

Информационные процессы / Information processes

Оригинальная статья / Original article

УДК 004.9:025.3

[https://doi.org/ 10.31432/1994-2443-2024-19-2-75-83](https://doi.org/10.31432/1994-2443-2024-19-2-75-83)**Анализ рынка труда с применением рандомизированных систем итерированных функций********А. К. Горностаев** ✉, **А. Л. Огарок***Российский Технологический Университет — МИРЭА, Москва, Российская Федерация*✉ agornostaevv@gmail.com

Аннотация. *Актуальность.* В настоящее время на рынке труда остро стоит задача поиска подходящих кандидатов. За прошедшие 5 лет рынок информационных технологий сильно изменился. В связи с оттоком и новыми способами стимуляции трудоустройства, компаниям стало сложнее отбирать квалифицированные кадры для выполнения необходимых задач. *Цель.* Применение рандомизированных систем итерированных функций для анализа рынка соискателей и подбора наиболее подходящих кандидатов по заданному профилю из сферы информационных технологий. *Материалы и методы.* Используются данные, собранные из открытых источников. Анализ проводился с применением рандомизированных систем итерированных функций. В работе использовался метод кластерного анализа, производился отбор типичных представителей каждого класса и применялись рандомизированные системы итерированных функций для построения фрактальных множеств. *Результаты.* Сформированы 3 группы кандидатов с различными характеристиками. *Выводы.* Применение рандомизированных систем итерированных функций позволяет эффективно анализировать рынок труда. Предложенный подход повышает качество подбора кандидатов.

Ключевые слова: классификация; рандомизированные системы итерированных функций; рынок труда; фрактальные множества

Финансирование. Финансирование отсутствовало.

Для цитирования: Горностаев А. К., Огарок А. Л. Анализ рынка труда с применением рандомизированных систем итерированных функций. *Информация и инновации.* 2024;19(2):75-83. <https://doi.org/10.31432/1994-2443-2024-19-2-75-83>

**** Статья написана по материалам доклада на международной научной конференции «Наука и технологии: источники данных и аналитические подходы в целях развития», 29–30 мая 2024, г. Москва, Россия.

Labor market analysis using randomized systems of iterated functions^{****}

A. K. Gornostaev ✉, A. L. Ogarok

Russian Technological University — MIREA, Moscow, Russian Federation

✉ agornostaevv@gmail.com

Abstract. *Relevance.* Currently, the labor market faces the pressing need to find suitable candidates. Over the past 5 years, the IT market has changed significantly. Due to the outflow and new ways to stimulate employment, it has become more difficult for companies to select qualified personnel to perform the necessary tasks.

Purpose. Using randomized systems of iterated functions to analyze the job seekers' market and select the most suitable candidates for a given profile from the field of information technology.

Materials and methods. The data collected from open sources were used. The analysis was carried out using randomized systems of iterated functions. In the work, the cluster analysis method was used, typical representatives of each class were selected, and randomized systems of iterated functions were used to construct fractal sets.

Results. Three groups of candidates with different characteristics were formed.

Conclusions. The use of randomized systems of iterated functions allows you to effectively analyze the labor market. The proposed approach improves the quality of candidate selection.

Keywords: classification; randomized systems of iterated functions; labor market; fractal sets

Funding. No funding.

For citation: Gornostaev A. K., Ogarok A. L. Labor market analysis using randomized systems of iterated functions. *Information and Innovations*. 2024;19(2):75-83. (In Russ.). <https://doi.org/10.31432/1994-2443-2024-19-2-75-83>

^{****} The article is based on the materials of the report at the international scientific conference «Science and Technology: Data Sources and Analytical Approaches for Development», May 29–30, 2024, Moscow, Russia.

Введение

В настоящее время актуальной нерешенной задачей является разработка научно-методических подходов к разработке методов поиска соискателей и отбора наиболее подходящих сотрудников по заданному профилю специализации в сфере разработки информационных технологий. Анализ подходов к решению поставленной актуальной научной задачи показал целесообразность использования аппарата рандомизированных систем итерированных функций, который подразумевает использование фрактальных множеств [1,2] и ультраметрического пространства для описания поведения данных.

Цель исследования — применение рандомизированных систем итерированных функций для анализа рынка соискателей и подбора наиболее подходящих кандидатов по заданному профилю из сферы информационных технологий при помощи воссоздания множества в ультраметрическом пространстве и перехода в признаковое для дальнейшего анализа.

