

НЕЧЕТКАЯ МОДЕЛЬ ДЛЯ ОЦЕНКИ ЭФФЕКТИВНОСТИ РАСПРЕДЕЛЕНИЯ ИНФОРМАЦИОННЫХ РЕСУРСОВ УЧЕБНОГО ЦЕНТРА

FUZZY MODEL FOR EVALUATING THE EFFECTIVENESS OF INFORMATION RESOURCE ALLOCATION OF THE TRAINING CENTER

**A. Sagalaeva
O. Romashkova
N. Rudnichenko**

Summary. This article discusses the specifics of evaluating the efficiency of computing resources allocation in the processing of information flows in the organizational infrastructure of the training center based on the application of the theory of soft computing and fuzzy logic. The general concept of the proposed fuzzy model is developed and described, mathematical formalization of the functional system of indicators of efficiency of information resources distribution of the training center, fuzzy subset for the output variable and defuzzification process is given. The conceptual structure of the information resources distribution efficiency evaluation model, the list and composition of model variable values, three-dimensional surfaces for reflecting interdependence of input variables on the output value of the target efficiency evaluation variable were developed, the study of using the developed model was carried out, and further ways of developing the created model were described.

Keywords: fuzzy logic, soft computing, distributed systems and technologies, information resources.

Сагалаева Анна Игоревна

Аспирант, ГАОУ ВО «Московский городской педагогический университет (МГПУ)» г. Москва, Россия
omegaanya@gmail.com

Ромашкова Оксана Николаевна

Д.т.н., профессор, Российская академия народного хозяйства и государственной службы при Президенте РФ (РАНХиГС)»
Г. Москва

ox-rom@yandex.ru

Рудниченко Николай Дмитриевич

К.т.н., доцент, Одесский национальный политехнический университет, г. Одесса
nikolay.rud@gmail.com

Аннотация. В данной статье рассмотрена специфика оценки эффективности распределения вычислительных ресурсов при обработке информационных потоков в организационной инфраструктуре учебного центра на основе применения теории мягких вычислений и нечеткой логики. Приведено обоснование актуальности предметной тематики исследования, разработана и описана общая концепция предлагаемой нечеткой модели, дана математическая формализация функционала системы показателей эффективности распределения информационных ресурсов учебного центра и процесса дефаззификации. Разработаны концептуальная структура модели оценки эффективности распределения информационных ресурсов, перечень и состав значений переменных модели, построены трехмерные поверхности для отражения взаимозависимости входных переменных на выходное значение целевой переменной оценки эффективности, проведено исследование использования разработанной модели, описаны дальнейшие пути развития созданной модели.

Ключевые слова: нечеткая логика, мягкие вычисления, оценка эффективности, распределенные системы и технологии, информационные ресурсы.

Введение

Все большую актуальность приобретает задача управления процессами обработки и анализа данных как на уровне конфигурации физическо-аппаратного оборудования, так и с использованием средств и технологий виртуализации [1]. Повышение требований к уровню быстродействия работы интел-

лектуальных информационных систем при предоставлении образовательного контента конечным пользователям, осуществляющим большое число запросов на его получение, способствует росту вычислительной нагрузки на серверную инфраструктуру, что требует выработки рациональных подходов к распределению ресурсов и обеспечению параллелизма обработки данных [2,3].

Следует отметить, что в контексте анализа рассматриваемой проблематики, процессы динамической обработки больших объемов информационных ресурсов (ИР) различной природы и структуры, формируемых в процессе функционирования современных учебных центров (УЦ), являются плохо прогнозируемыми в силу стохастической природы внешних воздействий как на управляющую систему, так и на вычислительные ресурсы, аппаратное и программное обеспечение ее работы [4,5].

