

УДК 004.89

DOI: [10.26102/2310-6018/2020.31.4.021](https://doi.org/10.26102/2310-6018/2020.31.4.021)

Метод формирования критериальных оценок морфологических признаков технических систем

Д.М. Коробкин

*Волгоградский государственный технический университет,
Волгоград, Российская Федерация*

Резюме: Цель работы заключается в разработке метода формирования критериальных оценок морфологических признаков технических систем на основе анализа тенденций, выявленных в патентном массиве. Одним из распространенных способов генерации новых технических решений или усовершенствования функциональной структуры технической системы является процедура морфологического синтеза. Для оценки нового синтезированного технического решения актуальным подходом является применение критериальных оценок составляющих его морфологических признаков (технических функций элементов системы). В результате применения методов извлечения технических функций в формате DGH («Действие» - «Объект» - «Ограничение») из русскоязычных патентов и в формате SAO («Субъект» - «Действие» - «Объект») из англоязычных патентов формируется терм-документная матрица. Содержимое терм-документной матрицы модернизируется на основе разработанного алгоритма сравнения технических функций, использующего статистический анализ патентного массива с помощью технологии Word2Vec (выявление контекстных синонимов). Разработан метод формирования критериальных оценок технических функций на основе патентных трендов, выявленных посредством кластеризации патентного массива на основе терм-документной матрицы. Сформирован метод определения критериальной оценки значимости технической функции в будущем временном периоде посредством прогнозирования временных рядов на базе метода ARIMA.

Ключевые слова: морфологический синтез, DGH, SAO, патенты, анализ текста, тренд, ARIMA

Для цитирования: Коробкин Д.М. Метод формирования критериальных оценок морфологических признаков технических систем. *Моделирование, оптимизация и информационные технологии*. 2020;8(4). Доступно по: <https://moitvvt.ru/ru/journal/pdf?id=867>
DOI: 10.26102/2310-6018/2020.31.4.021

Method for formation of criterial assessments of morphological signs of technical systems

D.M. Korobkin

Volgograd State Technical University, Volgograd, Russian Federation

Abstract: The purpose of the study is to develop a method for the formation of criterion assessments of morphological features of technical systems based on the analysis of trends identified in the patent array. One of the common ways to generate new technical solutions or improve the functional structure of a technical system is the procedure of morphological synthesis. To evaluate a new synthesized technical solution, an urgent approach is the use of criterion assessments of its constituent morphological features (technical functions of system elements). As a result of the application of methods for extracting technical functions in the DGH format ("Action" - "Object" - "Restriction") from Russian-language patents and in the SAO format ("Subject" - "Action" - "Object"), a term document is formed from

English-language patents. the matrix. The content of the term-document matrix is modernized on the basis of the developed algorithm for comparing technical functions, using statistical analysis of the patent array using Word2Vec technology (identifying contextual synonyms). A method for the formation of criterion estimates of technical functions based on patent trends identified by clustering a patent array based on a term document matrix has been developed. A method for determining the criteria-based assessment of the significance of a technical function in the future time period by means of forecasting time series based on the ARIMA method has been formed.

Keywords: morphological analysis, DGH, SAO, patents, fact extraction, trend, ARIMA

For citation: Korobkin D.M. Method for formation of criterial assessments of morphological signs of technical systems. *Modeling, optimization and information technology*. 2020;8(4). Available from: <https://moitvvt.ru/ru/journal/pdf?id=867> DOI: 10.26102/2310-6018/2020.31.4.021(In Russ).

Введение

Одним из распространенных способов генерации новых технических решений [1] или усовершенствования функциональной структуры [2] технической системы является процедура морфологического синтеза. Для оценки нового синтезированного технического решения актуальным подходом является применение критериальных оценок составляющих его морфологических признаков (технических функций элементов системы [3]).

Ранее разработанные на кафедре САПР и ПК методы [4] и автоматизированные системы (АС) [5] позволили извлекать технические функции из патентных документов. Учитывая тот факт, что информация, содержащаяся в патентном массиве, не может быть оторвана от уровня техники, он сам представляет собой огромный объем информации по уже реализованным техническим решениям, то актуальным является получение критериальных оценок технических функций посредством анализа патентного массива. Данная информация должна служить для оптимального выбора альтернатив при морфологическом синтезе.

Цель работы – разработать метод формирования критериальных оценок морфологических признаков технических систем на основе анализа тенденций, выявленных в патентном массиве.

