

В. А. ВИШНЯКОВ, Е. И. ПОЛОСКО

ПРОГРАММНЫЕ СРЕДСТВА АВТОМАТИЗАЦИИ РАСЧЕТОВ ДЛЯ ФАКТОРНОГО, РЕГРЕССИОННОГО, КОРРЕЛЯЦИОННОГО АНАЛИЗОВ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Аннотация. Предметом исследований является анализ применения средств автоматизации расчетов для использования факторного, регрессионного, корреляционного анализов. Цель статьи: приведение известного и авторского опыта использования языка Пайтон для применения факторного, регрессионного и корреляционного анализов и опыт авторов в их применении. Рассмотрены основные аспекты теоретических основ данных методов анализа. Проведена классификация программных средств для автоматизации статистических расчетов. Рассмотрены преимущества и ограничения языков программирования (Python, R), коммерческих платформ (SPSS, SAS, Stata), BI-платформ (Tableau, Power BI, Zoho Analytics), открытых платформ (KNIME, RapidMiner), специализированных статистических пакетов (JASP, Jamovi). Детализирована методология применения автоматизированного статистического анализа, включая подготовку исходных данных, выбор метода, настройку параметров, выполнение расчетов и получение первичных результатов. Приведена авторская методика автоматизации комплексного анализа для оценки использования нейронных сетей, Интернета вещей (IoT) в учебном процессе университета.

Ключевые слова: факторный, регрессионный, корреляционный, анализ, автоматизация, языки, платформы, системы

Введение

Современная эпоха характеризуется беспрецедентным ростом объемов данных, что делает необходимым применение специализированного программного обеспечения для их анализа. Автоматизация статистических расчетов становится не просто удобным инструментом, а обязательным условием эффективного исследования в экономике, технических науках, социологии и других областях. Факторный, регрессионный и корреляционный анализы как методы, требующие сложных математических вычислений и обработки значительных массивов информации, особенно выигрывают от автоматизации [1]. Использование специализированных программ позволяет не только ускорить процесс получения результатов, но и минимизировать человеческие ошибки, обеспечить воспроизводимость исследований и сосредоточить усилия исследователя на интерпретации результатов, а не на рутинных вычислениях. В статье рассматривается современное состояние программных средств автоматизации для выполнения указанных видов анализа, их классификация, возможности и практическое применение, авторские разработки в этой области.

Теоретические основы методов анализа

Корреляционный анализ представляет собой статистический метод, позволяющий определить наличие и силу зависимости между переменными. В автоматизированных системах этот метод реализуется через вычисление коэффициентов корреляции, которые количественно выражают степень и направление

связи. Как отмечается в исследованиях, коэффициент корреляции может принимать значения от -1 до $+1$, где значения близкие к $+1$ указывают на положительную линейную связь, близкие к -1 указывают на отрицательную линейную связь, а около нуля – на отсутствие статистической зависимости [2].

Современное программное обеспечение позволяет автоматически вычислять различные виды коэффициентов корреляции (Пирсона, Спирмена), визуализировать корреляционные матрицы, оценивать статистическую значимость полученных связей. Особую ценность представляет возможность быстрого анализа множества переменных одновременно и выявления наиболее значимых взаимосвязей в многомерных данных, что практически невозможно при ручном расчете.

Регрессионный анализ – более сложный метод, который не только устанавливает наличие связи между переменными, но и позволяет построить ее математическое выражение. В отличие от корреляционного анализа, который лишь отвечает на вопрос о существовании зависимости, регрессия дает ее количественное описание в виде уравнения, связывающего зависимую переменную с одной или несколькими независимыми переменными [3].

Автоматизация регрессионного анализа в программных средствах включает подбор вида регрессионной модели (линейная, нелинейная, множественная), расчет параметров уравнения с оценкой их статистической значимости, проверку адекватности модели и анализ остатков. Современные инструменты поддерживают сложные виды регрессии, включая логистическую, полиномиальную регрессию и многие другие, реализация которых вручную потребовала бы

значительных вычислительных ресурсов и времени.

