

ИНФОРМАТИКА, ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И УПРАВЛЕНИЕ

УДК 004.891

ЭКСПЕРИМЕНТАЛЬНАЯ ПРОВЕРКА МОДЕЛИ ОЦЕНКИ ИННОВАЦИОННОСТИ ОБЪЕКТА

В.К. ИВАНОВ, канд. техн. наук

Тверской государственный технический университет, 170026, Тверь,
наб. Аф. Никитина, д. 22, e-mail: mtivk@mail.ru

© Иванов В.К., 2020

В статье рассматривается подход к количественной оценке инновационности продуктов и технологий. Результаты такой оценки могут быть использованы при создании хранилища данных для описаний объектов со значительным инновационным потенциалом. Модель расчета индекса инноваций основана на понятиях новизны, актуальности и имплементируемости объекта. Даны формальные определения этих показателей, описана методика их расчета. Используются нечеткие методы для обработки неполной информации из многочисленных источников и для получения вероятностных оценок инноваций. Представлены экспериментальные данные проверки модели, в том числе расчеты локальных критериев и глобального аддитивного оценочного критерия. Установлены цикличность динамических изменений показателей, их взаимозависимость, некоторые общие особенности продвижения продуктов. Полученные экспериментальные данные согласуются с экспертными оценками исследуемых продуктов. Анализ локальных критериев дает основание утверждать о правильном использовании аддитивной n -мерной функции полезности. Подтверждается адекватность предположений и формальных выражений, которые применяются в вычислительных алгоритмах отбора информации для размещения в хранилище данных.

Ключевые слова: хранилище данных, инновационность, аддитивный критерий, функция полезности, поисковый запрос.

DOI: 10.46573/2658-5030-2020-4-54-63

ВВЕДЕНИЕ

В проекте «Организация и поддержка хранилища данных на основе интеллектуализации поискового агента и эволюционной модели отбора целевой информации» [1] предлагается технология создания хранилища описаний объектов (продуктов и технологий), обладающих значимым инновационным потенциалом. Инновационный потенциал количественно оценивается индексом инновационности, модель вычисления которого основана на показателях новизны, востребованности и имплементируемости объекта. Указанные показатели вычисляются для лингвистической модели объекта, создаваемой с помощью архетипов.

Архетипы объекта (АО) – это концепции предметной области для рассматриваемого объекта. Они реализуются термами, определяющими ключевые свойства объекта, и группируются в три дескриптивных класса: структура объекта, условия применения и результаты функционирования. Область определения АО

определяется термом-маркером. Могут также существовать дополнительные локальные ограничения: умолчания, синонимы термов, веса термов, предельное количество запросов, количество термов в запросе и т.п. Классы АО, маркер и локальные ограничения образуют лингвистическую модель объекта, которая используется как поисковый паттерн для генерации набора запросов на поиск информации о потенциально инновационном объекте.

Поисковые запросы – логические выражения, где множества операндов есть различные комбинации термов-архетипов и маркера. Релевантные документы, найденные после исполнения всех сгенерированных запросов, используются для вычисления показателей инновационности объекта.

Цель статьи – подтвердить основные предположения, используемые при разработке модели для вычисления аддитивного оценочного критерия инновационности объектов, представив данные экспериментальной проверки этой модели.

РАБОТЫ ПО ТЕМАТИКЕ ИССЛЕДОВАНИЯ

Обсуждаемые в настоящей статье глобальный и частные критерии применяются в разрабатываемой технологии создания хранилища описаний инновационных объектов. Анализ источников [2, 3], посвященных различным аспектам инновационного развития, показывает, что в понятие «инновация» всегда включаются такие коннотаты, как новый, научный, повышающий эффективность, приносящий прибыль. Исходя из этого формализуем понятие «инновационность».

Теоретическую основу формирования многокритериальных скалярных оценок, включая глобальный аддитивный критерий, составляет аксиоматическая теория полезности [4]. Одним из ее важных положений является доказанное утверждение, что из аддитивности оценочной функции следует взаимная независимость факторов, влияющих на нее [5]. Анализ используемых в исследовании частных критериев дает основание говорить о корректном использовании аддитивной n -мерной функции полезности.

В связи с ожидаемой неполнотой и неточностью информации об инновационном потенциале объектов, полученной из различных источников, использованы подход и методы теории свидетельств [6, 7].

Настоящая статья является продолжением предыдущих работ автора и его коллег о различных аспектах технологии создания хранилища описаний инновационных объектов [8, 9 и др.].