В рамках исследования описываются рандомизированные системы итерированных функций (РСИФ) — математические модели, которые используются для построения фрактальных множеств. Данная модель является системой итерированных функций, в которой на каждом шаге итерации с заданным вероятностным распределением реализуется только одна из функций из заданного набора.

Методические подходы к решению задачи

Существует два способа построения РСИФ [3].

Первый подход, $F1$, подразумевает определение набора линейных функций $\{f_1(x), f_2(x), \dots, f_K(x)\}$, в котором каждая функция имеет вид:

$$f^{(j)}(x_i) = (1 - \xi)x_i + \xi Z_j^i, \quad (1)$$

где: ξ — фиксированный параметр;

$\{Z_1, Z_2, \dots, Z_K\}$ — параметры, определяющие вид линейной функции.

Задается вероятностное распределение $\{p_1, p_2, \dots, p_K\}$ на множестве функций $f^{(j)}$. Для выбранной начальной точки x_0 на каждой итерации вначале случайно выбирается функция в соответствии с вероятностным распределением, а затем применяется выбранная функция $f^{(j)}$ к текущей точке $x_{i+1} = f^{(j)}(x_i)$, где $j = 1, \dots, N$.

Множество точек $\{x_1, x_2, \dots, x_N\}$, полученных в результате N итераций, называется предфракталом. Предельное множество, к которому сходится последовательность $\{x_1, x_2, \dots, x_N\}$, при неограниченном увеличении N , называется аттрактором РСИФ.

Во втором подходе, именуемым $F2$, рассматривается совокупность членов ряда

$$\frac{1}{\mu} \sum_{i=1}^{\infty} \xi^i = 1, \quad (2)$$

где $\mu = \frac{1}{1-\xi}$ — нормировочная константа.

Эта совокупность разбивается на K заранее заданное число классов $\{S_1, S_2, \dots, S_K\}$ таким образом, что каждый элемент матрицы A принадлежит ровно одному классу. Для каждого класса S_i определяется его состав элементов входящих в этот класс — a_i .

В результате, полученное множество строк $\{A_1, A_2, \dots, A_K\}$ будет представлять собой классификационное пространство A . Матрица X , представляющая фрактальное множество ультраметрического пространства в пространстве признаком, получается умножением A на Z .

Стоит также отметить, что в рамках данной работы решение задачи осуществляется в двух пространствах:

– признаковое — где существуют исходные данные;

– классификационное — ультраметрическое пространство, где существует матрица A .

На строках матрицы A задается метрика, которая, как показано в [3, 4], является ультраметрикой. Переход из классификационного пространства в признаковое осуществляется посредством перемножения матрицы A и матрицы протофрактала: $X = AZ$.

Данные представляют собой резюме соискателей, собранные в рамках одного профиля. Объем выборки, после предварительного анализа составляет 96 респондентов. Совместно с экспертами было выделено 16 признаков, по которым можно оценить человека при поиске работы на конкретную должность. Общее множество признаков было разделено на две группы [5]:

hard-skills — признаки, описывающие работу человека с конкретными технологиями или со специфичным программным обеспечением;

soft-skills — признаки, описывающие отношение человека к работе в целом, например, насколько в его деятельности выражена трудовая мобильность, а также социально-демографические данные.

В качестве целевой переменной выступает оценка специалиста по подбору персонала каждого резюме по десяти бальной системе. В конкретной задаче предлагается свести бальную систему к трем группам оценок:

- оценка от 1 до 4 — кандидат не подходит на должность;
- оценка от 5 до 7 — с кандидатом можно поговорить, но не в первую очередь;
- оценка от 8 до 10 — кандидат подходит под все или многие критерии.

К описанным выше данным применяется алгоритм кластеризации k -Means

для разделения на 3 кластера. Визуализация данных в признаковом пространстве представлена на рис. 1.

На рис. 1 приняты следующие условные обозначения: квадрат — группа номер 1, треугольник — группа номер 2, круг — группа номер 3. Крестиками отмечены типичные представители для каждой группы.

При использовании подхода РСИФ важно выбрать типичных представителей каждого класса. В большинстве случаев используется подход, при котором в каждом кластере находится центроид, который берется в качестве такого элемента. Однако стоит отметить, что типичным представителем является элемент, наиболее полно описывающий свойства своего класса. Так, в работе [6, с. 13–19] показывается, что для группы «Бедных» (по уровню дохода) типичным представителем будет самый бедный, для «Богатых» — самый богатый, а для «Среднего класса» — средний. Такой выбор типичных объектов в полной мере соответствует фрактальной теории [7].