Факторный анализ [1] представляет собой метод сокращения размерности данных, позволяющий выявить скрытые переменные (факторы), объясняющие структуру корреляционных связей между наблюдаемыми переменными. Этот метод особенно важен в психологии, социологии, маркетинговых исследованиях, где исследователи имеют дело с множеством взаимосвязанных параметров.

Автоматизация факторного анализа включает процедуры извлечения факторов (метод главных компонент, максимального правдоподобия), определение оптимального количества факторов, их ротацию и интерпретацию. Современные программы позволя-

ют автоматизировать наиболее сложные аспекты анализа, включая оценку качества факторизации и визуализацию результатов в виде факторных диаграмм.

Анализ программных средств автоматизации

Современный рынок программного обеспечения предлагает множество инструментов для автоматизации статистических расчетов, которые можно классифицировать по различным признакам: коммерческие и открытые, программные среды и платформы, облачные и десктопные решения. Рассмотрим классификацию основных инструментов с их ключевыми возможностями (табл. 1) [4 – 6].

Таблица 1. Программные средства для автоматизации статистических расчетов

Категория инструментов	Представители	Возможности для анализа	Преимущества
Языки программирования	Python, R	Реализация всех видов анализа через специализированные библиотеки (pandas, scikit-learn, numpy – для Python; stats, psych – для R)	Гибкость, воспроизводимость, возможность тонкой настройки алгоритмов, интеграция в скрипты автоматизированной обработки данных
Коммерческие платформы	SPSS, SAS, Stata	Комплексная реализация методов с графическим интерфейсом, автоматическая генерация отчетов	Удобный интерфейс, минимальный порог вхождения, техническая поддержка, стабильность работы
BI-платформы	Tableau, Power BI, Zoho Analytics	Визуализация результатов корреляционного и регрессионного анализа, возможность построения прогнозов	Интерактивные дашборды, интеграция с различными источниками данных, удобство для бизнес-пользователей
Открытые платформы	KNIME, RapidMiner	Визуальное проектирование аналитических workflow, включая все рассматриваемые виды анализа	Открытость, модульность, возможность расширения функциональности, активное сообщество пользователей
Специализированные статистические пакеты	JASP, Jamovi	Полный набор методов статистического анализа с акцентом на удобство интерфейса	Бесплатность, ориентированность на академическое сообщество, современный интерфейс

Инструменты на основе языков программирования. Язык *R* был создан специально для статистического анализа и на сегодняшний день обладает наиболее полным набором методов для выполнения факторного, регрессионного и корреляционного анализа. Благодаря наличию специализированных пакетов (factoextra для факторного анализа, corrplot для визуализации корреляций, regression для регрессионного моделирования), *R* остается одним из наиболее востребованных инструментов среди профессиональных статистиков [6]. Сильной стороной *R* является возможность детальной настройки всех параметров анализа и генерации публикационно-качественных графиков.

Python как универсальный язык программирования также предоставляет широкие возможности для автоматизации статистических расчетов через библиотеки, такие как pandas для обработки данных, scikit-learn для построения моделей, statsmodels для статистического тестирования. Преимущество *Python* перед *R* в его универсальности и лучших возможностях интеграции с веб-приложениями и системами управления данными, что делает его предпочтительным для внедрения аналитических моделей в производственные системы [6].

Коммерческие аналитические платформы [7]. *SPSS* (*Statistical Package for the Social Sciences*) является одним из старейших и наиболее авторитетных программных продуктов в области статистического анализа. Сильной стороной *SPSS* является сбалансированность между мощными аналитическими возможностями и относительно простым графическим интерфейсом, что делает его популярным в академической среде и среди исследователей, не являющихся профессиональными статистиками. Программа предоставляет комплексные средства для выполнения всех рассматриваемых видов анализа с возможностью автоматической генерации отчетов.