МОДЕЛЬ ВЫЧИСЛЕНИЯ ИНДЕКСА ИННОВАЦИОННОСТИ ОБЪЕКТА

Основные понятия. Ниже сформулированы понятия новизны, востребованности и имплементируемости как составных частей критерия инновационности искомого объекта. Количественная оценка этих показателей основывается на предположении об адекватности отображения жизненного цикла продуктов в виде информационных объектов, размещенных в различных хранилищах данных.

Новизна определяет значительные улучшения, новый способ использования или предоставления объекта. Предполагается, что для новых объектов количество найденной информации будет меньше, чем для давно существующих.

Востребованность (или спрос) – это осознанная потенциальным потребителем необходимость в этом объекте. Оценка востребованности основывается на значении частоты пользовательских запросов информации об объекте, находящейся в хранилищах данных. Примеры: частота выполнения запросов, похожих на запросы из поискового паттерна; показатель цитирования материалов об объекте; количество продаж объекта.

Имплементируемость определяет технологическую обоснованность, физическую осуществимость и способность объекта быть интегрированным в какую-либо систему для получения желаемых эффектов. Оценка имплементируемости основывается на значении среднего периода восстановления уровня новизны и/или востребованности объекта после их потери. Чем быстрее это происходит за счет новых технологий, конструкций, улучшенных функциональных и потребительских характеристик, тем выше имплементируемость.

Базовые вычисления. Индикатор технологической новизны Nov вычисляется следующим образом:

$$Nov = 1 - 1/N \sum_{k=1}^N f_n^{01}(R_k, \dots), \quad (1)$$

где N – общее количество выполненных запросов; R_k – число документов, найденных в результате выполнения k -го запроса; $f_n^{01}(R_k, \dots)$ – вариативная функция, нормирующая значение R_k на диапазон $[0; 1]$.

Индикатор востребованности Dem рассчитывается по формуле

$$Dem = 1/S \sum_{k=1}^S f_n^{01}(F_k, \dots), \quad (2)$$

где S – общее количество выполненных запросов; F_k – частота выполнения k -го запроса; $F_{01}(F_k, \dots)$ – вариативная функция, нормирующая значение F_k на диапазон $[0; 1]$.

Индикатор имплементируемости Imp находится как

$$Imp = 1 - 1/2(\Delta t_N(Nov(t)) + \Delta t_D(Dem(t))), \quad (3)$$

где $Nov(t)$ – функция, показывающая зависимость Nov от времени на временном интервале $[t_0; t_m]$; $Dem(t)$ – функция, показывающая зависимость Dem от времени на том же $[t_0; t_m]$; Δt_N и Δt_D – средние расстояния между двумя последовательными точками временного ряда $t_i, t_{i+1} \in [t_0; t_m]$ локальных максимумов функций $Nov(t)$ и $Dem(t)$. При этом Nov , Dem и Imp рассчитываются для точки t_{m+1} .

Индекс инновационности Ix имеет вид аддитивного критерия:

$$Ix = w_{Nov}Nov + w_{Dem}Dem + w_{Imp}Imp, \quad (4)$$

где w_{Nov} , w_{Dem} , w_{Imp} – весовые коэффициенты для Nov , Dem и Imp соответственно и $w_{Nov} + w_{Dem} + w_{Imp} = 1$.

Вычисления при неполной и неточной информации об объектах. В этом случае в проекте вводятся нечеткие показатели вероятности того, что объект обладает новизной, востребованностью и имплементируемостью. Для вычисления указанных вероятностей применяется теория свидетельств Демпстера – Шафера [6, 7]. Определяются базовые вероятности m попадания результатов измерения Nov , Dem и Imp в k -й интервал значений A_k ; результаты из различных источников комбинируются

с помощью разработанного специализированного алгоритма групповой обработки результатов измерений. Комбинирование выполняется рекурсивно по парам источников: из двух источников свидетельств образуется один условный источник, свидетельства которого комбинируются с очередным фактическим источником.