Среди существующих подходов к повышению уровня определенности распределения вычислительных ресурсов в сложных, многозвенных корпоративных образовательных инфраструктурах все чаще преобладают методики на основе использования машинного обучения и вычислительного интеллекта, позволяющие обеспечить автоматизацию ряда задач при принятии управляющих решений [2,6]. Однако, не все из них могут быть гибко адаптированы для рассматриваемой специфики ИР УЦ в силу отсутствия достаточных наборов статистических данных, экспертных мнений и оценок, сложности формализации всех бизнес-процессов, связанных с обработкой данных на разных этапах их использования. Высокую значимость приобретает оценивание степени влияния значимых параметров системы на итоговые значения выходного признака, позволяющего охарактеризовать эффективность использования вычислительных ресурсов аппаратной инфраструктуры.

В связи со сложностью формализации и функциональной оценки степени эффективности, а также — качества распределения и использования вычислительных ресурсов для управления ИР УЦ, предлагаемую модель целесообразно основывать с использованием аппарата нечеткой логики и мягких вычислений [7,8]. Данный подход оправдан благодаря возможностям снижения уровня неопределенности в процессах оценивания путем внедрения лингвистических шкал и формирования логических продукционных правил, способных учесть характер динамики изменений в инфраструктуре вычислительных кластеров УЦ. Это позволяет выполнить комплексное моделирование и оценку эффективности использования вычислительных ресурсов при различных сценариях нагрузки с учетом объемов обрабатываемых данных [9–11].

Цель статьи

Целью данной статьи является разработка и исследование нечеткой модели для оценки эффективности распределения ИР УЦ на основе применения теории мягких вычислений. Практическая значимость исследования заключается в формализации уровня неопре-

деленности управления процессами анализа объемов аппаратных и виртуальных ресурсов, необходимых для обеспечения деятельности УЦ.

Описание концепции модели

С целью построения модели для оценки эффективности распределения ИР в УЦ сформирована система входных и выходных показателей, которые определяют состав модели. К ним относятся:

1. Надежность используемого в вычислительной инфраструктуре аппаратного обеспечения (H).
2. Надежность внедряемого в рамках отдельных вычислительных узлов программного обеспечения (S).
3. Тип виртуализации, обеспечивающий функционирование рабочего окружения и процессов управления ИР (I).
4. Стоимость поддержки и обслуживания инфраструктуры УЦ (P).

Интегральным выходным показателем (целевой переменной) является эффективность распределения ИР УЦ (eff).

Каждый из интегральных входных показателей формируется на базе набора критериев, характеризующих в совокупности соответствующее значение показателя.

На значение H оказываются влияния такие критерии, как:

- ♦ *уровень износа* H_w , характеризуемый лингвистическими терминами «Высокий», «Средний», «Низкий», «Критический»;
- ♦ *пропускная способность* H_T , характеризуемый лингвистическими терминами «Высокая», «Средняя», «Низкая», «Критическая»;
- ♦ *быстродействие* H_p , характеризуемый лингвистическими терминами «Высокое», «Среднее», «Низкое», «Критическое».

На значение S оказываются влияния такие критерии, как:

- ♦ *архитектура приложения* S_A , характеризуемый лингвистическими терминами «Клиент-серверная», «Монолитная», «Смешанная»;
- ♦ *кроссплатформенность* S_C , характеризуемый лингвистическими терминами «Полная», «Частичная», «Отсутствует»;
- ♦ *степень требований к вычислительным ресурсам* S_H , характеризуемый лингвистическими терминами «Высокая», «Средняя», «Низкая»;
- ♦ *поддержка многопоточности* S_M , характеризуемый лингвистическими терминами «Присутствует», «Отсутствует»;

- ♦ степень обособленности S_S , характеризуемый лингвистическими терминами «Высокая», «Средняя», «Низкая».