SAS (*Statistical Analysis System*) представляет собой более мощную и сложную платформу, традиционно используемую в крупных корпорациях и научных учреждениях. Благодаря своей надежности и возможности обработки чрезвычайно больших объемов данных, *SAS* остается стандартом в фармацевтической промышленности и банковском секторе. Модуль *SAS Statistics* предоставляет исчерпывающий набор процедур для сложных видов регрессионного и факторного анализа.

BI-платформы с аналитическими возможностями [4, 8]. Современные бизнес-платформы, такие как *Tableau* и *Power BI*, постепенно интегрируют в себя функции статистического анализа, делая их доступными для бизнес-пользователей. Хотя эти инструменты не предоставляют такой же глубины анализа, как специализированные статистические пакеты, они позволяют быстро выполнять основные виды корреляционного и регрессионного анализа непосредственно в процессе визуализации данных.

Zoho Analytics выделяется среди аналогичных продуктов наличием AI-ассистента, позволяющего задавать вопросы о данных на естественном языке и автоматически получать аналитические отчеты, включая выявление корреляционных зависимостей и построение прогнозов на основе регрессионных моделей. Это делает сложные методы анализа доступными для пользователей без специальной статистической подготовки.

Открытые платформы с визуальным программированием [4, 5], такие как *KNIME*, *RapidMiner* представляют собой платформы для визуального проектирования аналитических процессов, которые сочетают удобство графического интерфейса с мощностью современных алгоритмов анализа данных. Особенностью этих платформ является модульный принцип организации workflow, где каждый узел представляет определенную операцию по обработке или анализу данных.

Эти инструменты особенно эффективны для автоматизации сложных многошаговых процессов анализа, когда необходимо комбинировать различные методы и последовательно применять их к обновляемым данным. Наличие версий с открытым исходным кодом делает их доступными для образовательных учреждений и малого бизнеса.

Методология внедрения автоматизированного анализа

Эффективное использование программных средств автоматизации требует соблюдения определенной методологии, которая включает несколько последовательных этапов.

Подготовка исходных данных. Качество результатов любого статистического анализа напрямую зависит от качества исходных данных. Современные аналитические платформы предоставляют инструменты для очистки и трансформации данных, а именно: заполнение пропущенных значений; обнаружение и обработка выбросов; проверка на соответствие нормальному распределению; стандартизация и нормализация данных.

Инструменты вроде *OpenRefine* специализируются на автоматизации процессов очистки структурированных данных, что особенно важно при работе с большими объемами информации. Для обеспечения достоверности результатов необходимо использовать методы теории вероятностей и математической статистики для оценки качества данных [9].

Процесс автоматизированного анализа. После подготовки данных осуществляется непосредственно анализ с использованием выбранного программного средства. Этот процесс включает:

- выбор метода анализа, соответствующего поставленным задачам и характеру данных;
- настройка параметров алгоритма (например, выбор метода вращения в факторном анализе или критерия отбора переменных в регрессии);
- выполнение расчетов и получение первичных результатов;
- проверка адекватности модели и соответствия статистическим предположениям.

Важным преимуществом автоматизации является возможность быстрого сравнения альтернативных моделей и методов анализа для нахождения оптимального решения.

Визуализация и интерпретация результатов

Современные аналитические средства предоставляют богатый арсенал для наглядного представления результатов анализа [10]: корреляционные матрицы (heatmaps); диаграммы рассеяния с линиями регрессии; факторные нагрузки и диаграммы вращения; остатки регрессионных моделей и диагностические графики. Эффективная визуализация не только облегчает интерпретацию результатов, но и помогает выявить скрытые закономерности и потенциальные проблемы с моделью.

Заключительным этапом является использование полученных моделей для решения практических задач [11, 12]: разработка рекомендаций по принятию управленческих решений; создание прогнозных систем на основе регрессионных уравнений; построение метрик и показателей на основе факторов; интеграция моделей в производственные системы. Результаты моделирования переносятся на реальный объект, и разрабатываются конкретные рекомендации по решению практических задач.