Материалы и методы

В результате применения методов извлечения технических функций в формате DGH («Действие» - «Объект» - «Ограничение») [6] из русскоязычных патентов и в формате SAO («Субъект» - «Действие» - «Объект») [7] из англоязычных патентов формируется терм-документная матрица. Из совокупности всех документов $\{P\}$ патентного массива извлекаются термы (функции технического объекта DGH_i) и определяется частота их встречаемости в патентных документах.

Метрика TF i -ой технической функции DGH_i в j -ом патентном документе P_j (раздел патента, относящийся к основному назначению изобретения):

$$TF_{ji} = f_{ij}, \quad (1)$$

где f_{ij} – частота встречаемости DGH_i в P_j ,

Таблица 1. Матрица встречаемости технических функций в патентах
Table 1. Matrix of technical functions in patents

Терм Документ \	DGH_1	DGH_2	...	DGH_M
P_1	TF_{11}	TF_{21}	...	TF_{N1}
P_2	TF_{12}	TF_{22}	...	TF_{N2}
...
P_N	TF_{1M}	TF_{2M}	...	TF_{NM}

На начальном этапе данная таблица в ячейках будет содержать 0, если DGH_i не встречается P_j , либо некоторое целочисленное значение.

Алгоритм сравнения DGH

Далее содержимое терм-документной матрицы модернизируется на основе разработанного алгоритма сравнения DGH. Если техническая функция DGH_i встречается в j -ом патентном документе P_j , то на основе DGH_i осуществляется поиск подмножества технических функций $\{DGH\}$, для которых коэффициент подобия DGH_i выше введенного порогового значения.

Рассмотрим пример сравнения двух DGH, извлеченных из патентов (Рисунок 1):

- SU245727A1 «Установка для очистки воздуха» - «Установка предназначена для очистки воздуха в производственных помещениях...»

- SU28860A1 «Фильтр для воздуха» - «Предлагаемый фильтр для воздуха, предназначенный для очищения пыльного теплого воздуха в жилых и промышленных помещениях...».



Рисунок 1 – Сравнение двух извлеченных DGH (слева – ПОЗ, справа ПОД)
Figure 1 - Comparison of two extracted DGH

а) На первом этапе идет сравнение по «Действию» (D). Если вершина первого DGH (поисковый образ запроса, ПОЗ) и второго DGH (поисковый образ документа, ПОД) не совпадают, то используется технология Word2Vec для выявления семантической близости термов «Действия».

Набор алгоритмов word2vec [8] для расчета векторных представлений слов реализует две основные архитектуры:

- Continuous Bag of Words: с учётом четырёх ближайших соседей термина (два предыдущих и два последующих слова) без учёта порядка следования.

- Skip-gram: последовательность длиной n , где элементы находятся на расстоянии не более, чем k друг от друга.

На вход подается корпус патентных тестов (формулы изобретений 200 тыс. русскоязычных патентов), а на выходе получается набор векторов слов.

В результате работы алгоритмов формируются следующие результаты:

- Word2Vec в режиме «CBoW»: на выходе список слов, характеризующих данное.
- Word2Vec в режиме «skipgrams»: на выходе список слов, взаимозаменяемых с данным (контекстные синонимы).

Для слова «Очистка» получаем согласно технологии Word2Vec следующий список контекстных синонимов с частью речи и коэффициентом Cosine similarity:

промывка - NOUN - 0.61
 очищение - NOUN - 0.59
 фильтрация - NOUN - 0.58
 очищать - VERB - 0.56
 экстракция - NOUN - 0.54
 утилизация - NOUN - 0.54
 сортировка - NOUN - 0.54
 осаждение - NOUN - 0.54
 сушка - NOUN - 0.53
 сточный - ADJ - 0.53

Среди первых 10 контекстных синонимов имеется терм «очищение».

Введем понятие коэффициента подобия DGH по элементу «Действие»:

$$K^D = \begin{cases} 1, \text{при совпадении D в ПОЗ и ПОД} \\ \text{Cosine similarity, при наличии контекстных синонимов} \\ 0, \text{при несовпадении} \end{cases} \quad (2)$$

Рассмотрим пример на Рисунке 3.1. В ПОЗ D-вершина «очистка», в ПОД – «очищение», соответственно, K^D (коэффициент подобия DGH по элементу «Действие») равен 0,59 (Cosine similarity).