В рамках данной задачи для каждой группы были выбраны типичные представители (см. рис. 1):

- 1–4 — объекты, заявленный уровень которых, не устраивает экспертов;
- 5–7 — объекты, в отношении которых, с достаточной уверенностью, невозможно определить их уровень;
- 8–10 — объекты, заявленный уровень которых, полностью соответствует требованиям.

Типичные элементы являются элементами протофрактала Z , на основе которого формируется матрица A методом $F2$. Можно получить представление классификационного пространства в признаковом пространстве путем перемножения $X = AZ$ (рис. 2).

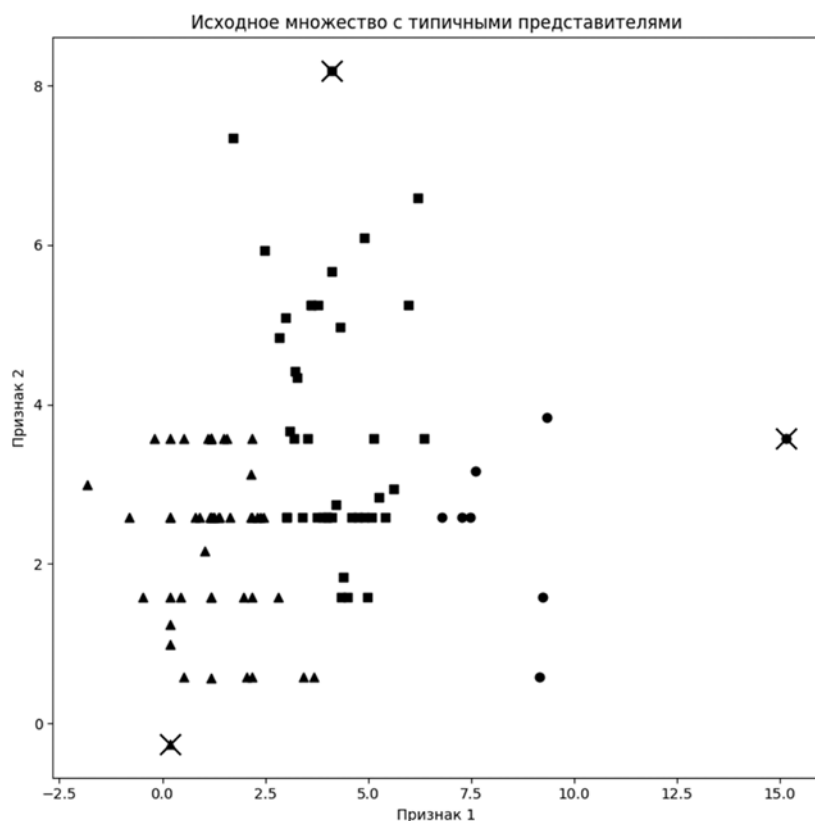


Рис. 1. Визуализация данных в признаковом пространстве

Fig. 1. Visualization of data in feature space

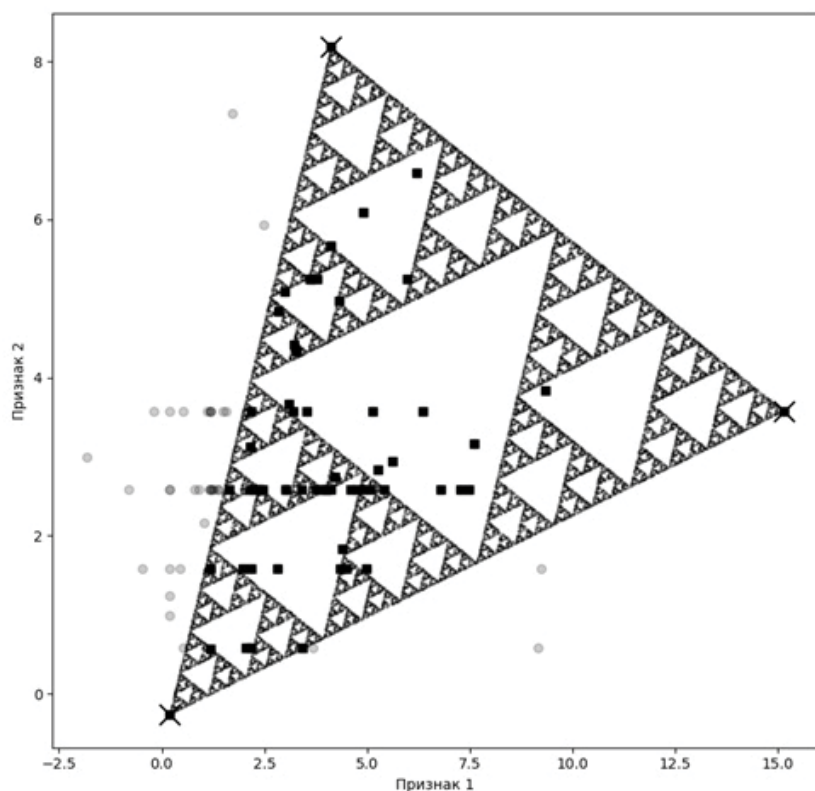


Рис. 2. Представление классификационного пространства в признаковом пространстве

Fig. 2. Representation of the classification space in the feature space

На рис. 2 фрактальное множество, построенное на элементах протофрактала (крестики), разбитого на три группы. Так как фрактальное множество образует треугольник Серпинского, то для выбора точек, находящихся внутри данной фигуры, можно использовать барицентрические координаты (рис. 3).