Система автоматизации расчетов. Авторская система автоматизации включает пользовательский интерфейс, ядро (модули: данных, модели, оптимизации, анализа, визуализации и формирования отчетов), хранилище данных. Она обеспечивает гибкость, масштабируемость и возможность дальнейшего расширения функциональности (рис.1). Система работает в следующей последовательности: инициализация – загрузка конфигурации, создание объектов, подготовка данных; настройка модели – загрузка данных для расчета, построение матриц расстояний; оптимизация – поиск оптимальных объемов поставок для минимизации затрат; анализ чувствительности – исследование влияния параметров цифровизации; сценарный анализ – сравнение различных сценариев развития; визуализация – построение графиков и диаграмм; генерация отчетов – сохранение результатов в различных форматах.

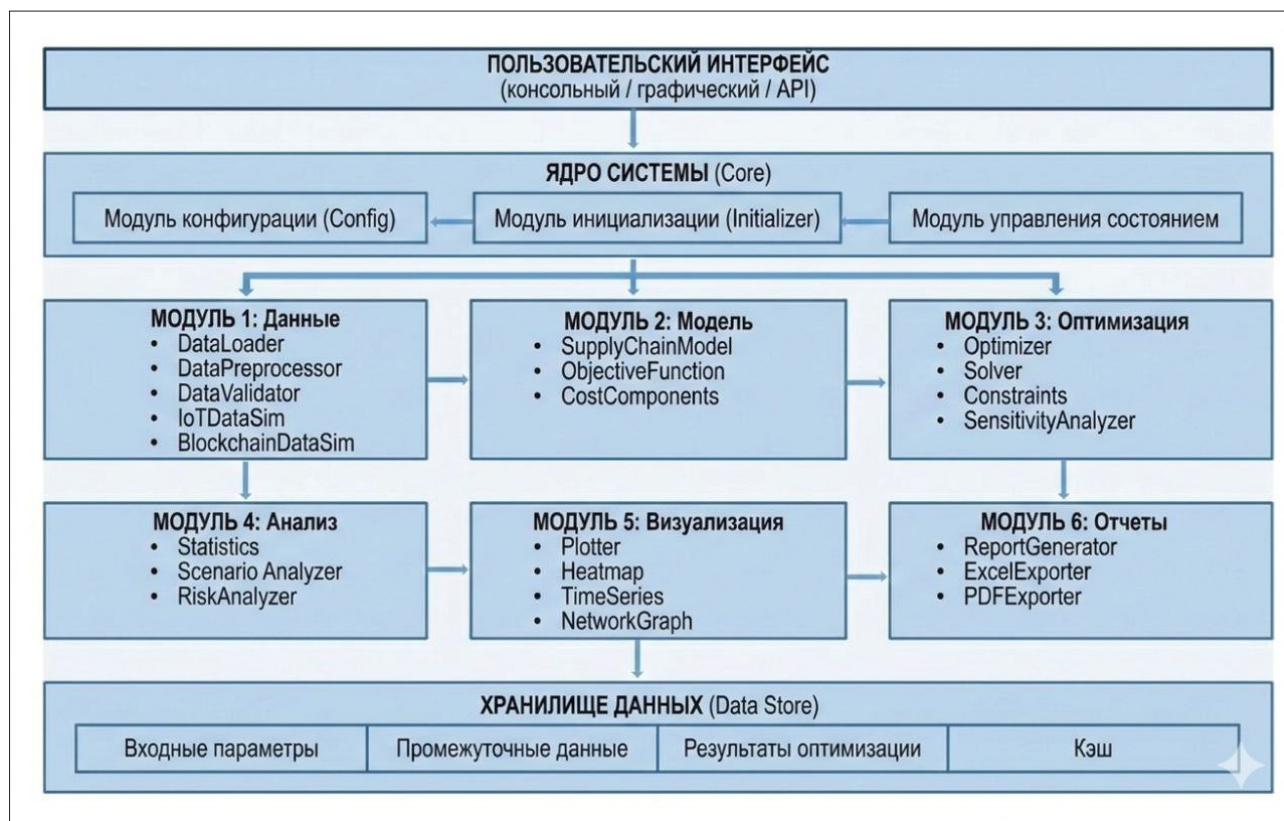


Рисунок 1. Структура системы автоматизации расчетов

Проведенные авторами комплексные анализы с использованием системы на языке *Python*, включающие факторную декомпозицию, регрессионное моделирование и корреляционную оценку для оценки использования нейронных сетей, Интернета вещей в учебном процессе университета, позволили раскрыть структуру восприятия цифровых технологий в образовательной среде [12, 13].

Заключение

1. Автоматизация расчетов факторного, регрессионного и корреляционного анализа представляет собой тенденцию, обусловленную как ростом объемов обрабатываемых данных, так и усложнением методов статистического анализа. Современные программные средства, начиная от специализированных статистических пакетов и заканчивая универсальными языками программирования, упрощают доступ к сложным аналитическим методам, делая их доступными не только для профессиональных статистиков, но и для исследователей с различным уровнем подготовки.

2. Ключевым преимуществом автоматизации является не только ускорение вычислений, но и обе-