б) На втором этапе идет проверка подобия «Объекта» - сравнение вершин G и связанных с ними дочерних структур.

Если G-вершины ПОЗ и ПОД не совпадают, то коэффициент подобия DGH по «Объекту» равен 0. Если G-вершины ПОЗ и ПОД совпадают, то сравниваются их дочерние элементы, связанные с данной G-вершиной.

На каждом уровне при несовпадении термов (слов) происходит проверка на значимость. Проверка значимости терма производится на основе заранее подготовленной таблицы, в которой определены коэффициенты IDF (обратная документная частота, с которой термы встречается во всех документах патентной базы). Если IDF терма меньше порогового значения, то слово значимым не является и не учитывается в расчете коэффициента подобия.

Введем понятие коэффициента подобия DGH по элементу «Объект»:

$$K^G = \begin{cases} 1, \text{при совпадении G в ПОЗ и ПОД} \\ \text{Cosine similarity, при наличии контекстных синонимов} \\ 0, \text{при несовпадении} \end{cases} \quad (3)$$

Введем понятие коэффициента подобия дочерних элементов, связанных с G-вершиной («Объект»):

$$K_C^G = \begin{cases} \frac{\sum_{i=1}^{N_{ПОЗ}} S(t_1, t_2)}{N_{ПОЗ}}, \\ 1, \text{если } N_{ПОЗ} = 0 \end{cases}, \quad (4)$$

где $N_{ПОЗ}$ – количество дочерних элементов G-вершины ПОЗ с учетом IDF;
 $TG_{ПОЗ}, TG_{ПОД}$ – набор связанных с G-вершиной дочерних элементов для ПОЗ и ПОД соответственно (с учетом IDF);
 $S(t_1, t_2)$ – оператор, определяющий совпадение дочерних элементов $t_1 \in TG_{ПОЗ}$ и $t_2 \in TG_{ПОД}$;

$$S(t_1, t_2) = \begin{cases} 1, \text{при совпадении } t_1 \text{ и } t_2 \\ \text{Cosine similarity, при совпадении } t_2 \text{ с контекстны м синонимом } t_1 \\ 0, \text{при несовпадении} \end{cases} \quad (5)$$

Рассмотрим пример на Рисунке 3.1. Коэффициент K^G подобия DGH по «Объекту» равен 1. В ПОЗ нет дочерних элементов для G-вершины, в ПОД – 2 («пыльный», «теплый»), соответственно, $N_{ПОЗ}=0$. Коэффициент K_C^G равен 0.

в) На третьем этапе идет проверка подобия «Ограничения» - сравнение вершин H и связанных с ними дочерних структур.

Если H-вершины ПОЗ и ПОД не совпадают, то коэффициент подобия DGH по «Ограничению» равен 0. Если H-вершины ПОЗ и ПОД совпадают, то сравниваются их дочерние элементы, связанные с данной H-вершиной.

Введем понятие коэффициента подобия DGH по элементу «Ограничение»:

$$K^H = \begin{cases} 1, \text{при совпадении H в ПОЗ и ПОД} \\ \text{Cosine similarity, при наличии контекстны х синонимов} \\ 0, \text{при несовпадении} \end{cases} \quad (6)$$

Введем понятие коэффициента подобия дочерних элементов, связанных с H-вершиной («Ограничение»):

$$K_C^H = \begin{cases} \frac{\sum_{i=1}^{N_{ПОЗ}} S(t_1, t_2)}{N_{ПОЗ}}, \\ 1, \text{если } N_{ПОЗ} = 0 \end{cases}, \quad (7)$$

где $N_{ПОЗ}$ – количество дочерних элементов H-вершины ПОЗ с учетом IDF;
 $TH_{ПОЗ}, TH_{ПОД}$ – набор связанных с H-вершиной дочерних элементов для ПОЗ и ПОД соответственно (с учетом IDF);
 $S(t_1, t_2)$ – оператор, определяющий совпадение дочерних элементов $t_1 \in TH_{ПОЗ}$ и $t_2 \in TH_{ПОД}$;

$$S(t_1, t_2) = \begin{cases} 1, \text{при совпадении } t_1 \text{ и } t_2 \\ \text{Cosine similarity, при совпадении } t_2 \text{ с контекстны м синонимом } t_1 \\ 0, \text{при несовпадении} \end{cases} \quad (8)$$

Рассмотрим пример на Рисунке 3.1. Коэффициент K^H подобия DGH по «Ограничению» равен 1. В ПОЗ 1 дочерний элемент для G-вершины

(«производственный»), в ПОД – 2 («жилой», «промышленный»), соответственно, $N_{ПОЗ}=1$. Терм «производственный» проходит фильтр по параметру IDF.