На значение V оказываются влияния такие критерии, как:

- ♦ тип инфраструктуры V_I , характеризуемый лингвистическими терминами «Гиперконвергентная», «Конвергентная»;
- ♦ масштабируемость V_M , характеризуемый лингвистическими терминами «Высокая», «Низкая»;
- ♦ прозрачность вычислительных процессов V_T , характеризуемый лингвистическими терминами «Высокая», «Низкая»;
- ♦ балансировка ресурсов V_B , характеризуемый лингвистическими терминами «Автоматическая», «Полуавтоматическая», «Ручная».

На значение P оказываются влияния такие критерии, как:

- ♦ удельная стоимость профилактических работ за установленный период P_P , характеризуемый лингвистическими терминами «Высокая», «Средняя», «Низкая»;
- ♦ удельная стоимость замены оборудования P_C , характеризуемый лингвистическими терминами «Высокая», «Средняя», «Низкая»;
- ♦ удельная стоимость резервирования данных P_R , характеризуемый лингвистическими терминами «Высокая», «Средняя», «Низкая»;
- ♦ удельная стоимость настройки рабочего окружения P_S , характеризуемый лингвистическими терминами «Высокая», «Средняя», «Низкая».

Таким образом общий функционал системы показателей эффективности распределения ИР в УЦ может быть определен в следующем виде

$$eff = [H, S, V, P] = [\{H_W, H_P, W_V\}, \{S_A, S_C, S_M, S_S\}, \{V_I, V_M, V_T, V_B\}, \{P_P, P_C, P_R, P_S\}] \quad (1)$$

В полученном алгоритме присутствуют следующие шаги:

- ♦ установка правил программирования в виде продукционных правил (формата «ЕСЛИ,... ТО»), задающих зависимость уровней входных данных и риска на выходе системы;
- ♦ задание функции принадлежности входных переменных и получение первичного результата их оценок;
- ♦ фаззификация оценок входных переменных — определение конкретных значений заданной функций принадлежности;

- ♦ агрегирование — проверка истинности условий путем ряда преобразований функций принадлежности;
- ♦ активизация заключений — нахождение весовых коэффициентов для каждого из правил и функций истинности;
- ♦ аккумуляция заключений — нахождение функции принадлежности для выходных переменных;
- ♦ дефаззификация — нахождение четких значений каждой выходной переменной.

Разработка нечеткой модели

Предлагаемая модель выполнена в системе Matlab с использованием модуля Fuzzy Logic Toolbox. Структура модели оценки эффективности распределения ИР УЦ на базе нечеткой логики приведена на рисунке 1.

При помощи использования операции МАКСИМУМ (max) было получено итоговое нечеткое подмножество для выходной переменной с функцией принадлежности вида

$$\mu_{\Sigma}(y_{eff}) = \mu_{x_i}(y_{eff}) = \max_{y_{eff}} [\mu_{x_{H_i}}(y_{eff}), \mu_{x_{S_i}}(y_{eff}), \mu_{x_{V_i}}(y_{eff}), \mu_{x_{P_i}}(y_{eff})], \quad (2)$$

где $\mu_{\Sigma}(y_{eff})$ —

итоговое нечеткое подмножество для выходной переменной эффективности распределения ИР УЦ (y_{eff});

$\mu_{x_i}(y_{eff})$ —

нечеткие множества, входящие в состав подмножества $\mu_{\Sigma}(y_{eff})$;

$\mu_{x_{H_i}}(y_{eff})$ —

множество надежности аппаратного обеспечения, используемого в вычислительной инфраструктуре;

$\mu_{x_{S_i}}(y_{eff})$ —

нечеткое множество надежности внедряемого в рамках отдельных вычислительных узлов программного обеспечения;