спечение воспроизводимости результатов, стандартизация процедур анализа и минимизация субъективного фактора. Программные средства являются именно инструментом, эффективность которого определяется квалификацией и статистической грамотностью исследователя. Интерпретация результатов и содержательные выводы остаются прерогативой человека. Приведена структура системы автоматизации расчетов, которая включает пользовательский интерфейс, ядро (модули данных, модели, оптимизации, анализа, визуализации и формирования отчетов), хранилище данных.

3. Проведенные авторами комплексные анализы с использованием системы на языке *Python*, включающие факторную декомпозицию, регрессионное моделирование и корреляционную оценку для оценки использования нейронных сетей, Интернета вещей в учебном процессе университета, позволили раскрыть структуру восприятия цифровых технологий в образовательной среде. Перспективы развития программных средств автоматизации статистического анализа связаны с дальнейшей интеграцией искусственного интеллекта для автоматического выбора оптимальных методов анализа.

ЛИТЕРАТУРА

1. Хацкевич, Г. А. Эконометрика / Г.А. Хацкевич, Т. В. Русилко. Минск : РИВШ, 2021. 450 с.
2. Аблямин, Р. Ф. Применение корреляционного и регрессионного анализа в системе прогнозирования тех-

нологического процесса / Р. Ф. Абляминов // Научные высказывания. 2023. №21 (45). С. 24–28. URL: https://nvjournal.ru/article/Primenenie_korreljatsionnogo_i_regressionnogo_analiza_v_sisteme_prognozirovaniya_tehnologicheskogo_protsesta (дата обращения: 02.09.2025).

3. Прикладная статистика. Основы эконометрики : в 2 т. 2-е изд., испр. М. : ЮНИТИ-ДАНА, 2001. Т. 1 : Теория вероятностей и прикладная статистика / С. А. Айвазян, В. С. Мхитарян. 656 с.

4. The Top 10 Data Analysis Tools for Perfect Data Management // Software Testing Help. URL: <https://www.softwaretestinghelp.com/data-analysis-tools%20/> (date of access: 02.09.2025).

5. 30 Big Data Tools to Improve Your Analysis in 2025 // Octoparse Blog. URL: <https://www.octoparse.com/blog/top-30-big-data-tools-for-data-analysis-in-2021> (date of access: 02.09.2025).

6. Statistical Consulting // UCLA. URL: <https://stats.oarc.ucla.edu/> (date of access: 15.09.2025).

7. Top 15 Automation Tools for Data Analytics // GeeksforGeeks. URL: <https://www.geeksforgeeks.org/data-science/top-15-automation-tools-for-data-analytics/> (date of access: 20.09.2025).