Для слова «производственный» получаем согласно технологии Word2Vec следующий список контекстных синонимов с частью речи и коэффициентом Cosine similarity:

производство - NOUN - 0.68
 предприятие - NOUN - 0.65
 технологический - ADJ - 0.65
 технический - ADJ - 0.63
 проектный - ADJ - 0.60
 промышленный - ADJ - 0.60
 лесозаготовительный - ADJ - 0.59
 оборудование - NOUN - 0.58
 эксплуатационный - ADJ - 0.57

Среди первых 10 контекстных синонимов имеется терм «промышленный».

Коэффициент K_C^H равен 0,6.

г) Введем понятие коэффициента подобия 2 структур ВПР:

$$K_{DGH} = \frac{K^D + \frac{K^G + K_C^G}{2} + \frac{K^H + K_C^H}{2}}{3} \quad (9)$$

Для примера на Рисунке 3.1. $K_{DGH} = (0,59+2/2+1,6/2)/3 = 0,8$. Два DGH можно назвать подобными, если их коэффициент подобия больше некоего порогового значения. Максимальное значение коэффициента равно 1.

Далее для каждой DGH_i осуществляется поиск подобных структур.

Если на основе DGH_i выявили схожую DGH_j с коэффициентом подобия K_{DGH} , то характеристика TF для DGH_j в строках терм-документной матрицы, соответствующих патентам, в которых $TF(DGH_i)$ больше 0, примет значение $TF(DGH_i) * K_{DGH}$.

Например, если на основе DGH_1 выявили подобную структуру DGH_2 с коэффициентом $K_{DGH}=0,8$, а для DGH_2 коэффициент подобия DGH_1 равен 0,58. В таблицах 2 и 3 приведены примеры пересчета характеристики TF.

Таблица 2. Пример исходной матрицы встречаемости технических функций в патентах
Table 2. Example of the initial matrix of the occurrence of technical functions in patents

Терм Документ \	DGH_1	DGH_2	...	DGH_M
P_1	1	0	...	0
P_2	0	2	...	1
...
P_N	0	0	...	1

Таблица 3. Пример модернизированной матрицы встречаемости технических функций в патентах

Table 3. Example of an upgraded matrix of technical functions occurrence in patents

Терм Документ \	DGH_1	DGH_2	...	DGH_M
P_1	1	1*0,8	...	0
P_2	2*0,58	2	...	1
...
P_N	0	0	...	1

Глобальная оценка DGH

В качестве глобального взвешивания DGH_i выбрана инвертированная документная частота, которая равна 0, если техническая функция появляется во всех патентных документах массива, и этот вес увеличивается по мере того, как количество документов, в которых встречается техническая функция, уменьшается:

$$IDF_i = \log \left(\frac{N}{\sum_{j=1}^N \chi(f_{ij})} \right), \quad (10)$$

где N – количество документов в патентном массиве,
 f_{ij} – частота встречаемости DGH_i в патенте P_j ,

$$\chi(f_{ij}) = \begin{cases} 1, & \text{при } f_{ij} > 0, \\ 0, & \text{при } f_{ij} = 0 \end{cases} \quad (11)$$

Таким образом, инвертированная документная частота (IDF) будет одной из критериальных оценок морфологических признаков технических систем, которая позволит выявить «общеупотребимые» технические функции и снизить их значимость.

Критериальная оценка присутствия DGH в ключевых современных технологиях

Патентный массив анализируется по временным промежуткам (кварталам), выявляются кластеры с максимальной нарастающей активностью (патентный тренд), формируется набор ключевых технических функций, соответствующих современным технологиям.

Осуществляется нечеткая кластеризация матрицы «Патент-Техническая функция» для выявления ключевых современных технологий и изобретений. При этом учитываются патенты, зарегистрированные в течение последних 10 лет (2009-2019 гг.). Строятся матрицы «Патент-Техническая функция» для каждого временного квартала (I квартал 2009 г., II квартал 2009 г., ... III квартал 2019 г.). Кластеризация происходит на основе модели LDA и построенная модель используется для получения векторов распределения по кластерам (неименованным топикам). Для этого применим фреймворк распределенных вычислений Spark и его библиотеку машинного обучения MLlib.