$\mu_{x_{V_i}}(y_{eff})$ —

Таблица 1. Функции принадлежности лингвистических переменных

	Наименование переменной	Логика переменной	Область определения функции принадлежности
Интегральные показатели модели (Входные)	Надежность используемого в вычислительной инфраструктуре аппаратного обеспечения (H)	Критическая	$Y, 0, 0.125, 0.25$
		Низкая	$Y, 0.125, 0.265, 0.38$
		Средняя	$Y, 0.25, 0.475, 0.7$
		Высокая	$Y, 0.65, 0.825, 1$
	Надежность внедряемого в рамках отдельных вычислительных узлов программного обеспечения (S)	Критическая	$Y, 0, 0.1, 0.2$
		Низкая	$Y, 0.15, 0.265, 0.38$
		Средняя	$Y, 0.35, 0.475, 0.6$
		Высокая	$Y, 0.5, 0.75, 1$
	Тип виртуализации (V)	Программная	$Y, 0, 0.125, 0.25$
		Аппаратная	$Y, 0.125, 0.265, 0.38$
		Контейнерная	$Y, 0.3, 0.525, 0.75$
		Смешанная	$Y, 0.7, 0.85, 1$
Стоимость поддержки обслуживания инфраструктуры УЦ (P)	Низкая	$Y, 0, 0.125, 0.25$	
	Средняя	$Y, 0.125, 0.265, 0.38$	
	Высокая	$Y, 0.25, 0.45, 0.65$	
	Критическая	$Y, 0.62, 0.81, 1$	
Выходная целевая переменная	Эффективность распределения ИР УЦ (eff)	Низкая	$Y, 0, 0.25$
		Средняя	$Y, 0.125, 0.38$
		Достаточная	$Y, 0.25, 0.75$
		Высокая	$Y, 0.65, 1$

