9836>, [muljanov@mail.ru](mailto:muljanov@mail.ru)

**Sedyakin Ilya**, student of the Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University (1, Leninskie gory, GSP-1, Moscow 119991, Russia), ORCID: <http://orcid.org/0000-0002-9118-8517>, [ilya.sedyakin@mail.ru](mailto:ilya.sedyakin@mail.ru)



This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted reuse, distribution, and reproduction in any medium provided the original work is properly cited.