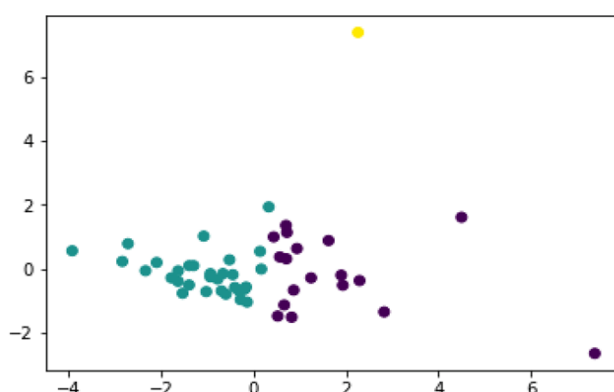


Рисунок 2 - Распределение векторов по неименованным топикам
 Figure 2 - Distribution of vectors by unnamed topics

Терм-документная матрица представляет собой математическую матрицу, в которой строки являются патентами, а столбцы - термами. Данная матрица описывает статистический набор документа, описывает частоту появления каждого термина в

документе. В нашем случае под терминами понимаются технические функции в формате SAO, под документами – патенты.

Для построения модели Латентного размещения Дирихле (LDA) [9] используется матрица на основе TF (term frequency), которая отображает частоту встречаемости DGH в текстах патентов.

Используется словарь всех технических функций в формате DGH, извлеченных из патентного массива, $voc = \{t_1, t_2, t_3, \dots, t_n\}$, где t_i – очередной уникальный трендовый SAO патентного массива. Тогда каждая i -ая строка терм-документной матрицы будет представлять собой полученный словарь, а каждый столбец будет представлять собой количество вхождений структуры DGH (включая подобные структуры, найденные согласно разработанному алгоритму) в i -ый патент.

На основе модели LDA можно получить одновременную кластеризацию патентов и технических функций на одном и том же множестве кластеров, называемых темами. В результате строится «мягкая» кластеризация, то есть патент (и техническая функция) может принадлежать нескольким темам в различной степени. Будем в дальнейшем использовать в качестве порогового значения степень принадлежности нечеткому кластеру равную 0,7.

$$DGH_Cl_i = \{DGH_1\}, P_Cl_i = \{P_k\} \quad (12)$$

где $i = 1, \dots, cl$ - индекс кластера;

cl – заданное количество кластеров (на основе проведенных экспериментов устанавливаем количество кластеров равное 150);

DGH_Cl_i - i кластер технических функций в формате DGH;

$DGH_1 \in T_{DGH}$ — 1-ая техническая функция в кластере DGH_Cl_i ;

P_Cl_i - i кластер патентов;

$P_k \in T_P$ — k -ый патент в кластере P_Cl_i .

Далее ранжируем полученные кластеры по количеству патентов и соотносим принадлежащие им технические функции. Выбираем первые 30 кластеров с максимальным количеством патентов и минимальным расстоянием от центроида до самого удаленного патента (Евклидово расстояние).

Таким образом, для каждого временного квартала получаем набор так называемых «трендовых» технических функций в формате DGH, составляющих первые 30 кластеров из ранжированного списка.

Множество трендовых DGH за определенный промежуток времени определяется следующим образом:

$$T_{DGH}^t = \{DGH_1, DGH_2, \dots, DGH_{N_{DGH}^t}\} \quad (13)$$

Критериальная оценка присутствия технической функции в ключевых современных технологиях формируется как ее частота встречаемости во множествах трендовых DGH для всех 39 временных кварталов (I квартал 2009 г., II квартал 2009 г., ... III квартал 2019 г.):

$$TR_i = \sum_{j=1}^{29} \chi(f_{ij}), \quad (14)$$

где f_{ij} – частота встречаемости DGH_i в списке трендовых технических функций T_{DGH}^j ,

$$\chi(f_{ij}) = \begin{cases} 1, & \text{при } f_{ij} > 0, \\ 0, & \text{при } f_{ij} = 0, \end{cases} \quad (15)$$

Прогнозная оценка DGH

Следующая критериальная оценка технической функции формируется на основе прогноза ее встречаемости в будущем временном периоде.