Рассчитываются функция доверия $Bel(A) = \sum_{A_k: A_k \subseteq A} m(A_k)$ и функция правдоподобия $Pl(A) = \sum_{A_k: A_k \cap A} m(A_k)$, которые определяют верхнюю и нижнюю границу вероятности того, что объект обладает свойством, заданным соответствующим фактором. Тогда выражение (4) приобретает вид мультипликативного оценочного критерия:

$$I_x = [Bel_{Nov}(A), Pl_{Nov}(A)]^{w_{Nov}} \times [Bel_{Dem}(A), Pl_{Dem}(A)]^{w_{Dem}} \times [Bel_{Imp}(A), Pl_{Imp}(A)]^{w_{Imp}}, \quad (5)$$

которое сводится логарифмированием I_x к аддитивному критерию:

$$\ln I_x = w_{Nov} \times \ln([Bel_{Nov}(A), Pl_{Nov}(A)]) + w_{Dem} \times \ln([Bel_{Dem}(A), Pl_{Dem}(A)]) + w_{Imp} \times \ln([Bel_{Imp}(A), Pl_{Imp}(A)]). \quad (6)$$

Поскольку $\ln()$ – возрастающая функция, рассуждения, касающиеся I_x , справедливы для $\ln(I_x)$.

МЕТОДИКА И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

Методика проведения экспериментов. Цель экспериментов заключалась в проверке модели вычислений локальных и глобального оценочных критериев и тем самым подтверждении адекватности формальных выражений (1)–(3), которые используются в вычислительных алгоритмах отбора информации для размещения в хранилище данных.

В качестве объектов с очевидным инновационным потенциалом были выбраны смартфоны популярных моделей известных производителей. После этого экспертами были сформированы две лингвистических модели: *iPhone X* и *Samsung Galaxy S*. Эти модели были использованы для генерации поисковых запросов. Пример поисковых запросов (маркер – «смартфон iPhone»; АО: «камера», «экран», «музыка», «производительность», «аккумулятор»; ограничение – три термина в запросе):

смартфон AND iPhone AND камера;
смартфон AND iPhone AND экран;
смартфон AND iPhone AND производительность;
смартфон AND iPhone AND музыка;
смартфон AND iPhone AND аккумулятор.

В качестве источников данных об объектах использовались следующие хранилища данных: Google Scholar (<https://scholar.google.ru>), AliExpress (<https://aliexpress.com>), ACM Digital Library (<https://dlnext.acm.org>), IEEE Explore Digital Library (<https://ieeexplore.ieee.org>).

Измеряемые показатели: *Nov*, *Dem* (усреднялись среднеарифметическим и медианным значениями, нормализовались на диапазон [0; 1] функциями для линейной и экспоненциальной нормализации). Вычисляемые показатели: *Imp*, *I_x*.

Результаты экспериментов. На рис. 1 представлен график изменения индикатора новизны исследуемых объектов за 10-летний период. На этом и последующих графиках аппроксимирующие кривые показаны пунктиром.

Сравнение графиков изменения индикатора новизны объектов в зависимости от вида нормирующей функции показано на рис. 2.

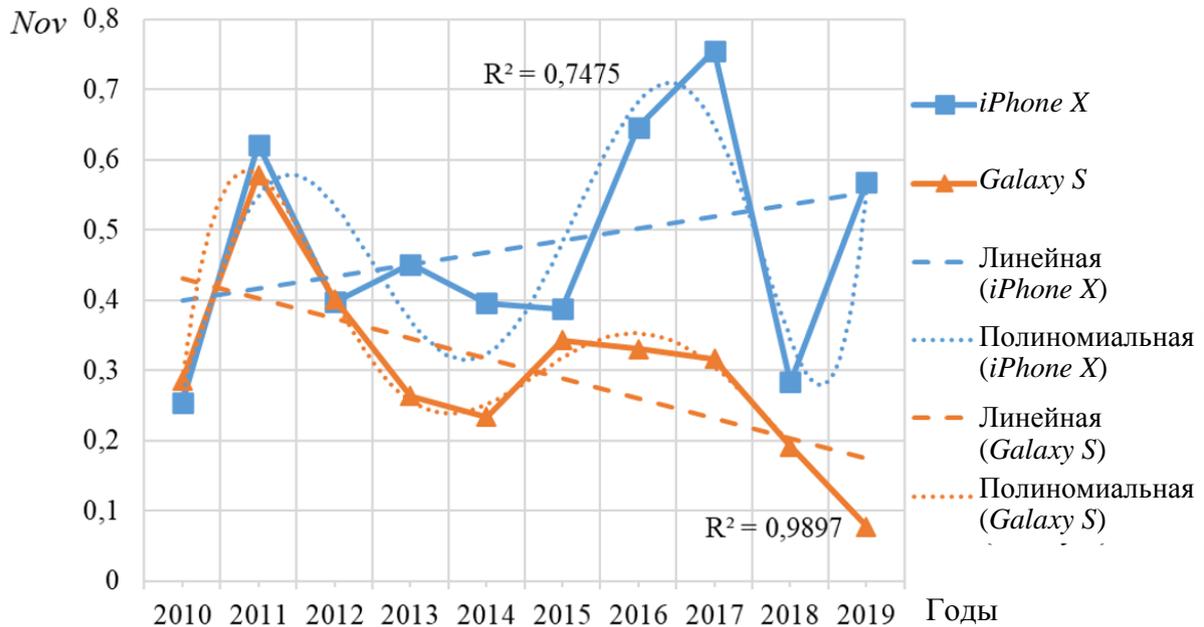


Рис. 1. Динамика изменения новизны объектов *iPhone X* и *Galaxy S* за 10-летний период

