ISSN 2076-2429 (print)
ISSN 2223-3814 (on line)

УДК 004.056.55
А.В. Соколов, магистр, Одес. нац. политехн. ун-т

БЫСТРОДЕЙСТВУЮЩИЙ ГЕНЕРАТОР КЛЮЧЕВЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ НА ОСНОВЕ КЛЕТОЧНЫХ АВТОМАТОВ

Введение. Эффективность работы большинства современных криптографических систем во многом определяется качеством входящих в их состав генераторов псевдослучайных ключевых последовательностей (ГПКП). ГПКП используются для формирования ключей, векторов инициализации криптографических алгоритмов, а также как основной элемент в современных алгоритмах поточного шифрования (АПШ).

В последнем случае ГПКП представляет собой основообразующий элемент АПШ, генерирующий поток битов z_i, используемых в качестве элементов Г-последовательности, которая далее перемешивается с потоком битов исходного сообщения x_i, в результате чего на выходе АПШ формируются элементы зашифрованного сообщения y_i. Таким образом, в АПШ исходное сообщение $\{x_i\}$, зашифрованное сообщение $\{y_i\}$ и Г-последовательность связаны соотношениями, основанными на свойствах операции сложения по модулю 2 и определяющими процедуры шифрования и дешифрования [1],

$$\begin{cases} y_i = x_i \oplus z_i \\ x_i = y_i \oplus z_i \end{cases}, \ i = 1, 2, ..., N, \qquad (1)$$

где N — длина сообщения;

$\oplus$ — знак операции сложения по модулю 2.

Анализ последних исследований и публикаций. Анализ (1) позволяет сделать вывод об определяющем влиянии качества Г-последовательности, формируемой ГПКП и используемой в АПШ, на криптостойкость и быстродействие последнего. Таким образом, актуальной задачей современной криптографии является создание быстродействующих ГПКП, которые бы формировали Г-последовательности, высококачественные с точки зрения нелинейности и стохастических свойств.

С точки зрения современной теории анализа Г-последовательностей к подобным стохастическим свойствам, помимо нелинейности алгоритма формирования Г-последовательностей, позволяющим противостоять атакам Берлекэмпа-Мэсси [2], основанным на знании нескольких периодов Г-последовательности, относят такие свойства, как сбалансированность, случайный внешний вид графического представления последовательности, равномерное распределение гистограммы, случайное распределение на плоскости, монотонность, линейная сложность, низкий уровень боковых лепестков автокорреляционной функции (АКФ), хорошие спектральные свойства, соответствие свойствам k-граммного распределения, т.е. распределения серий длины k бит на периоде последовательности [3].

Одной из лучших схем формирования Г-последовательностей в настоящее время является ГПКП, основанный на применении свойств максимальной нелинейности булевых бент-функций [2] и получивший дальнейшее развитие за счет использования свойств дуальных пар бент-функций [4].

Тем не менее, как показал анализ, несмотря на высокую эффективность, ГПКП [4] не лишен недостатков, связанных со стремительным ростом сложности реализации схемы ГПКП при

DOI: 10.15276/opu.1.43.2014.31

применении бент-функций от числа переменных $k > 4$: каждая переменная бент-функции требует подключения отдельного регистра сдвига с обратной связью (РСЛОС), программная реализация которых крайне неэффективна из-за необходимости избегать излишне разреженных генераторных полиномов для предотвращения ухудшения корреляционных свойств генерируемой последовательности, тогда как более плотные генераторные полиномы приводят к существенному замедлению работы ГПКП. Другой недостаток ГПКП, использующих РСЛОС, заключается в том, что для наращивания криптографической защищенности ГПКП приходится применять генераторные полиномы высоких степеней, поиск которых затруднен с вычислительной точки зрения.

Проведенный анализ показал, что указанные недостатки могут быть устранены за счет применения в ГПКП математического аппарата клеточных автоматов [5], который позволяет значительно упростить схемы ГПКП, адаптировать их реализацию под архитектуру современной вычислительной техники, а также добиться более высоких показателей стохастического качества.

Целью работы является построение эффективного с точки зрения стохастических и криптографических свойств ГПКП, основанного на математическом аппарате клеточных автоматов и совершенных алгебраических конструкциях.

Изложение основного материала. В простейшем случае одномерным клеточным автоматом называется дискретная структура, включающая в себя решетку ячеек памяти длины l, такую, что для каждой ячейки a_i определена окрестность из r ячеек, взаимодействующих по правилу эволюции P, определяющему переход ячеек из текущего состояния в новое с течением дискретного времени t [5].