Строятся временные ряды на основе вектора встречаемости технической функции в патентных документах (используется характеристика TF из терм-документной матрицы «Патент-Техническая функция»). При этом учитываются патенты, зарегистрированные в течение последних 30 лет (1989-2018 гг.). Строятся матрицы «Патент-Техническая функция» для каждого временного квартала (I квартал 1989 г., II квартал 1989 г., ... IV квартал 2018 г.).

Таким образом, для каждой технической функции в формате DGH получаем временной ряд, где фиксируется количество упоминаний TF в патентах, выданных в определенный временной квартал. Это будет набор данных №1, который используется для реализации предложенного метода, а набор данных №2 (I квартал 2019 г., II квартал 2019 г., III квартал 2019 г.) будет использоваться для проверки эффективности предложенного метода.

Формула для модели прогнозирования развития технологии, выраженной в терминах DGH, может быть получена с помощью регрессионного анализа.

Используется 3 показателя (2 количественных и 1 качественный):

- T1 оценивает популярность (частоту упоминаний) DGH в течение соответствующего периода времени, основанную на метрике TF
- T2 показывает потенциал развития упоминаний DGH в будущем. Используем наклон кривой в последнем ретроспективном временном периоде, чтобы показать тенденцию роста кривой в будущем
- Качественный критерий T3 формируется экспертами, используется для минимизации ожидаемого отклонения, вызванного двумя предыдущими количественными показателями. Данный критерий используется при выборе технического решения, построенного на морфологической таблице (пункт 5 Отчета). Экспертам предлагается оценить важность каждой реализуемой технической функции (TF) путем распределения баллов (по 10-балльной шкале) для каждого компонента технического объекта (ТО), и тогда TF с большим вкладом в решение конкретных проблем, которые предполагается решить с помощью нового технического решения, могут получить более высокие баллы, чем другие. Среднее значение всех оценок экспертов определяется как T3.

Критериальная оценка значимости технической функции в будущем временном периоде формируется посредством прогнозирования временных рядов на базе метода ARIMA [10].

Обсуждение

Теоретическая значимость работы заключается в разработанных методах сравнения технических функций для формирования терм-документной матрицы на основе извлеченных из патентов технических функций в формате «Действие-Объект-Ограничение» / «Субъект-Действие-Объект», кластеризации патентного массива на основе терм-документной матрицы, формирования критериальных оценок технических функций на основе патентных трендов, вычисления прогнозных оценок.

Заключение

В данной статье был рассмотрен метод формирования критериальных оценок морфологических признаков технических систем на основе анализа тенденций, выявленных в патентном массиве.

Благодарности

Работа выполнена при финансовой поддержке РФФИ (грант № 18-07-01086 а), РФФИ и Администрации Волгоградской области (гранты №№ 19-47-340007 р_а, 19-41-340016 р_а).

ЛИТЕРАТУРА

1. Коробкин Д.М., Фоменков С.А., Колесников С.А. Метод синтеза функциональной структуры новых технических решений на основе данных патентных массивов. *Моделирование, оптимизация и информационные технологии*. 2019;7(2):135-148
2. Коробкин Д.М., Фоменков С.А., Колесников С.Г. Метод верификации синтезированной функциональной структуры посредством построения физического принципа действия технической системы. *Моделирование, оптимизация и информационные технологии*. 2019;7(2):97-109.
3. Шабанов Д.В., Коробкин Д.М., Фоменков С.А., Колесников С.Г. Формирование матрицы "физические эффекты - технические функции" на основе данных анализа патентных массивов. *Математические методы в технике и технологиях - ММТТ*. 2019;7:94-99
4. Korobkin D., Fomenkov S., Kolesnikov S., Lobeyko V., Golovanchikov A. Modification of physical effect model for the synthesis of the physical operation principles of technical system. *Communications in Computer and Information Science* (см. в книгах). 2015; 535:367-377
5. Korobkin D.M., Fomenkov S.A., Kolesnikov S.G., Golovanchikov A.B. Technical function discovery in patent databases for generating innovative solutions. В сборнике: *Proceedings of the International Conferences on ICT, Society, and Human Beings 2016, Web Based Communities and Social Media 2016, Big Data Analytics, Data Mining and Computational Intelligence 2016 and Theory and Practice in Modern Computing 2016 - Part of the Multi Conference on Computer Science and Information Systems 2016*. 2016:241-245.
6. Korobkin D.M., Vasiliev S.S., Fomenkov S.A., Lobeyko V.I. Extraction of structural elements of inventions from russian-language patents. В сборнике: *Multi Conference on Computer Science and Information Systems, MCCSIS 2019 - Proceedings of the International Conferences on Big Data Analytics, Data Mining and Computational Intelligence 2019 and Theory and Practice in Modern Computing 2019*. 4. 2019:159-166
7. Kharitonov A., Korobkin D., Fomenkov S., Kolesnikov S. Extraction of morphological features of technical systems from russian patent. В сборнике: *CEUR Workshop Proceedings. IS 2019 - Proceedings of the 14th International Conference on Interactive Systems: Problems of Human-Computer Interaction*. 2019:205-213.
8. Serge, Sonfack. Word2Vec how it works. 2019. DOI: 10.13140/RG.2.2.12524.51844.
9. Zhang, Meng & Li, Wei & Zhang, Bo. A Novel Method to Solve the Separation Problem of LDA. 2020. DOI: 10.1007/978-3-030-46931-3_6.
10. Sanjoyo, Sanjoyo. ARMA ARIMA ARCA GARCH. 2020. DOI: 10.13140/RG.2.2.29139.81445.