8. Top 21 Data Analysis Tools for 2025 // OWOX Blog. URL: <https://www.owox.com/blog/articles/data-analysis-tools> (date of access: 02.10.2025).

9. Использование методов корреляционно-регрессионного анализа в бизнесе // StudFiles. URL: <https://studfile.net/preview/717158/page:20/> (date of access: 05.10.2025).

10. IT Data Analytics Software // ManageEngine Analytics Plus URL: <https://www.manageengine.com/analytics-plus/> (date of access: 10.09.2025).

11. Best Product Analytics Tools for Product and Marketing Teams in 2025 // Userpilot Blog. URL: <https://userpilot.com/blog/product-analytics-tools/> (date of access: 12.10.2025).

12. Вишняков, В. А. Комплексный анализ для оценки использования сетей интернета вещей в учебном процессе университета информатики и радиоэлектроники / В. А. Вишняков, Г. А. Хацкевич, Е. И. Полоско // Статистика и экономика. 2025. Т. 22, № 4. С. 52–60. DOI: 10.21686/2500-3925-2025-4-52-51.

13. Вишняков, В.А. Факторный, регрессионный и корреляционный анализы для оценки использования нейронных сетей в учебном процессе университета / В. А. Вишняков, Е. И. Полоско // Системный анализ и прикладная информатика. 2025. № 4. С. 56–63. DOI: 10.21122/2309-4923-2025-4-56-63.

REFERENCES

1. Khatskevich G.A., Rusilko T.V. *Ekonometrika [Econometrics]*. Minsk: RIVSh; 2021. 452 p. (in Russian).

2. Aplyaminov R.F. *Primeneniye korrelyatsionnogo i regressionnogo analiza v sisteme prognozirovaniya tekhnologicheskogo protsesta [Application of correlation and regression analysis in the process forecasting system]*. *Nauchnye vyskazyvaniya*. 2023;21(45):24–28 (in Russian). Available at: https://nvjournal.ru/article/Primenenie_korreljatsionnogo_i_regressionnogo_analiza_v_sisteme_prognozirovaniya_tehnologicheskogo_protsesta (accessed 09 September 2025).

3. *Prikladnaya statistika. Osnovy ekonometriki [Applied Statistics. Fundamentals of Econometrics]*: In 2 vol. 2nd ed. Vol. 1: Ayvazyan S.A., Mkhitaryan V.S. Probability Theory and Applied Statistics. Moscow: UNITY-DANA; 2001. 656 p. (in Russian).

4. The Top 10 Data Analysis Tools for Perfect Data Management. Software Testing Help. Available at: <https://www.softwaretestinghelp.com/data-analysis-tools%20/> (accessed 09 September 2025).

5. 30 Big Data Tools to Improve Your Analysis. Octoparse Blog. Available at: <https://www.octoparse.com/blog/top-30-big-data-tools-for-data-analysis-in-2021> (accessed 09 September 2025).

6. Statistical Consulting. UCLA. Available at: <https://stats.oarc.ucla.edu/> (accessed 15 September 2025).

7. Top 15 Automation Tools for Data Analytics. GeeksforGeeks. Available at: <https://www.geeksforgeeks.org/data-science/top-15-automation-tools-for-data-analytics/> (accessed 20 September 2025).

8. Top 21 Data Analysis Tools for 2025. OWOX Blog. Available at: <https://www.owox.com/blog/articles/data-analysis-tools> (accessed 02 October 2025).

9. Using methods of correlation and regression analysis in business. StudFiles. Available at: <https://studfile.net/preview/717158/page:20/> (accessed 05 October 2025).

10. IT Data Analytics Software. ManageEngine Analytics Plus. Available at: <https://www.manageengine.com/analytics-plus/> (accessed 10 September 2025).

11. Best Product Analytics Tools for Product and Marketing Teams in 2025. Userpilot Blog. Available at: <https://userpilot.com/blog/product-analytics-tools> (accessed 12 October 2025).

12. Vishniakou U.A., Khatskevich G.A., Polosko E.I. Comprehensive analysis to evaluate the use of internet of things networks in the educational process of the university of informatics and radio electronics. *Statistics and Economics*. 2025;22(4):52–60 (in Russian). <https://doi.org/10.21686/2500-3925-2025-4-52-51>.

13. Vishniakou U.A., Polosko E.I. Factorial, regression and correlation analyses to evaluate the use of neural networks in the university educational process. *System analysis and applied information science*. 2025;4:56–63 (in Russian). <https://doi.org/10.21122/2309-4923-2025-4-56-63>.

U. A. VISHNIAKOU, E. I. POLOSKO

SOFTWARE TOOLS FOR AUTOMATING CALCULATIONS FOR FACTORIAL, REGRESSION, AND CORRELATION ANALYSES

Belarusian State University of Informatics and Radioelectronics
Minsk, Republic of Belarus

Abstract. The subject of research is the analysis of the use of calculation automation tools for the use of factorial, regression, and correlation analyses. The purpose of the article is to present the well-known and author's experience of using the Python language for applying factorial, regression and correlation analyses and the experience of the authors in their application. The main aspects of the theoretical foundations of these methods of analysis are considered. The classification of software tools for automation of statistical calculations is given. The advantages and limitations of programming languages (Python, R), commercial platforms (SPSS, SAS, Stata), BI platforms (Tableau, Power BI, Zoho Analytics), open platforms (KNIME, RapidMiner), specialized statistical packages (JASP, Jamovi) are considered. The methodology of the application of automated statistical analysis is detailed, including the preparation of initial data, the choice of method, tincture of parameters, performing calculations and obtaining primary results. The author's methodology for automating complex analysis for evaluating the use of neural networks, the Internet of Things and blockchain in the university's educational process is presented.

Keywords: factorial, regression, correlation, analysis, automation, languages, platforms, systems



Вишняков Владимир Анатольевич

Белорусский государственный университет информатики и радиоэлектроники (БГУИР), г. Минск, Республика Беларусь.

Доктор технических наук, профессор; профессор кафедры инфокоммуникационных технологий БГУИР. Область научных интересов: информационное управление и безопасность, электронный бизнес, интеллектуальные системы управления, нейронные сети, интернет вещей, блокчейн. Член двух докторских Советов по защите диссертаций. Автор более 560 научных работ, в том числе 9 монографий (2 на английском языке), 4 учебных пособий с грифами Министерства образования Республики Беларусь, 1 учебного пособия с грифом УМО, 8-томного учебного комплекса «Информационный менеджмент», 215 научных статей, 21 авторского свидетельства и 3 патентов.

Uladzimir Anatolyevich Vishnyakov

Belarusian State University of Informatics and Radioelectronics (BSUIR), Minsk, Republic of Belarus.

Doctor of Technical Sciences, Professor; Professor of the Department of infocommunication technologies of the BSUIR. Research interests: information management and security, electronic business, intelligent control systems, IoT, network, blockchain. Member of 2 doctoral Councils for the defense of dissertations. Author of more than 560 scientific papers, including 9 monographs (2 in English), 5 textbooks, 8-volume educational complex "Information Management", 215 scientific articles, 24 patents.

E-mail: vish@bsuir.by

Полоско Екатерина Ивановна

Белорусский государственный университет информатики и радиоэлектроники (БГУИР), г. Минск, Республика Беларусь.

Старший преподаватель кафедры «Экономическая информатика» БГУИР, аспирантка. Область научных интересов: комплексный статистический анализ, автоматизация его применения.

Ekaterina Polosko

Belarusian State University of Informatics and Radioelectronics (BSUIR), Minsk, Republic of Belarus.

Senior Lecturer of the Department of Economic Informatics of the BSUIR, Postgraduate Student. Research interests: Comprehensive statistical analysis, automation of its application.

E-mail: kafei@bsuir.by