Пример работы простейшего клеточного автомата. Пусть задано исходное состояние решетки длины $l = 8$ ячеек памяти $a = \{1,0,0,0,1,0,1,1\}$, $i = 1,2,...1,$ окрестность $r = 1$, а также правило эволюции P, для удобства представленное в виде алгебраической конструкции

$$P = \begin{array}{|c|c|c|c|c|c|c|c|} \hline 111 & 110 & 101 & 100 & 011 & 010 & 001 & 000 \\ \hline 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ \hline \end{array}. \qquad (2)$$

Тогда преобразование данных за один период дискретного времени t можно проиллюстрировать (рис. 1).

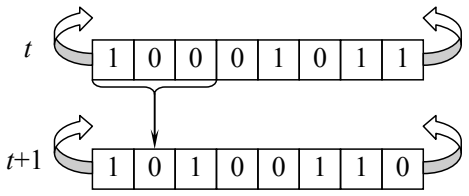


Рис. 1. Эволюция ячеек памяти a_i в клеточном автомате за один период дискретного времени t

Очевидно, что конструкция обладает простейшей аппаратной и программной реализацией, тем не менее преобразования, происходящие внутри решетки, с течением дискретного времени t могут быть нелинейными в зависимости от вида правила эволюции P.

Предлагается схема ГПКП на основе математического аппарата клеточных автоматов (рис. 2).

В схеме используется решетка ячеек памяти a_i длиной $l = 64$, что соответствует 64-битной аппаратной платформе современной вычислительной техники и позволяет генерировать за один такт работы ГПКП 64 бита (8 байт) Г-последовательности, что значительно увеличивает быстродействие генератора.

Генератор использует *Mg*-последовательности со свойством *k*-граммного распределения, т.е. такие, в которых каждая серия из k бит встречается на замкнутом цикле точно один раз [6].

Схематически изображенное правило эволюции P клеточного автомата основано на применении дуальной пары бент-функций $\{f_1, f_2\}$ от $2r = 6$ переменных, соответственно длины $N=2^n=2^6=64$. Бент-функцией называется булева функция, таблица истинности которой представляет собой бинарную, т.е. полученную в результате однозначного преобразования между двоичным и бинарным кодом $0 \leftrightarrow +$, $1 \leftrightarrow -$, последовательность $\mathbf{B}$, спектр Уолша-Адамара которой равномерно распределен по модулю,

ISSN 2076-2429 (print)
ISSN 2223-3814 (on line)

$$W_B(\omega) = \mathbf{BA} = \{\pm 2^{n/2}\},\ \omega = \overline{0, 2^n - 1},\ \omega = \overline{0, 2^n - 1}, \quad (3)$$

где **A** — матрица Уолша-Адамара порядка 2^n [4].