REFERENCES

1. Korobkin D.M., Fomenkov S.A., Kolesnikov S.A. Metod sinteza funktsional'noy strukturyi novyih tehnikeskikh resheniy na osnove dannyih patentnyih massivov. *Modelirovanie, optimizatsiya i informatsionnyie tehnologii*. 2019;7(2):135-148
2. Korobkin D.M., Fomenkov S.A., Kolesnikov S.G. Metod verifikatsii sintezirovannoy funktsional'noy strukturyi posredstvom postroeniya fizicheskogo printsipa deystviya tehnikeskoy sistemyi. *Modelirovanie, optimizatsiya i informatsionnyie tehnologii*. 2019;7(2):97-109.
3. Shabanov D.V., Korobkin D.M., Fomenkov S.A., Kolesnikov S.G. Formirovanie matritsyi "fizicheskie efektyi - tehnikeskie funktsii" na osnove dannyih analiza patentnyih massivov. *Matematicheskie metodyi v tehnike i tehnologiyah - MMTT*. 2019;7:94-99.
4. Korobkin D., Fomenkov S., Kolesnikov S., Lobeyko V., Golovanchikov A. Modification of physical effect model for the synthesis of the physical operation principles of technical system. *Communications in Computer and Information Science (sm. v knigah)*. 2015;535:367-377.
5. Korobkin D.M., Fomenkov S.A., Kolesnikov S.G., Golovanchikov A.B. Technical function discovery in patent databases for generating innovative solutions. V sbornike: *Proceedings of the International Conferences on ICT, Society, and Human Beings 2016, Web Based Communities and Social Media 2016, Big Data Analytics, Data Mining and Computational Intelligence 2016 and Theory and Practice in Modern Computing 2016 - Part of the Multi Conference on Computer Science and Information Systems 2016*. 2016:241-245
6. Korobkin D.M., Vasiliev S.S., Fomenkov S.A., Lobeyko V.I. Extraction of structural elements of inventions from russian-language patents. V sbornike: *Multi Conference on Computer Science and Information Systems, MCCSIS 2019 - Proceedings of the International Conferences on Big Data Analytics, Data Mining and Computational Intelligence 2019 and Theory and Practice in Modern Computing 2019*. 4. 2019:159-166.
7. Kharitonov A., Korobkin D., Fomenkov S., Kolesnikov S. Extraction of morphological features of technical systems from russian patent. V sbornike: *CEUR Workshop Proceedings. IS 2019 - Proceedings of the 14th International Conference on Interactive Systems: Problems of Human-Computer Interaction*. 2019:205-213
8. Serge, Sonfack. Word2Vec how it works. 2019. DOI: 10.13140/RG.2.2.12524.51844.
9. Zhang, Meng & Li, Wei & Zhang, Bo. A Novel Method to Solve the Separation Problem of LDA. 2020. DOI: 10.1007/978-3-030-46931-3_6.
10. Sanjoyo, Sanjoyo. ARMA ARIMA ARCA GARCH. 2020. DOI: 10.13140/RG.2.2.29139.81445.

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Дмитрий Михайлович Коробкин, канд. техн. наук, доцент кафедры САПРиПК, Волгоградский государственный технический университет, Волгоград, Russian Federation
Dmitriy M. Korobkin, PhD, Associate Professor of the CAD Department Volgograd State Technical University, Volgograd, Russia
e-mail: dkorobkin80@mail.ru