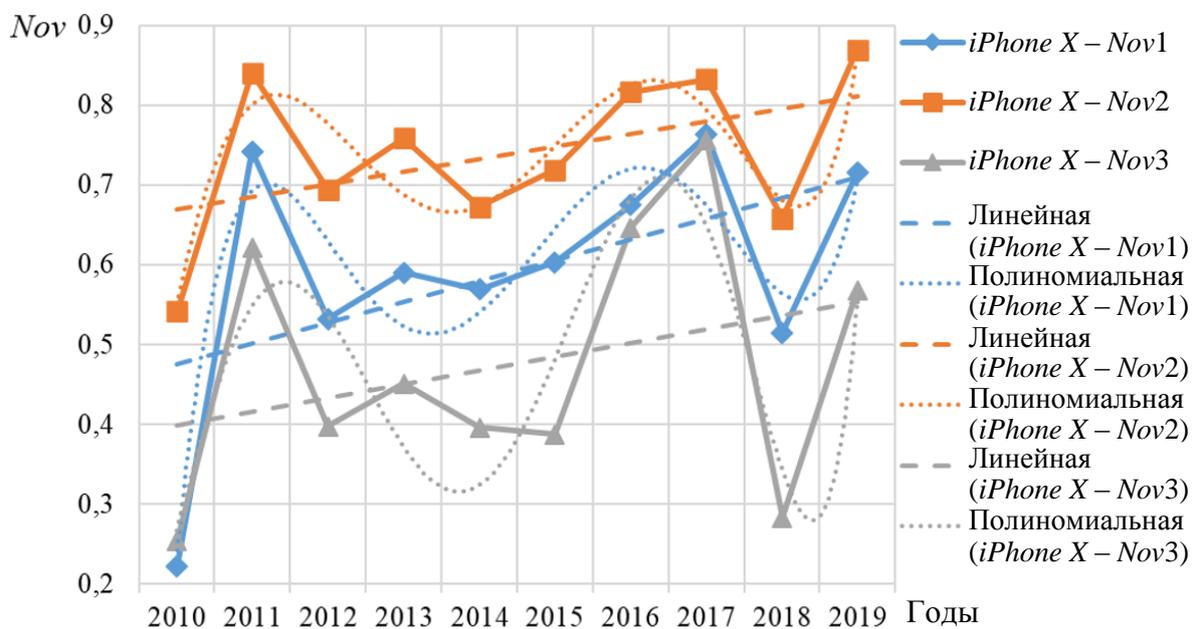


Рис. 2. Динамика изменения новизны объекта *iPhone X* для различных нормирующих функций

Рис. 3 и 4 иллюстрируют изменения индикатора востребованности исследуемых объектов за 10-летний период (для случаев, когда F_k – медиана и среднее арифметическое).