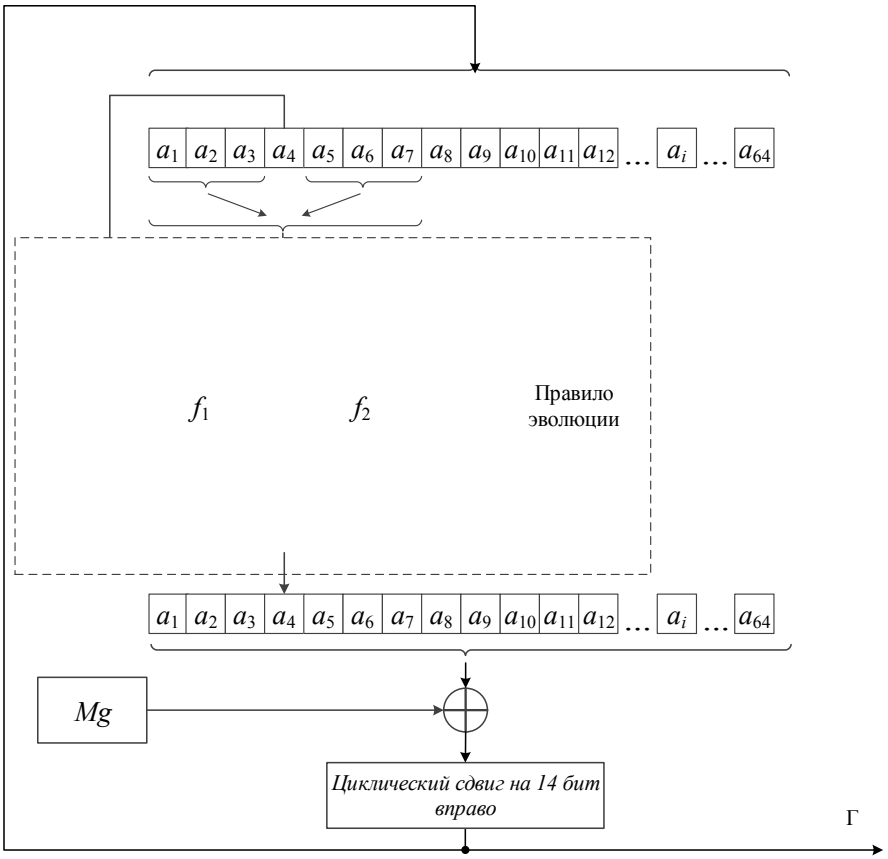


Рис. 2. Предлагаемая схема ГПКП на основе клеточных автоматов

Известны методы синтеза бент-функций от $n \le 6$ переменных, в частности, методы синтеза бент-функций от $n = 6$ переменных [7].

Использование бент-последовательностей в предлагаемой схеме ГПКП позволяет достичь наилучших показателей нелинейности генерируемой Г-последовательности, что обусловливает высокую резистивность предлагаемого ГПКП к атакам линейного криптоанализа, а также к атаке Берлекэмпа-Мэсси [2]. Использование дуальной пары бент-функций, т.е. такой, для которой количество отрицательных $K^{(-)}$ и положительных $K^{(+)}$ элементов в объединении $f_1 \cup f_2$ одинаково, т.е. $K^{(-)} = K^{(+)} = 64$, позволяет достичь равномерного распределения символов 0 и 1 в генерируемой Г-последовательности, что является необходимым условием ее стохастического качества. В качестве входов блоков бент-функций f_1 и f_2 используются ячейки памяти в окрестности $2r = 6$ по отношению к обрабатываемой ячейке памяти a_i, в то время как сама обрабатываемая ячейка памяти a_i определяет выбор бент-функции в дуальной паре. Применение столь большой окрестности r клеточного автомата существенно улучшает нелинейные свойства генерируемой Г-последовательности.

Известным недостатком применения клеточных автоматов в криптографии является наличие векторов инициализации, приводящих к плохим стохастическим свойствам генерируемой Г-последовательности, а также угроза возможной остановки эволюции клеточного автомата, его перемещение в некоторое устойчивое состояние. Для противостояния подобным явлениям в

предлагаемой схеме ГПКП используется такая совершенная алгебраическая конструкция, как последовательность со свойством *k*-граммного распределения. Свойства строгого соответствия данной последовательности *k*-граммному распределению, в частности, ее сбалансированность, позволяют путем ее комплементарного наложения на решетку ячеек памяти клеточного автомата на каждом такте поддерживать сбалансированность сегмента Г-последовательности и обеспечивать большие значения ее периода.

На выходе правила эволюции P клеточного автомата в целях улучшения автокорреляционных свойств генерируемой Г-последовательности применяется блок циклического сдвига вправо на 14 бит.

Результаты. Пусть, например, задано нулевое исходное значение клеточного автомата $a_i = 0$, $i = 1,\ 2,\ ...,\ 64$, дуальная пара бент-функций, определяемых их таблицами истинности

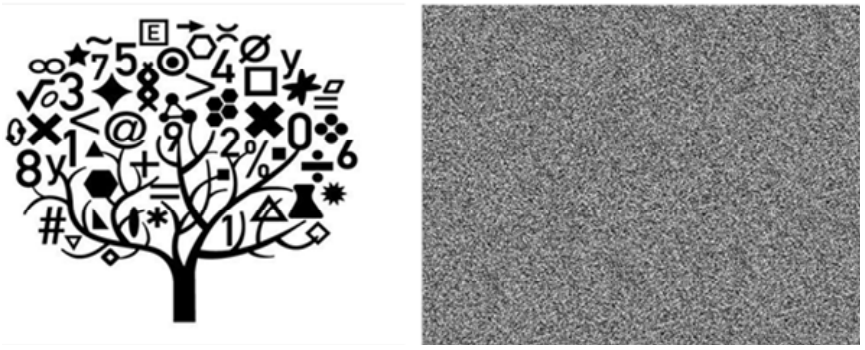
$$\begin{cases} f_1 = \{0111011100100010010001000001000101101001001111000101101000001111\}; \\ f_2 = \{0010101111101000110100011010100111111111100001110100101011001 10\}, \end{cases}$$

а также последовательность со свойством *k*-граммного распределения

$$Mg = \{1111110000001000011000101001111011101000111001001011011001101010.$$

Основываясь на заданных конкретных параметрах схемы (см. рисунок 2) возможно осуществить генерацию Г-последовательности длиной 16384 бит для проведения исследований ее стохастических свойств в соответствии с тестами [3].