Элементы исходного множества, находящиеся внутри фрактального множества (квадратики), элементы не вошедшие (круги) и типичные представители (крестики). Сгруппируем элементы исходного множества, которые попали внутрь выпуклой оболочки множества протофрактала, относительно типичных представителей по расстоянию Евклида (рис. 4).

На рис. 4 группа номер 1 — обозначена треугольником, группа номер 2 — кругом, группа номер 3 — квадратом. Анализируя каждую получившуюся группу [8], можно заметить, что во вторую группу попали элементы с высоким исходным рейтингом. В первой группе находятся кандидаты с высокой заработной платой, с опытом работы в крупных компаниях и часто меняющие работу, то есть демонстрируют трудовую мобильность, что может свидетельствовать о нестабильности таких кандидатов при найме и высоком риске их ухода из компании. В третьей группе оказались соискатели с более низкими зарплатными ожиданиями, с меньшим опытом работы в крупных компаниях, но при этом частота смены работы ниже (де-

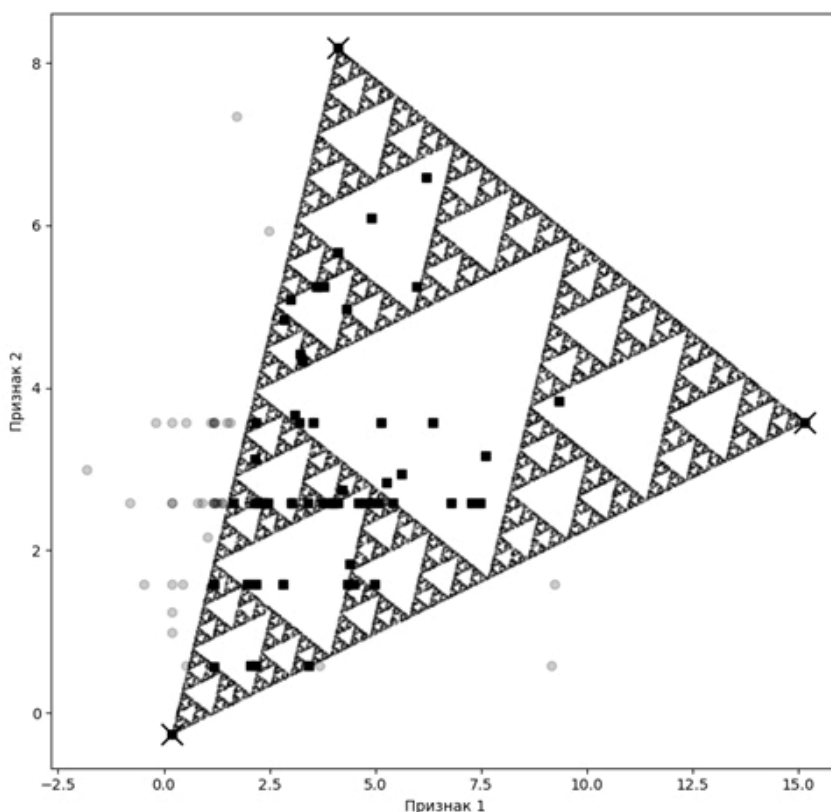


Рис. 3. Представление классификационного пространства в признаковом пространстве с барицентрическими координатами