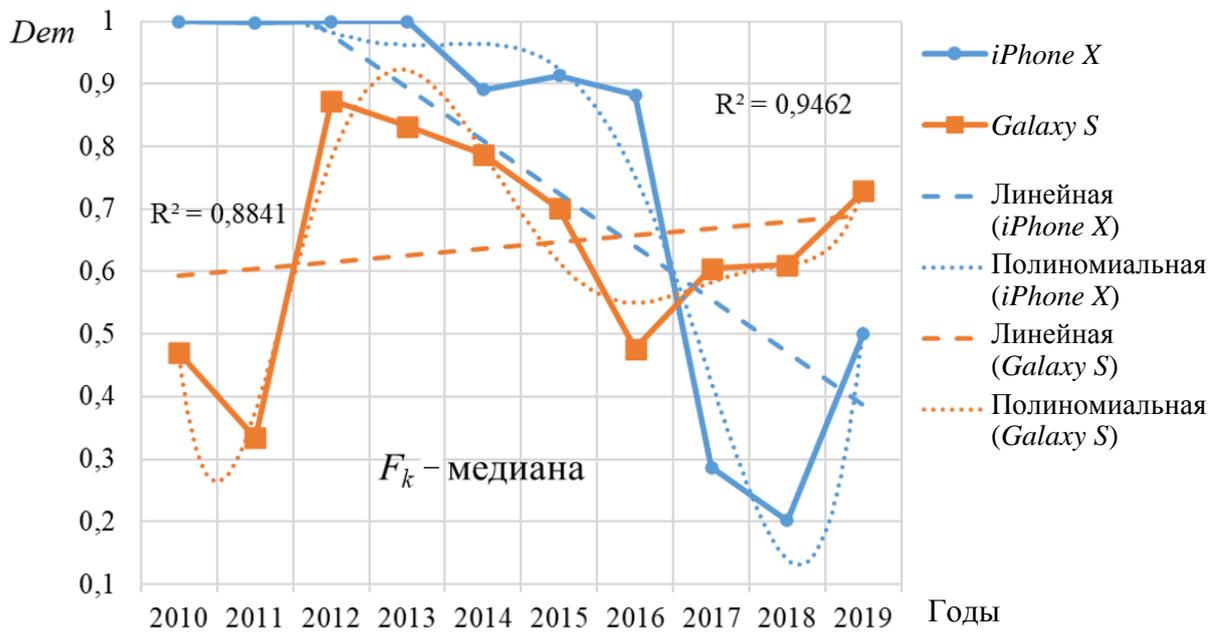


Рис. 3. Динамика изменения востребованности объектов *iPhone X* и *Galaxy S* за 10-летний период (F_k – медиана)

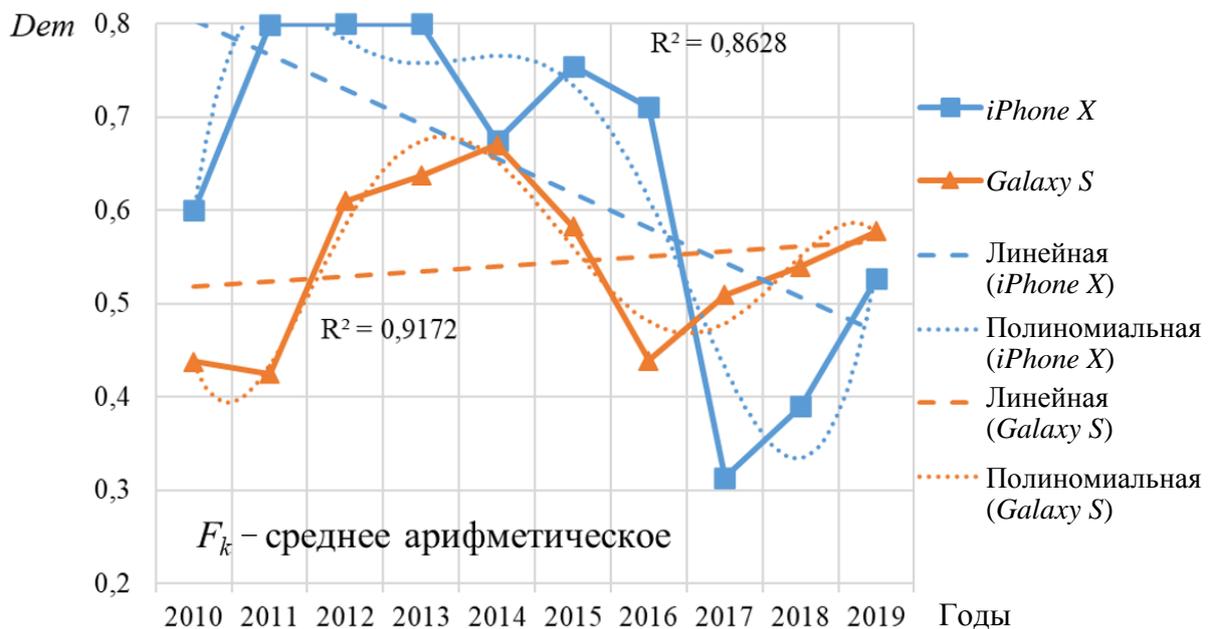


Рис. 4. Динамика изменения востребованности объектов *iPhone X* и *Galaxy S* за 10-летний период (F_k – среднее арифметическое)

На рис. 5 представлена сравнительная диаграмма индикаторов новизны, востребованности и имплементируемости исследуемых объектов, построенная по

данным из двух источников. Диаграмма индикатора востребованности исследуемых объектов в сравнении с экспертной оценкой рейтинга смартфонов (<https://rskrf.ru/ratings/tekhnika-i-elektronika/electronic/smartfony->) «Российской системы качества» показана на рис. 6.

В ходе проведения экспериментов был получен достаточный объем данных для первичного анализа динамики изменений показателей инновационности объектов и сравнения их с экспертными оценками. На рис. 1–6 приведена часть этих данных.

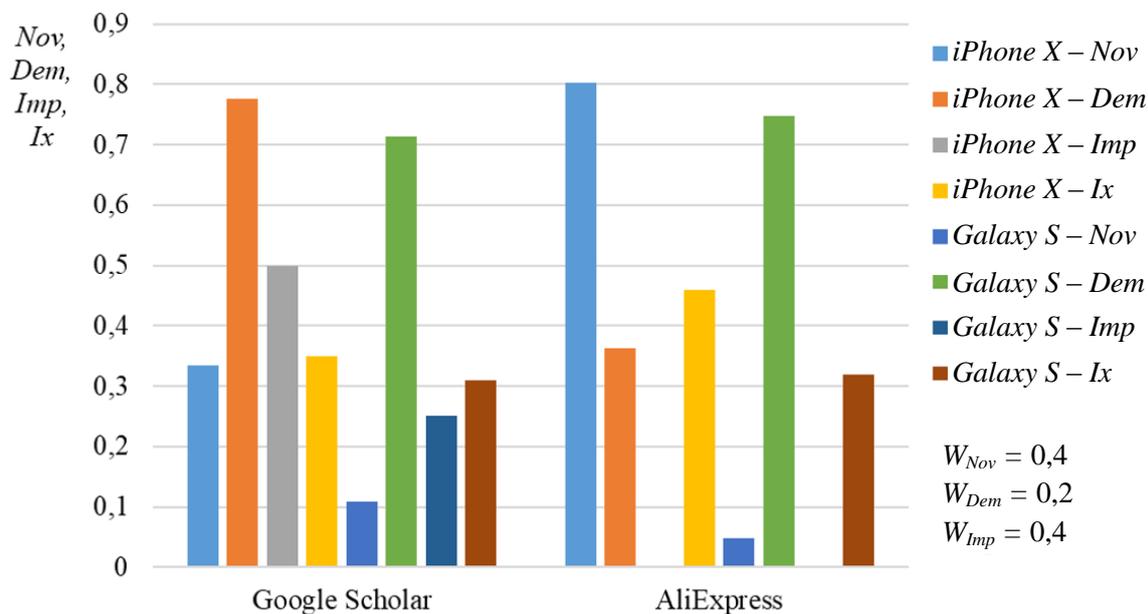


Рис. 5. Индикаторы инновационности исследуемых объектов *iPhone X* и *Galaxy S*

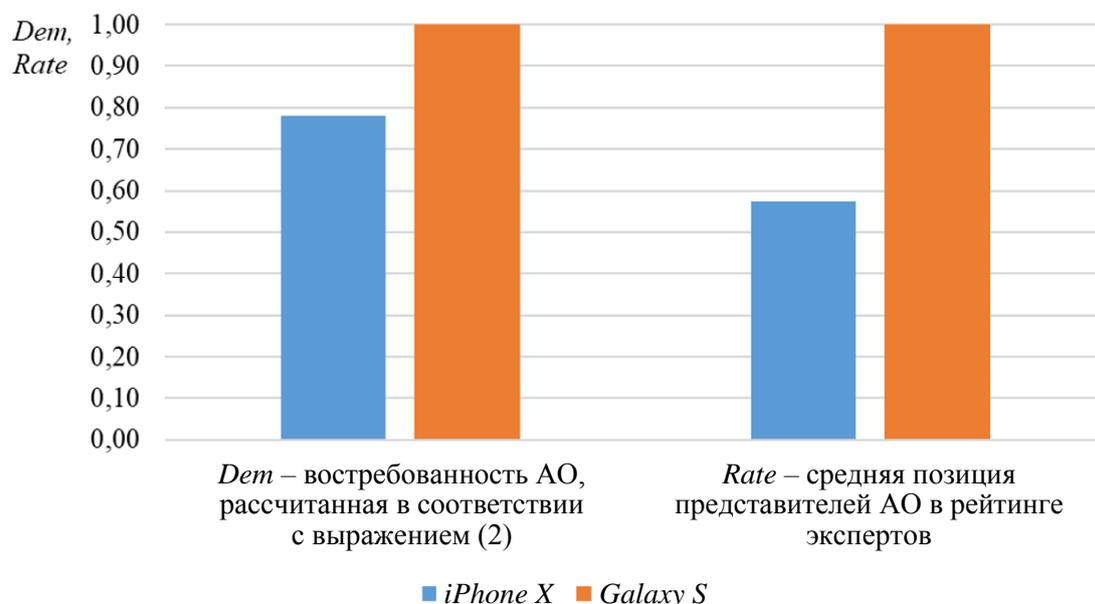


Рис. 6. Вычисленный индикатор востребованности *Dem* в сравнении с экспертной оценкой

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Очевидна цикличность изменений анализируемых показателей (см. рис. 1–4), которая соответствует предположению о наличии инновационных циклов в конкретной области применения.

Для рассмотренной области применения (смартфоны) установленная цикличность появления инновационных продуктов составляет 5–6 лет (см. рис. 1–4). Эта периодичность требует дополнительного анализа.

Анализ соотношений показателей новизны о востребованности объектов выявляет определенные закономерности: падение востребованности вызывает рост новизны объекта, а рост востребованности объекта приводит к падению его новизны, что подтверждают рис. 1–3.

Полученные данные позволяют оценить некоторые особенности производства и продвижения продуктов. Так, восходящий тренд новизны и нисходящий тренд востребованности продуктов *iPhone X* показывают высокий производственный потенциал производителя и наличие резервов в маркетинге. Для продуктов *Samsung Galaxy S* идентифицируется наличие производственных резервов (нисходящий тренд новизны) и стабильный уровень маркетинга (флэт востребованности продуктов) (см. рис. 1, 3 и 4).

Особенностью аддитивного критерия является возможность взаимной компенсации значений локальных критериев при вычислении глобального оценочного критерия. Эта ситуация показана на рис. 5. Здесь значения показателей *Nov* и *Dem* для продуктов *iPhone X* существенно отличаются для различных источников исходных данных. Кроме того, соотношения между значениями как *Nov*, так и *Dem* обратны для этих источников. При этом значения глобального критерия I_x для обоих источников сопоставимы.

Аппроксимация полученных значений показателей инновационности позволяет оценить (и в определенном смысле прогнозировать) их динамику.

Чувствительность модели вычислений к виду нормирующей функции очевидна: существенно различаются значения *Nov* (см. рис. 2). Отметим некоторые особенности. Линейная аппроксимация дает практически одинаковый угловой коэффициент для всех трех вариантов, но различные величины сдвига, полиномиальная показывает одинаковую периодичность *Nov* для всех вариантов, что говорит об отсутствии влияния нормирующей функции на вычисление значения *Imp*. Более точная оценка условий применения методов нормирования требует проведения дополнительных исследований.

Полученные экспериментальные данные хорошо согласуются с экспертными оценками исследуемых объектов (см. рис. 6). Показано соотношение показателей востребованности разных объектов, вычисленных в соответствии с формулой (2). Аналогичное соотношение наблюдается при анализе результатов экспертизы. Результаты вычислений и экспертизы востребованности объектов вполне сравнимы, так как методикой экспертных оценок смартфонов предусматривается выбор параметров оценки и их значимости исходя из предпочтений пользователей.

Теоретическое обоснование типа оценочного критерия базируется на положениях теории полезности и теории свидетельств. Их достаточная апробированность позволяет надеяться на адекватность разработанных моделей.

ЗАКЛЮЧЕНИЕ

Представленная в статье модель используется при разработке вычислительных алгоритмов для отбора информации об объектах с инновационным потенциалом. Экспериментальные данные позволили усовершенствовать методику расчета пока-

зателей инновационности объектов, включая совместную обработку нечетких данных, полученных из разных поисковых систем. Результаты исследования подтвердили основные предположения, используемые при разработке модели. В дальнейших исследованиях предполагается расширить экспериментальную базу для получения более статистически значимых результатов.

БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке РФФИ
(проект № 18-07-00358 и проект № 20-07-00199).