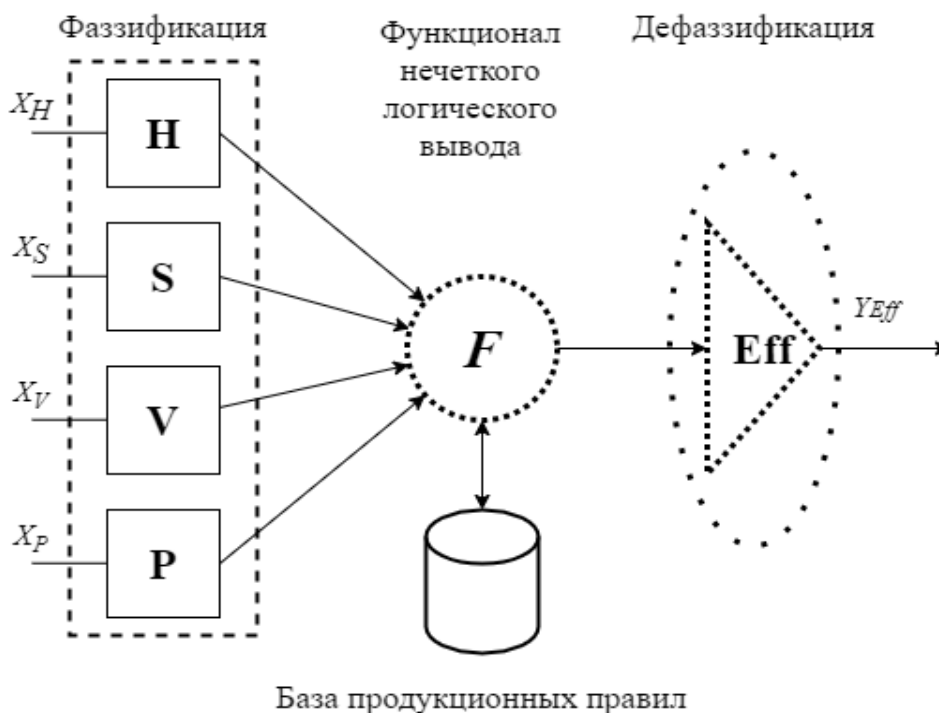


Рис. 1. Структура модели для оценки эффективности распределения ИР УЦ

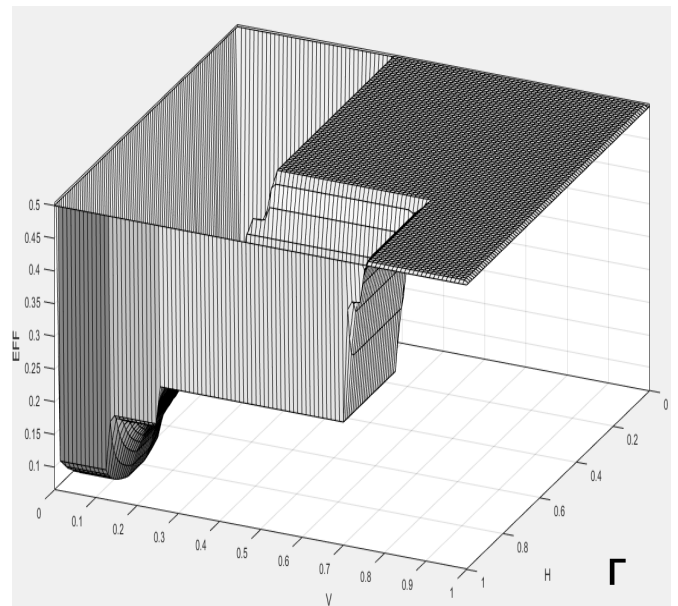
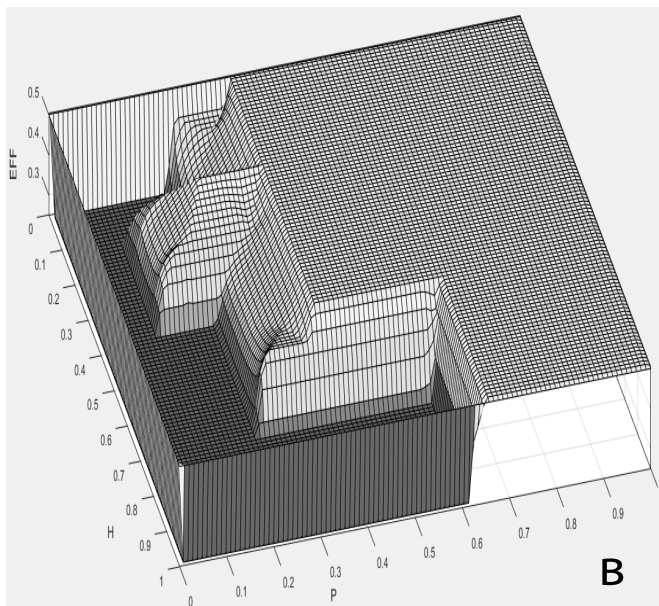
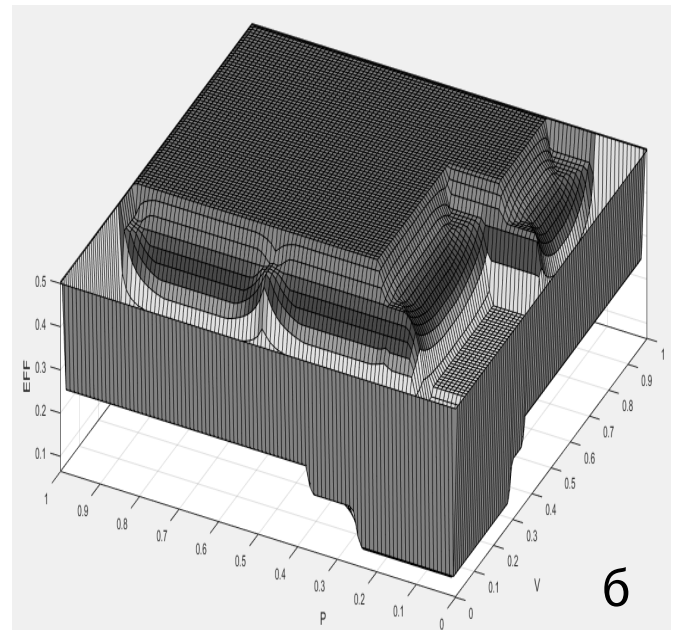
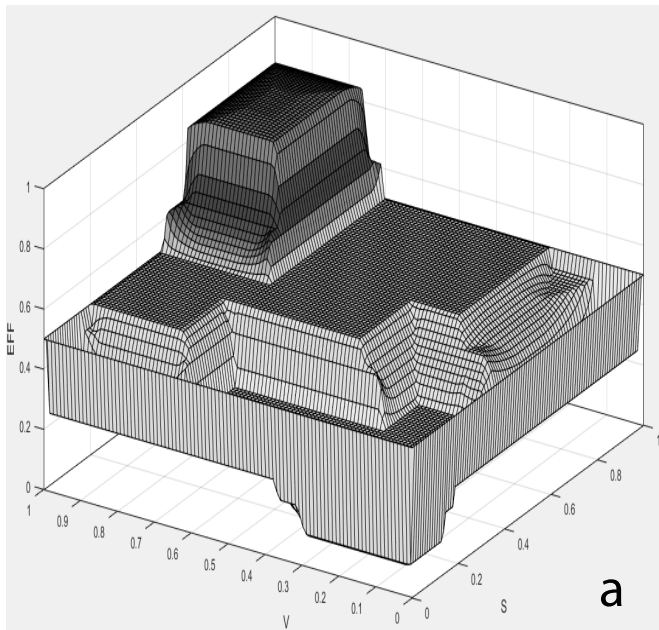