Fig. 3. Representation of the classification space in the feature space with barycentric coordinates

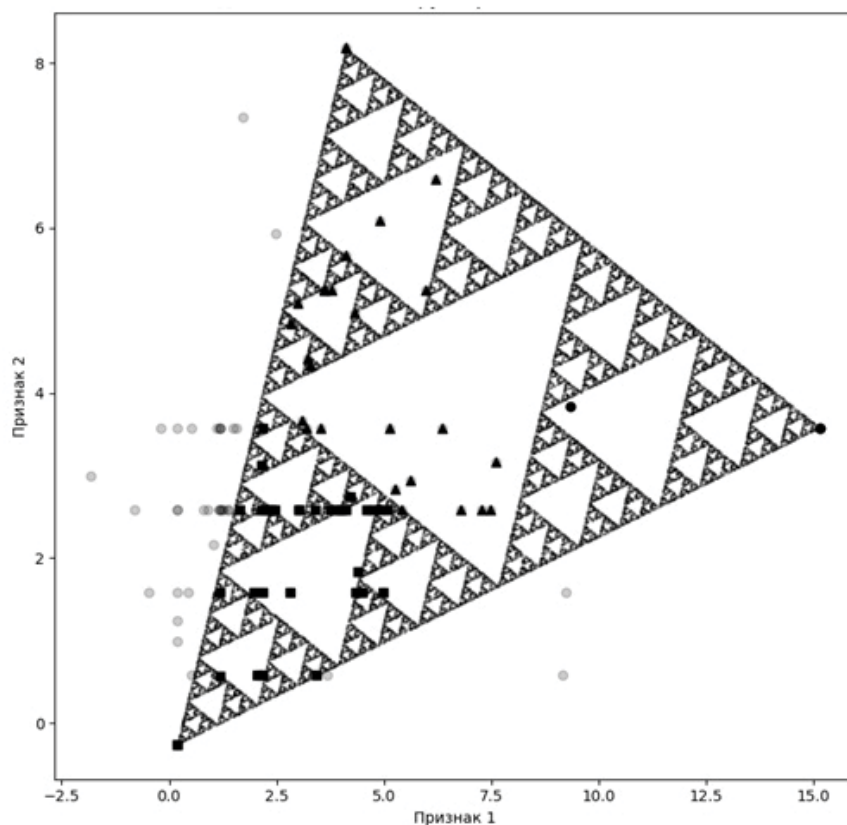


Рис. 4. Элементы исходного множества, сгруппированные по цветам относительно типичных представителей каждого класса

Fig. 4. Elements of the original set, grouped by colors relative to typical representatives of each class

монстрируют трудовую стабильность), то есть данный класс описывает людей более склонных к стабильной планомерной работе [9].

Прикладные результаты решения задачи

Подводя итоги решения поставленной научной задачи можно отметить, что с использованием представления профиля человека в признаковом пространстве признаков в результате анализа удалось сформировать 3 группы с различными свойствами. Каждую из групп можно анализировать в рамках конкретной вакансии и в конкретной компании. Так, например, для срочных кратковремен-

ных проектов, подойдут кандидаты из первой группы, а для длительной работы лучше выбрать людей из третьей группы. Если задача требует решения грамотным специалистом с опытом, стоит рассмотреть кандидатов из второй группы. Стоит также отметить, что в группе номер 2 оказалось меньше кандидатов по сравнению с остальными группами, что подтверждает положение о том, что найти грамотных специалистов трудно.

Заключение

В данном исследовании описан научно-методический подход к формализации постановки и решения актуальной задачи анализа рынка труда

на примере соискателей Российской Федерации и подбора наиболее подходящих кандидатов по заданному профилю из сферы IT с применением рандомизированных систем итерированных функций. Были выбраны типичные представители, относительно которых, в дальнейшем, формировались группы, полученные с помощью РСИФ. Произведен анализ образованных групп и даны рекомендации для дальнейшего принятия решений при поиске релевантных кандидатов. Для проведения исследования использован экспертный и аналитический подход к формализации и решению научной задачи. Научная новизна предложенного подхода состоит в применении апробированного научно-методического аппарата класте-

ризации данных и теории матриц к прикладной области анализа рынка труда с применением рандомизированных систем итерированных функций. Практическая значимость исследования заключается в том, что предложенный научно-методический подход позволяет повысить качество подбора кандидатов на вакансии применительно к различным критериям отбора специалистов. Апробация указанного научно-методического подхода подтвердила высокую эффективность предложенного научно-методического подхода, повышение качества кластеризации данных и минимизацию времени обработки данных, что особенно актуально для автоматической обработки массивов данных большой размерности.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы заявляют об отсутствии конфликта интересов.