ЛИТЕРАТУРА

1. Организация и поддержка хранилища данных на основе интеллектуализации поискового агента и эволюционной модели отбора целевой информации. Описание проекта. URL: https://www.rfbr.ru/rffi/ru/project_search/o_2071601 (дата обращения: 30.05.2020).
2. Tucker R.B. Driving growth through innovation: how leading firms are transforming their futures. 2nd ed. Berrett-Koehler Publishers, San Francisco, 2008. 224 p.
3. Mortensen P.S., Bloch C.W. Oslo Manual-Guidelines for collecting and interpreting innovation data // Organisation for Economic Cooperation and Development. Paris: OECD, 2005.
4. Фишберн П. Теория полезности для принятия решений. М.: Наука, 1978. 352 с.
5. Кини Р.Л., Райфа Х. Принятие решений при многих критериях: предпочтения и замещения: пер. с англ. / под ред. И.Ф. Шахнова. М.: Радио и связь, 1981. 560 с.
6. Shafer G. A Mathematical theory of evidence. Princeton, N.J.: Princeton University Press, 1976.
7. Yager R., Liping Liu. Classic works of the Dempster – Shafer theory of belief functions. London: Springer, 2010.
8. Ivanov V.K., Palyukh B.V., Sotnikov A.N. Efficiency of genetic algorithm for subject search queries // *Lobachevskii Journal of Mathematics*. 2016. V. 37. № 3. P. 244–254.
9. Ivanov V.K., Palyukh B.V., Sotnikov A.N. Features of data warehouse support based on a search agent and an evolutionary model for innovation information selection. *Proceedings of the Fourth International Scientific Conference “Intelligent Information Technologies for Industry” (ITI’19)*, Springer Nature Switzerland AG, 2020. P. 1–11.

Для цитирования: Иванов В.К. Экспериментальная проверка модели оценки инновационности объекта // Вестник Тверского государственного технического университета. Серия «Технические науки». 2020. № 4 (8). С. 54–63.

EXPERIMENTAL VERIFICATION OF MODEL FOR OBJECT INNOVATION EVALUATION

V.K. IVANOV, Cand. Sc.

Tver State Technical University, 22, Af. Nikitin emb., 170026, Tver,
Russian Federation, e-mail: mtivk@mail.ru

The article discusses the approach for evaluating the innovation index of the products and technologies. The evaluation results can be used to create a warehouse for the object descriptions with significant innovation potential. The innovation index calculation model is

*Вестник Тверского государственного технического университета.
Серия «Технические науки». № 4 (8), 2020*

based on the concepts of novelty, relevance, and implementability of the object. Formal definitions of these indicators are given and a methodology for their calculation are described. The fuzzy methods to process (incomplete) data from numerous sources and to obtain probabilistic innovation assessments are used. The experimental data of the model verification including the calculations of local criteria and global additive evaluation criterion are presented. The cyclical nature of dynamic changes in indicators, their interdependence and some general features of product promotion have been established. The obtained experimental data check with expert judgement of the investigated products. The analysis of the local criteria used in the research affords ground to assert about the correct use of the additive n-dimensional utility function. The adequacy of assumptions and formal expressions that are used in computational algorithms for selection information for data warehouse is confirmed.

Keywords: data warehouse, innovation, additive criterion, utility function, search query.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research, project № 18-07-00358 and project № 20-07-00199.

REFERENCES

1. Data Warehousing Based on Search Agent Intellectualization and Evolutionary Model of Target Information Selection. The project summary. URL: https://www.rfbr.ru/rffi/ru/project_search/o_2071601 (data accessed: 07.05.2020).
2. Tucker R.B. Driving growth through innovation: how leading firms are transforming their futures. 2nd ed. Berrett-Koehler Publishers, San Francisco, 2008. 224 p.
3. Mortensen P.S., Bloch C.W. Oslo Manual-Guidelines for collecting and interpreting innovation data. Organization for Economic Cooperation and Development. Paris: OECD, 2005.
4. Fishburn P. Teoriya poleznosti dlya prinyatiya resheniy. [The theory of utility for decision making]. Moscow: Nauka, 1978. 352 p.
5. Kini R.L., Raifa X. Prinyatiye resheniy pri mnogikh kriteriyakh: predpochteniya i zameshcheniya. [Decision making under many criteria: preferences and substitutions] Ed. I.F. Shakhnov. Moscow: Radio i svyaz, 1981. 560 p.
6. Shafer G. A Mathematical theory of evidence. Princeton, N.J.: Princeton University Press, 1976.
7. Yager R., Liping Liu. Classic works of the Dempster – Shafer theory of belief functions. London: Springer, 2010.
8. Ivanov V.K., Palyukh B.V., Sotnikov A.N., Efficiency of genetic algorithm for subject search queries. *Lobachevskii Journal of Mathematics*. 2016. V. 37. No. 3. Pp. 244–254.
9. Ivanov V.K., Palyukh B.V., Sotnikov A.N. Features of data warehouse support based on a search agent and an evolutionary model for innovation information selection. *Proceedings of the Fourth International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’19)*, Springer Nature Switzerland AG. 2020, pp. 1–11.

Поступила в редакцию/received: 17.07.2020; после рецензирования/revised: 16.10.2020; принята/accepted 15.11.2020