Рис. 2. Трехмерные поверхности, отражающие взаимозависимость и влияние сочетаний входных переменных V и S (а), P и V (б), H и P (в), V и H (г) на выходное значение целевой переменной eff

нечеткое множество типов виртуализации;

$$\mu_{x_P} (y_{eff}) —$$

нечеткое множество стоимости поддержки и обслуживания инфраструктуры УЦ.

Лингвистические переменные и их функции принадлежности в структурной форме приведены в таблице 1.

Операция дефаззификации выполнена в Matlab при помощи метода нахождения центра тяжести плоской фигуры, ограниченной осями координат и графиком функции принадлежности нечеткого множества (метода центра тяжести centroid), согласно

$$\overline{y_{eff}} = \frac{\int_{\min}^{\max} y_{eff} \cdot \mu(y_{eff}) dy_{eff}}{\int_{\min}^{\max} \mu(y_{eff}) dy_{eff}}, \quad (3)$$

где $\overline{y_{eff}}$ — результат дефаззификации; y_{eff} — переменная, заданная для выходной лингвистической переменной y_{eff} ; $\mu(y_{eff})$ — функция принадлежности нечеткого множества для выходной переменной y_{eff} после этапа аккумуляции; \min и \max — левая и правая точки интервала носителя нечеткого множества рассматриваемой выходной переменной y_{eff} .

В ходе моделирования получены трехмерные визуализации поверхностей модели для оценки эффективности распределения ИР в УЦ для разных сочетаний входных переменных (рисунок 2), использованные для

оценивания и формализации представления специфики влияния H, S, V, P на eff .

Разработанная модель позволяет получить значение eff для каждой точки, принадлежащей трехмерной поверхности, а также отражает качественные переходы между значениями параметров в виде «впадин» и «всплесков».

Модель комбинирует и преобразует функции принадлежности всех входных переменных, что позволяет отразить специфику взаимосвязи H, S, V, P между собой. Анализ полученных трехмерных визуализаций созданных поверхностей модели для оценки эффективности распределения ИР в УЦ наибольший рост уровня eff наблюдается при значениях V в диапазоне от 0,455 до 1.

Заключение

Разработанная нечеткая модель для оценки эффективности распределения ИР УЦ на основе применения теории мягких вычислений позволяет выявить взаимосвязь между сформированным набором входных параметров и формализовать их влияние на значение выходной переменной, что может быть использовано в качестве основы интеллектуального модуля платформы поддержки принятия управленческих решений в ИР УЦ.