CONFLICT OF INTERESTS

The authors declare no relevant conflict of interests.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Hutcthinson J.E. Fractals and self similarity. *Indiana University Mathematics Journal*. 1981;30(5):713–747. <https://doi:10.1887/0750304006/b384c2>
2. Barnsley F. Hawley Rising. *Fractals Everywhere*. Academic Press Professional, 1993. 531 p.
3. Буховец А.Г., Бирючинская Т.Я. Структура аттрактора рандомизированных систем интегрированных линейных функций. 2016;(2):5–10. *Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии*. EDN WGBHOT
Bukhovets A. G., Biryuchinskaya T. Ya. The structure of the attractor of randomized systems of iterated linear functions. *Bulletin of Voronezh State University. Series: System Analysis and Information Technology*. 2016;5.
4. Хренников А. Ю. Моделирование процессов мышления в p-адических системах координат. М.: ФИЗМАТЛИТ, 2004. 296 с.
Khrennikov A. Y. Modeling of thinking processes in mathematical coordinate systems. М.: FIZMATLIT, 2004. 296 p.
5. Lamri J, Lubart T. Reconciling Hard Skills and Soft Skills in a Common Framework: The Generic Skills Component Approach. *Journal of Intelligence*. 2023 Jun 1;11(6):107. doi: 10.3390/jintelligence11060107.

6. Буховец А.Г., Бирючинская Т.Я., Лаврова Е. О. Выбор типичных объектов в классификационных задачах. Социологические методы в современной исследовательской практике: Сборник статей, посвященный памяти первого декана факультета социологии НИУ ВШЭ А. О. Крыштановского, Москва, 22 февраля 2011 года. Москва: Национальный исследовательский университет «Высшая школа экономики», 2011. 558 с. EDN QBUUNJ
Bukhovets A. G., Biryuchinskaya T. Ya., Lavrova E. O. The choice of typical objects in classification problems. Sociological methods in modern research practice: A collection of articles dedicated to the memory of the first Dean of the Faculty of Sociology of the Higher School of Economics A. O. Kryshtanovsky, Moscow, February 22, 2011. Moscow: National Research University Higher School of Economics, 2011. 558 p.
7. Mandelbrot BB. Fractals and Scaling in Finance. Springer New York, 1997. 551 p. <https://doi.org/10.1007/978-1-4757-2763-0>
8. Буховец А.Г., Семин Е. А., Горностаев А. К. О новых подходах к прогнозированию урожайности зерновых культур. *Современная экономика: проблемы и решения*. 2023; № 2(158):8–19. EDN QGJBJO. <https://doi:10.17308/meps/2078-9017>
Bukhovets A. G., Semin E. A., Gornostaev A. K. On new approaches to forecasting the yield of grain crops. *Modern economy: problems and solutions*. 2023; No2(158):8–19. <https://doi:10.17308/meps/2078-9017>
9. Обухова Л. А. Трудовая мобильность как фактор формирования предложения на рынке труда. Вестник КРАСГАУ. 2006;(10):28–33.
Obukhova L. A. Labor mobility as a factor in the formation of supply in the labor market. Vestnik of KRASGAU. 2006;(10):28–33.

ИНФОРМАЦИЯ ОБ АВТОРАХ

Горностаев А. К., Российский Технологический Университет — МИРЭА, Москва, Российская Федерация, ORCID ID — 0009-0008-8033-7135, agornostaevv@gmail.com

Огарок А. Л., кандидат технических наук, Российский Технологический Университет — МИРЭА, доцент, старший научный сотрудник кафедры БК 231, Москва, Российская Федерация, ORCID ID — 0000-0002-3125-4594, aogarok@gmail.com

INFORMATION ABOUT THE AUTHOR

Gornostaev A. K., Master Student, Russian Technological University — MIREA, Moscow, Russian Federation, ORCID ID — 0009-0008-8033-7135, agornostaevv@gmail.com

Ogarok A. L., Candidate of Technical Sciences, Russian Technological University — MIREA, Moscow, Russian Federation, Associate Professor, Senior Researcher at the BK 231 Department, ORCID ID — 0000-0002-3125-4594, aogarok@gmail.com

Поступила в редакцию / Received 05.06.2024