Пример шифрования графического сообщения в соответствии с приведенными конкретными параметрами предлагаемого ГПКП. Исходное графическое сообщение представляет собой матрицу 500×422, каждый элемент которой является байтом (8 бит) информации, определяющим цвет (в градациях серого) соответствующего пикселя. При формировании из исходной Г-последовательности матрицы, аналогичной исходной, выполняется шифрование в соответствии с соотношениями (1), в результате чего получается зашифрованное сообщение.



а *б*

Рис. 3. Пример шифрования графического сообщения: исходное (а), зашифрованное (б)

Анализ зашифрованного сообщения позволяет сделать вывод о высоком качестве предлагаемого ГПКП, обеспечивающего полное сокрытие исходного сообщения.

Проведен сравнительный анализ характеристик предложенного ГПКП с известными аналогами (см. таблицу).

Анализ данных сравнения позволяет сделать вывод об очень высоком быстродействии предлагаемого ГПКП на основе клеточных автоматов при сохранении соответствия его стохастических характеристик современным требованиям к его стохастическому качеству. Так, быстродействие предложенного ГПКП в 2,33 раза выше быстродействия ГПКП криптоалгоритма RC4 [1], также основанного на схемах РСЛОС, тогда как быстродействие по сравнению с наилучшими ГПКП [4] на основе бент-функций выше в 11,99 раз.

ISSN 2076-2429 (print)
ISSN 2223-3814 (on line)

Сравнительный анализ характеристик предложенного ГПКП

Критерии качества	Предложенный ГПКП	ГПКП [4]	Генератор [2]	АПШ RC4 [1]
Сбалансированность	8288/8096	8250/8134	10210/ 6174	8082/8302
Случайный внешний вид сигнала	+	+	+	+
Равномерное распределение гистограммы	+	+	–	+
Случайное распределение на плоскости	+	+	–	+
k-х граммное распределение, *k*=2, 3, 4	+	+	–	+
Монотонность	+	+	+	+
Линейная сложность	+	+	+	+
Максимальный боковой лепесток битовой АКФ	~0,0303	$\sim 0,02$	$\sim 0,1$	$\sim 0,02$
Спектральный тест	+	+	–	+
Время генерации (мс)	93,8	1125,0	1078,1	218,8
Длина слова, генерируемого за один такт (бит)	64	1	1	8 (при $n = 8$)
Простота программной реализации	+	–	–	–

Очевидно, число уровней защиты предложенного генератора определяется разрядностью решетки ячеек памяти $l = 64$, т.е. его вектором инициализации, а также мощностями классов бент-функций длины $N = 64$, которых насчитывается $J_{bent} = 5\,425\,430\,528$ штук, и последовательностей со свойством *k*-граммного распределения, число которых соответственно $J_{Mg} = 67108864$. Таким образом, число уровней защиты

$$\Psi = 2^l \cdot \left(\frac{J_{bent}}{2}\right)^2 \cdot J_{Mg} = 2^{64}\left(\frac{5425430528}{2}\right)^2 67108864 \approx 2^{153}. \quad (4)$$

Соответственно, длина ключа составляет 153 бита. Следует отметить, что число уровней защиты (4) и соответственно длина ключа могут быть легко масштабированы за счет увеличения числа ячеек памяти в решетке генератора.

Выводы. Основные результаты проведенных исследований:

— получила дальнейшее развитие теория генерации Г-последовательностей, в рамках чего предложен ГПКП на основе клеточных автоматов, бент-функций и последовательностей со свойством *k*-граммного распределения, обладающий высокими стохастическими и криптографическими качествами, а также простой программной реализацией;

— проведенный сравнительный анализ предложенного и существующих ГПКП, а также расчет числа возможных уровней защиты предложенного генератора позволяют рекомендовать его к использований в современных АПШ;

— эмпирические эксперименты по шифрованию графического сообщения с помощью предложенного ГПКП позволили экспериментально подтвердить его эффективность.

Таким образом, предложенный ГПКП может быть рекомендован к использованию в современных высокопроизводительных приложениях, требующих высококачественные Г-последовательности.

Литература

1. Рябко, Б.Я. Основы современной криптографии и стеганографии / Б.Я. Рябко, А.Н. Фионов. — М.: Горячая линия — Телеком, 2010. — 232 с.

2. Агафонова, И.В. Криптографические свойства нелинейных булевых функций / И.В. Агафонова // Семинар по дискрет. гармон. анализу и геометр. моделированию. — СПб.: DHA & CAGD, 2007. — С. 1 — 24.
3. Иванов, М.А. Теория, применение и оценка качества генераторов псевдослучайных последовательностей / М.А. Иванов, И.В. Чугунков. — М.: КУДИЦ-ОБРАЗ, 2003. — 238 с.
4. Мазурков, М.И. Генератор ключевых последовательностей на основе дуальных пар бент-функций / М.И. Мазурков, Н.А. Барабанов, А.В. Соколов // Пр. Одес. політехн. ун-ту. — 2013. — Вип. 3(42). — С. 150 — 156.
5. Szaban, M Cryptographically Strong S-Boxes Based on Cellular Automata / M. Szaban, F. Seredynski // Lecture Notes in Computer Science. — 2008. — Vol. 5191. — PP. 478 — 485.
6. Мазурков, М.И. Методы синтеза двоичных псевдослучайных последовательностей со свойством k-граммного распределения / М.И. Мазурков, А.В. Соколов // Пр. Одес. політехн. ун-ту. — 2012. — Вип. 1(38). — С. 188 — 198.
7. Agievich, S.V. On the representation of bent functions by bent rectangles / S.V. Agievich // Probabilistic Methods in Discrete Mathematics: Proc. of the Fifth Intern. Petrozavodsk Conf., Petrozavodsk, June 1–6, 2000. — Utrecht, Boston: VSP, 2002. — PP. 121 — 135.

References

1. Ryabko, B.Ya. Osnovy sovremennoy kriptografii i stenografii [Foundations of modern cryptography and steganography] / B.Ya. Ryabko, A.N. Fionov. — Moscow, 2010. — 232 p.
2. Agafonova, I.V. Kriptograficheskie svoystva nelineynykh bulevykh funktsiy [Cryptographic properties of nonlinear Boolean functions] / I.V. Agafonova // Seminar po diskter. garmon. analizu i geometr. modelirovaniyu [Seminar on Discrete Harmon. Analysis and Geom. Modeling]. — St.Petersburg, 2007. — pp. 1 —24.
3. Ivanov, M.A. Teoriya, primenenie i otsenka kachestva generatorov psevdosluchaynykh posledovatel'nostey [Theory, application and quality estimation of pseudo-random sequence generators] / M.A. Ivanov, I.V. Chugunkov. — Moscow, 2003. — 238 p.
4. Mazurkov, M.I. Generator klyuchevykh posledovatel'nostey na osnove dual'nykh par bent-funktsiy [Key-sequence generator based on dual pairs of bent-functions] / M.I. Mazurkov, N.A. Barabanov, A.V. Sokolov // Pratsi Odeskoho politekhnichnoho universytetu [Proc. of Odesa Polytech. Univ.]. — 2013. — Iss. 3(42). — pp. 150 — 156.
5. Szaban, M Cryptographically Strong S-Boxes Based on Cellular Automata / M. Szaban, F. Seredynski // Lecture Notes in Computer Science. — 2008. — Vol. 5191. — pp. 478 — 485.
6. Mazurkov, M.I. Metody sinteza dvoichnykh psevdosluchaynykh posledovatel'nostey so svoystvom k-grammnogo raspredeleniya [Synthesis methods of pseudo-random binary sequences with the property of the k-gram distribution for encryption tasks] / M.I. Mazurkov, A.V. Sokolov // Pratsi Odeskoho politekhnichnoho universytetu [Proc. of Odesa Polytech. Univ.]. — 2012. — Iss. 1(38). — pp. 188 — 198.
7. Agievich, S.V. On the representation of bent functions by bent rectangles / S.V. Agievich // Probabilistic Methods in Discrete Mathematics: Proc. of the Fifth Intern. Petrozavodsk Conf., Petrozavodsk, June 1–6, 2000. — Utrecht, Boston: VSP, 2002. — pp. 121 — 135.

АНОТАЦІЯ / АННОТАЦИЯ / ABSTRACT

А.В. Соколов. **Швидкодіючий генератор ключових послідовностей на основі клітинних автоматів.** Одним з найбільш ефективних сучасних підходів до розв'язання завдання потокового шифрування є підхід, заснований на використанні для генерації псевдовипадкових ключових послідовностей математичного апарату булевих функцій. Пропонується нова конструкція генератора псевдовипадкових ключових послідовностей на основі клітинних автоматів і таких досконалих алгебраїчних конструкцій, як бент-функції і послідовності з властивістю *k*-граммного розподілу. Встановлено, що генератор генерує високоякісні з точки зору нелінійності та стохастичних властивостей ключові послідовності, при цьому в порівнянні з існуючими має просту програмну реалізацію. Таким чином, запропонований генератор псевдовипадкових ключових послідовностей може бути рекомендований для використання в сучасних криптографічних застосуваннях, наприклад, алгоритмах потокового шифрування, що вимагають високоякісні селестичні ключові послідовності.

Ключові слова: ключова послідовність, клітинний автомат, бент-функція, послідовність з властивістю *k*-грамного розподілу.

ISSN 2076-2429 (print)
ISSN 2223-3814 (on line)

А.В. Соколов. **Быстродействующий генератор ключевых последовательностей на основе клеточных автоматов.** Одним из наиболее эффективных современных подходов к решению задачи потокового шифрования является подход, основанный на использовании для генерации псевдослучайных ключевых последовательностей математического аппарата булевых функций. Предлагается новая конструкция генератора псевдослучайных ключевых последовательностей на основе клеточных автоматов и таких совершенных алгебраических конструкций, как бент-функции и последовательности со свойством *k*-граммного распределения. Установлено, что генератор генерирует высококачественные с точки зрения нелинейности и стохастических свойств ключевые последовательности, при этом по сравнению с существующими имеет простую программную реализацию. Таким образом, предложенный генератор псевдослучайных ключевых последовательностей может быть рекомендован для использования в современных криптографических приложениях, например, алгоритмах поточного шифрования, требующих высококачественные псевдослучайные ключевые последовательности.
Ключевые слова: ключевая последовательность, клеточный автомат, бент-функция, последовательность со свойством *k*-граммного распределения.

A.V. Sokolov. **Quick key sequences generator based on cellular automata.** One of the most effective modern approaches to the problem of stream encryption used for generation of pseudo-random key sequences is based on the Boolean functions mathematical instrument. A new design of the pseudo-random key sequences generator based on cellular automata and such perfect algebraic structures as bent functions and sequences with the property of *k*-gram distribution is proposed. It is shown that new generator generates a highly nonlinear and good stochastic quality pseudo-random key sequences compared with the existing high-speed generators, and have a simple program implementation. Thus, the proposed pseudo-random key sequences generator can be recommended for use in modern cryptographic applications, for example, stream encryption algorithms which require high-quality pseudo-random key sequences.
Keywords: key sequence, cellular automata, bent function, sequence with the property of *k*-gram distribution.

Рецензент д-р техн. наук, проф. Одес. нац. политехн. ун-та Мазурков М.И.

Поступила в редакцию 21 февраля 2014 г.

УДК 004.62

А.С. Коляда, магистр,
В.А Яковенко, магистр,
В.Д. Гогунский, д-р техн. наук, проф.,
Одес. нац. политехн. ун-т

ПРИМЕНЕНИЕ ЛАТЕНТНОГО РАЗМЕЩЕНИЯ ДИРИХЛЕ ДЛЯ АНАЛИЗА ПУБЛИКАЦИЙ ИЗ НАУКОМЕТРИЧЕСКИХ БАЗ ДАННЫХ

Введение. Проект по извлечению информации из наукометрических баз данных (НМБД) [1] подразумевает получение информации о публикациях из наиболее известных НМБД, которые принадлежат конкретному автору. Так как в мире может существовать несколько людей с одинаковыми ФИО, это поле не может быть уникальным идентификатором записи. Добавив к этому тот факт, что чаще всего публикации содержат только инициалы с фамилией, вероятность нахождения публикаций нескольких авторов с идентичными ФИО, еще выше. Для решения этой проблемы используется латентно-семантический анализ (ЛСА) [2], который позволяет вы-

DOI: 10.15276/opu.1.43.2014.32