Возможным направлением дальнейшего развития предложенной модели является выполнение исследований и вычислительных экспериментов, направленных на оптимальный подбор функций активации, а также гибридизация модели на основе применения методов машинного обучения и искусственного интеллекта, в частности, — аппарата искусственных нейронных сетей.

ЛИТЕРАТУРА

1. Биктимиров М.Р. Тенденции развития технологий обработки больших данных и инструментария хранения разноформатных данных и аналитики / М.Р. Биктимиров, А.М. Елизаров, А.Ю. Щербаков // Электронные библиотеки. — № 5. — Т. 19. — 2016. — С.390–406.
2. Сагалаев Ю.Р. Разработка проекта системы поддержки принятия решений для провайдера сетевых услуг на базе вычислительных кластеров / Ю.Р. Сагалаев, А.И. Сагалаева, О.Н. Ромашкова // Современная наука: актуальные проблемы теории и практики. Серия: Естественные и Технические Науки. — 2021. — № 10. — С. 109–115.
3. Шibaев Д.С. Оптимизация методов прогнозирования, обработки и анализа информации в разноструктурных хранилищах данных / Д.С. Шibaев, В.В. Вычужанин, Н.О. Шibaева, Н.Д. Рудниченко // Информатика и математические методы в моделировании. — 2018. — № 1. — С. 78–85.
4. Орлова, А.И. Применение средств виртуализации как один из критериев повышения эффективности использования ресурсов учебного центра / А.И. Орлова // Технологии информационного общества: Материалы XIII Международной отраслевой научно-технической конференции, Москва, 20–21 марта 2019 года. — Москва: ООО «Издательский дом Медиа паблишер», 2019. — С. 81–83.
5. Орлова, А.И. Возможности использования хранилища данных и платформы виртуализации в учебном центре на примере Virtuozzo infrastructure platform / А.И. Орлова // #ScienceJuice2019: Сборник статей и тезисов студенческой открытой конференции, Москва, 18 ноября — 09 2019 года. — Москва: Издательство ПАРАДИГМА, 2020. — С. 392–393.

6. Пырнова О.А. Технологии искусственного интеллекта в образовании / О.А. Пырнова, Р.С. Зарипова // Russian Journal of Education and Psychology.— 2019.— № 3.— С. 41–44.
7. Воронова М.В. Применение нечеткой логики, как средство оптимизации снабжения / М.В. Воронова, М.В. Токмачева // Актуальные проблемы авиации и космонавтики.— 2017.— № 13.— С. 596–598.
8. Вычужанин В.В. Построение модели оценок технического состояния систем на транспорте / В.В. Вычужанин, Н.Д. Рудниченко, А.В. Вычужанин, А.Е. Козлов // Актуальные проблемы автотранспортного комплекса: межвуз. сб. науч. статей (с междунар. участием).— Самара.— 2019.— С. 121–126.
9. Борисов В.В. Нечеткие модели и сети / В.В. Борисов, В.В. Круглов, А.С. Федулов.— М.: Горячая линия–Телеком, 2012.— 284 с.
10. Ермакова Т.Н. Модернизированная структура управления образовательной системой / Т.Н. Ермакова, О.Н. Ромашкова, Л.А. Пономарева // Вестник Брянского государственного технического университета.— 2019.— № 6 (79).— С. 84–91.
11. Ромашкова О.Н. Линейное ранжирование показателей оценки деятельности вуза / О.Н. Ромашкова, Л.А. Пономарева, И.П. Василюк // Современные информационные технологии и ИТ-образование.— 2018.— Т. 14.— № 1.— С. 245–255.

© Сагалаева Анна Игоревна (omegaanya@gmail.com),

Ромашкова Оксана Николаевна (ox-rom@yandex.ru), Рудниченко Николай Дмитриевич (nickolay.rud@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»



Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации