

ВЕСТНИК НОВОСИБИРСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

Научный журнал
Основан в ноябре 1999 года

Серия: Информационные технологии

2024. Том 22, № 2

СОДЕРЖАНИЕ

<i>Дымонт Н. А., Малаева Е. Д., Яхьяева Г. Э.</i> Алгоритмы проверки корректности интервальных экспертных оценок	5
<i>Евсиков И. А.</i> Информационные технологии для расчета антропогенного потока тепла в условиях городской застройки	20
<i>Мазайшвили Е. К., Авксентьева Е. Ю.</i> Метод цифровой подписи изображений без дополнительных файлов, с устойчивостью к JPEG-сжатию и удалению метаданных	33
<i>Пылов П. А., Майтак Р. В., Копылов Д. Е.</i> Разработка метода для оптимизации процесса получения значений числовых характеристик компонентов нефтяного сырья из графических данных хроматограмм.....	44
<i>Тимофеев В. С.</i> Использование цепей Маркова для автоматического завершения исходного кода программы	57
<i>Щелоков А. Е., Будников К. И.</i> Исследование архитектур нейронных сетей для определения концентраций газов по спектрам	68
Информация для авторов	79

V E S T N I K

NOVOSIBIRSK STATE UNIVERSITY

Scientific Journal
Since 1999, November
In Russian

Series: Information Technologies

2024. Volume 22, № 2

CONTENTS

<i>Dymont N. A., Malaeva E. D., Yakhyaeva G. E.</i> Algorithms for Verifying the Correctness of Interval Expert Evaluations	5
<i>Evsikov I. A.</i> Information Technologies for Calculating Anthropogenic Heat Flux in Urban Areas	20
<i>Mazaishvili E. K., Avksentieva E. Yu.</i> Digital Signature Method for Images without Additional Files, Robust to JPEG Compression and Metadata Removal	33
<i>Pylov P. A., Maitak R. V., Kopylov D. E.</i> Development of a Method for Optimization of the Process of Extraction of Values of Numerical Characteristics of Oil Feedstock Components from Graphical Data of Chromatographic Analysis	44
<i>Timofeev V. S.</i> Code Completion Using Markov Chains.....	57
<i>Shchelokov A. E., Budnikov K. I.</i> Study of Neural Network Architectures for Determining Gas Concentrations by Spectra.....	68
Instructions for Contributors.....	79

Editor in Chief M. M. Lavrentiev

Vice-Editor A. V. Avdeev

Executive Secretary D. P. Iksanova

Editorial Board of the Series

- I. V. Bychkov*, professor, academician (Irkutsk), *B. M. Glinsky*, professor (Novosibirsk)
A. N. Gorban, professor (Lester, GB), *E. P. Gordov*, professor (Tomsk)
B. S. Dobronets, professor (Krasnoyarsk), *A. M. Elizarov*, professor (Kazan)
G. N. Erokhin, professor (Kaliningrad), *A. I. Kamyshnikov*, professor (Khanty-Mansiysk)
G. P. Karev, professor (Maryland, USA), *N. A. Kolchanov*, professor, academician (Novosibirsk)
M. M. Lavrentjev, professor (Novosibirsk), *V. E. Malyshkin*, professor (Novosibirsk)
N. N. Mirenikov, professor (Aizu, Japan), *N. M. Oskorbin*, professor (Barnaul)
D. E. Palchunov, professor (Novosibirsk), *T. Pizansky*, professor (Ljubljana, Slovenia)
V. P. Potapov, professor (Kemerovo), *O. I. Potaturkin*, professor (Novosibirsk)
V. A. Serebryakov, professor (Moscow), *A. V. Starchenko*, professor (Tomsk)
S. I. Smagin, professor, corresponding member of RAS (Khabarovsk)
D. A. Tusupov, professor (Astana, Kazakhstan)
V. V. Shajdurov, professor, corresponding member of RAS (Krasnoyarsk)
Yu. I. Shokin, professor, academician (Novosibirsk)

*The journal is published quarterly in Russian since 1999
by Novosibirsk State University Press*

*The address for correspondence
Faculty of Information Technologies, Novosibirsk State University
1 Pirogov Street, Novosibirsk, 630090, Russia
Tel. +7 (383) 363 42 46*

E-mail address: inftech@vestnik.nsu.ru

On-line version: <http://elibrary.ru>

Научная статья

УДК 004.827

DOI 10.25205/1818-7900-2024-22-2-5-19

Алгоритмы проверки корректности интервальных экспертных оценок

Надежда Андреевна Дымонт¹, Елена Дмитриевна Малаева²
Гульнара Эркиновна Яхьяева³

Новосибирский государственный университет
Новосибирск, Россия

¹n.dymont@g.nsu.ru

²e.malaeva@g.nsu.ru

³g.iakhiaeva@g.nsu.ru

Аннотация

Экспертная вероятность используется для принятия решений, позволяет оценить риски, особенно полезна в условиях ограниченности и недоступности получения объективных данных. При этом возможны противоречия между оценками, что может затруднить анализ и принятие решений.

В статье описан алгоритм, который проверяет соответствие экспертных оценок с использованием вероятностных интервалов. Один из разделов посвящен разработке метода оценки новой формулы, основанной на исходной системе оценок. Описан пользовательский интерфейс, обеспечивающий взаимодействие с разработанными алгоритмами.

Ключевые слова

экспертные оценки, субъективная вероятность, согласованное означивание, нечеткая модель

Для цитирования

Дымонт Н. А., Малаева Е. Д., Яхьяева Г. Э. Алгоритмы проверки корректности интервальных экспертных оценок // Вестник НГУ. Серия: Информационные технологии. 2024. Т. 22, № 2. С. 5–19. DOI 10.25205/1818-7900-2024-22-2-5-19

Algorithms for Verifying the Correctness of Interval Expert Evaluations

Nadezhda A. Dymont¹, Elena D. Malaeva², Gulnara E. Yakhyaeva³

Novosibirsk State University,
Novosibirsk, Russia

¹n.dymont@g.nsu.ru

²e.malaeva@g.nsu.ru

³g.iakhiaeva@g.nsu.ru

Abstract

Expert probability is used for decision-making, allows you to assess risks, and is especially useful in conditions of limited and inaccessible objective data. At the same time, there may be contradictions between the estimates, which may complicate analysis and decision-making. This paper describes an algorithm that checks the compliance of expert estimates using probability intervals. One of the sections is devoted to the development of an evaluation method for a

© Дымонт Н. А., Малаева Е. Д., Яхьяева Г. Э., 2024

new formula based on the original evaluation system. The article describes the user interface that provides interaction with the developed algorithms.

Keywords

expert evaluations, subjective probability, coordinated evaluations, fuzzy model

For citation

Dymont N. A., Malaeva E. D., Yakhyaeva G. E. Algorithms for verifying the correctness of interval expert evaluations. *Vestnik NSU. Series: Information Technologies*, 2024, vol. 22, no. 2, pp. 5–19 (in Russ.) DOI 10.25205/1818-7900-2024-22-2-5-19

Введение

В наше время субъективная вероятность играет важную роль в принятии решений в различных областях, от финансового анализа до управления проектами. Этот подход позволяет оценивать риски, успешность завершения проектов и предвидеть возможные сценарии, особенно когда объективные данные недоступны или их неполнота делает использование объективных оценок нецелесообразным из-за высоких затрат на их получение.

Концепция субъективной вероятности, впервые введенная Фрэнком Рамсеем, базируется на уверенности эксперта в наступлении конкретного события [1]. Этот подход выражает личное мнение и может быть основан на опыте, интуиции или знаниях специалиста в определенной области. Однако важно учитывать, что субъективная вероятность также может быть подвержена влиянию различных факторов, таких как эмоции или предвзятость.

Чтобы создать более надежную модель, часто привлекают мнение нескольких экспертов [2]. При этом возможны противоречия между их оценками, что может затруднить анализ и принятие решений. Для обеспечения согласованности экспертных знаний используются различные инструменты, включая методы агрегирования мнений и анализ согласованности предсказаний. Одним из способов проверки согласованности является проведение экспертных сессий, где участники могут обсудить свои оценки и аргументировать точки зрения [3]. Правильное применение методов оценки согласованности может значительно повысить эффективность принятия решений в условиях неопределенности, обеспечивая более точные прогнозы и уменьшая риски.

Настоящее исследование базируется на алгоритме проверки корректности оценочных знаний, описанном в работе [4], где была предложена теоретико-модальная формализация субъективной и объективной интерпретаций вероятности. Также в контексте данной темы ранее было проведено исследование [5, 6], которое сосредоточилось на визуализации исходных данных в виде дерева и проверки согласованности высказываний с его помощью. В процессе исполнения программы строится дерево, в котором конъюнкты из СДНФ-формул на входе являются его листьями. При помощи дизъюнкции между собой они объединяются в исходные формулы, при этом в корне дерева возникает особенная вершина – дизъюнкция всех конъюнкций с вероятностью 1. Ребра в этой структуре являются отношением частичного порядка между вершинами дерева. У каждой вершины есть свое оценочное значение, и стоит отметить, что конъюнктам из СДНФ-формул по умолчанию присваивается значение в виде интервала $[0;1]$. Далее эти значения могут корректироваться в зависимости от введенных пользователем значений, что приводит к последующему сужению интервала.

Подобный подход не позволяет расширять знания о предметной области, лишь проверять их согласованность. Но также важно уметь оценивать и новые высказывания в имеющемся контексте, исследуя его новые стороны и углубляя свое понимание. Прогнозирование играет ключевую роль в развитии научного и практического прогресса во многих областях. Новые высказывания могут предложить и новые способы интерпретации данных или решения проблем, которые ранее были недоступны. Их оценка позволяет расширить возможности в области при-

нятия решений и анализа данных, она способствует развитию науки и углублению понимания проблем и явлений в различных областях знаний.

Одной из причин, почему эксперты могут испытывать затруднения в оценке новых формул, является их сложность и техническая специфика [7]. Новые подходы могут основываться на совершенно иных принципах или использовать новые технологии, с которыми эксперты могут быть не знакомы. Это может приводить к тому, что эксперты не имеют достаточного опыта или знаний для адекватной оценки. Важно также отметить, что даже если эксперт имеет достаточные знания в предметной области, он может столкнуться с трудностями в оценке новых формул из-за их абстрактности. Некоторые новые подходы могут быть труднопонимаемыми или требовать специальных знаний в области математики, статистики или информатики, что делает их оценку непростой задачей.

В существующих решениях также невозможна оценка события несколькими значениями, что представляет собой значительное ограничение. В большинстве случаев системы оценки предполагают использование одного численного значения вероятности для каждого события. Однако в реальной жизни разные эксперты могут иметь разные оценки вероятности для одного и того же события, основываясь на своем опыте, знаниях и восприятии. Когда такие различия все же возникают, системы, в которых одно и то же выражение имеет несколько несовпадающих вероятностных оценок, являются заведомо несовместными. Это создает серьезные препятствия для принятия эффективных решений и планирования на основе таких данных.

Одним из возможных способов преодоления этой проблемы является принятие вероятности события за интервал, который включает все оценки, полученные от различных экспертов. Такой подход позволяет учесть разнообразие мнений и взглядов и обеспечить более полное представление о вероятности наступления события. После того как вероятность события представлена интервалом, применяются методы интервального анализа для проверки согласованности системы высказываний. Такая проверка позволяет выявить возможные противоречия в оценках и принять решения на основе наиболее достоверных данных.

В данной статье представлен процесс разработки алгоритма, который направлен на проверку согласованности экспертных оценок с использованием вероятностных интервалов. В контексте исследования особое внимание уделяется разработке метода оценки новой формулы, основанной на изначальной системе оценок. Для обеспечения удобства взаимодействия с разработанными алгоритмами представлен пользовательский интерфейс. Этот интерфейс предоставляет пользователям интуитивно понятные инструменты для ввода экспертных оценок, проведения проверки согласованности оценок и анализа результатов. В заключении статьи сформулированы основные выводы и результаты работы. В целом, статья подчеркивает важность разработки и применения подобных инструментов в контексте принятия решений и анализа данных в условиях неопределенности.

1. Математические основы

При анализе сложных систем в условиях неопределенности часто применяются методы теории вероятности и математической статистики. Эти методы основаны на вероятностной интерпретации данных и выводов. Сегодня возрастает потребность в новых подходах к математическому описанию информации с высоким уровнем неопределенности. В данной работе используется теоретико-модельная формализация субъективной вероятности через аппарат нечетких моделей, описанная в работах [8–10].

Определение 1. Тройку $\mathfrak{A}_\mu = \langle A, \sigma, \mu \rangle$ будем называть **нечеткой моделью**, если истинностная функция $\mu: S(\sigma_A) \rightarrow [0, 1]$ является нечеткой мерой, определенной на алгебре Линденбаума – Тарского $S(\sigma_A) / \sim$, на которой выполняются следующие условия.

(A1) Если предложение $\varphi \in S(\sigma_A)$ тождественно истинно (тождественно ложно), то $\mu(\varphi) = 1$ ($\mu(\varphi) = 0$).

(A2) Для любой счетной последовательности предложений $\{\varphi_i \in S(\sigma_A) \mid i \in \mathbb{N}\}$ такой, что $\mu(\varphi_i \& \varphi_j) = 0$ для любых $i, j \in \mathbb{N}$, выполняется условие

$$\mu\left(\bigvee_{i \in \mathbb{N}} \varphi_i\right) = \sum_{i \in \mathbb{N}} \mu(\varphi_i).$$

(A3) Для любых предложений $\varphi, \psi \in S(\sigma_A)$ выполняется условие

$$\varphi\psi \Rightarrow \mu(\varphi) = \mu(\psi).$$

(A4) Для любой формулы $\varphi(x) \in F(\sigma)$ с одной свободной переменной выполняются условия

$$\mu(\forall x \varphi(x)) = \mu\left(\bigwedge_{a \in A} \varphi(a)\right), \quad \mu(\exists x \varphi(x)) = \mu\left(\bigvee_{a \in A} \varphi(a)\right).$$

Задавая нечеткую модель $\mathfrak{A}_\mu = \langle A, \sigma, \mu \rangle$ как формализацию предметной области, мы должны располагать знанием об истинностных значениях на модели $\mathfrak{A}_\mu$ всех бескванторных формул сигнатуры. Однако эксперт (или даже группа экспертов) может не обладать таким полным знанием. Следовательно, на практике мы можем столкнуться с частичным означиванием предложений сигнатуры σ . Более того, разные эксперты могут немного расходиться в своих оценках того или иного события, и, как следствие, мы будем вынуждены вместо точечных оценок рассматривать интервальные оценки.

Обозначим через $\mathbb{I}$ множество всех замкнутых подынтервалов интервала $[0; 1]$, т. е.

$$\mathbb{I} = \{[a; b] \mid 0 \leq a \leq b \leq 1\}.$$

Также будем считать, что $\emptyset \in \mathbb{I}$. Это нам позволит формализовать частичные (неполные) знания о предметной области.

Определение 2. Отображение $ev: S(\sigma_A) \rightarrow \mathbb{I} \cup \{\emptyset\}$ назовем **интервальным означиванием** множества предложений сигнатуры σ_A .

Означивание ev назовем **полным**, если для любого предложения $\varphi \in S(\sigma_A)$ имеем $ev(\varphi) \neq \emptyset$. В противном случае означивание ev будем называть **частичным**.

Означивание ev назовем **точечным**, если для любого предложения $\varphi \in S(\sigma_A)$ либо $ev(\varphi) = \emptyset$, либо $ev(\varphi)$ является вырожденным интервалом, т. е. числом из интервала $[0; 1]$.

Определение 3. Интервальное означивание $ev: S(\sigma_A) \rightarrow \mathbb{I}$ **согласуется** с нечеткой моделью $\mathfrak{A}_\mu = \langle A, \sigma, \mu \rangle$, если для любого предложения $\varphi \in S(\sigma_A)$ выполняется условие:

$$ev(\varphi) \neq \emptyset \Rightarrow \mu(\varphi) \in ev(\varphi).$$

Означивание называется **согласованным**, если существует хотя бы одна нечеткая модель, с которой оно согласуется.

Очевидно, что истинностная функция нечеткой модели $\mathfrak{A}_\mu$ является согласованным, полным, точечным означиванием.

Далее встает вопрос о проверке, является ли данное означивание согласованным или нет. В статье [11] была доказана следующая теорема.

Теорема 1. Пусть $ev: S(\sigma_A) \rightarrow \mathbb{I}$ точечное частичное означивание такое, что множество $S = \{\varphi \mid ev(\varphi) \neq \emptyset\}$ является конечным множеством бескванторных предложений. Означивание является согласованным тогда и только тогда, когда система линейных уравнений

$$\begin{cases} \sum_{j=1}^m \varepsilon_{ij} \omega_j = ev(\varphi_i), & i = 1, \dots, n; \\ \sum_{j=1}^m \omega_j = 1, \end{cases} \quad (1.1)$$

совместна и имеет решение при ограничениях $\omega_j \in [0; 1] (j = 1, \dots, m)$.

Поясним обозначения из теоремы 1. Пусть $S = \{\varphi_1, \dots, \varphi_n\}$ – конечное множество бескванторных предложений. Тогда система (1.1) будет содержать $n + 1$ уравнение. Также в системе (1.1) через ω_j обозначается полный конъюнкт сигнатуры σ_A , а сумма $\sum_{j=1}^m \varepsilon_{ij} \omega_j$ является разложением предложения φ_i в дизъюнкцию полных конъюнктов.

Данная теорема будет нами использоваться в параграфе 2.1 для описания алгоритма проверки согласованности интервального частичного означивания.

Определение 4. Означивание $ev_1: S(\sigma_A) \rightarrow \mathbb{I}$ является **расширением** интервального означивания $ev_2: S(\sigma_A) \rightarrow \mathbb{I}$, если для любого предложения $\varphi \in S(\sigma_A)$ выполняется условие:

$$ev_2(\varphi) \subseteq ev_1(\varphi).$$

Означивание ev_2 в этом случае будем называть **сужением** означивания ev_1 .

Очевидно, что если ev – полное согласованное означивание, тогда любое его расширение также согласованное. Однако если является частичным согласованным означиванием, то не всякое его расширение будет согласованным.

Определение 5. Пусть $ev: S(\sigma_A) \rightarrow \mathbb{I}$ – согласованное означивание. Отображение $ken_{ev}: S(\sigma_A) \rightarrow \wp([0; 1])$ будем называть **ядром** означивания ev , если для любого предложения $\varphi \in S(\sigma_A)$ выполняется условие:

$$ken_{ev}(\varphi) = \{\mu(\varphi) \mid \text{модель } \mathfrak{A}_\mu \text{ согласована с означиванием } ev\}.$$

Теорема 2. Ядро любого согласованного означивания является полным согласованным означиванием.

Доказательство. Рассмотрим согласованное означивание $ev: S(\sigma_A) \rightarrow \mathbb{I}$ и его ядро $ken_{ev}: S(\sigma_A) \rightarrow \wp([0; 1])$. Пусть нечеткие модели $\mathfrak{A}_{\mu_1} = \langle A, \sigma, \mu_1 \rangle$ и $\mathfrak{A}_{\mu_2} = \langle A, \sigma, \mu_2 \rangle$ согласуются с означиванием ev . Тогда для любого предложения $\varphi \in S(\sigma_A)$ имеем

$$\mu_1(\varphi), \mu_2(\varphi) \in ken_{ev}(\varphi).$$

Допустим, что $\mu_1(\varphi) < \mu_2(\varphi)$. Нам необходимо показать, что для любого числа α такого, что $\mu_1(\varphi) < \alpha < \mu_2(\varphi)$, можно построить нечеткую модель $\mathfrak{A}_\mu = \langle A, \sigma, \mu \rangle$, которая, с одной стороны, согласована с означиванием ev , а с другой стороны, $\mu(\varphi) = \alpha$.

Пусть

$$k_1 = \frac{\alpha - \mu_1(\varphi)}{\mu_2(\varphi) - \mu_1(\varphi)} \text{ и } k_2 = \frac{\mu_2(\varphi) - \alpha}{\mu_2(\varphi) - \mu_1(\varphi)}.$$

Определим отображение

$$\mu = k_2 \cdot \mu_1 + k_2 \cdot \mu_2.$$

Очевидно, что отображение μ является нечеткой мерой, отвечающей всем свойствам из определения 1. Следовательно, оно является истинностной функцией нечеткой модели $\mathfrak{A}_\mu = \langle A, \sigma, \mu \rangle$. Более того, $\mu(\varphi) = \alpha$.

Покажем, что построенная таким образом, нечеткая модель $\mathfrak{A}_\mu$ согласуется с означиванием ev .

Случай, когда $ev(\varphi) = \emptyset$ тривиален. Поэтому рассмотрим случай, когда $ev(\varphi) \neq \emptyset$.

Так как модели $\mathfrak{A}_{\mu_1}$ и $\mathfrak{A}_{\mu_2}$ согласуются с означиванием ev , то $[\mu_1(\varphi), \mu_2(\varphi)] \subseteq ev(\varphi)$. А так как $\mu_1(\varphi) < \alpha < \mu_2(\varphi)$, то $\alpha \in ev(\varphi)$. Следовательно, модель $\mathfrak{A}_\mu$ также согласуется с означиванием ev .

Теорема доказана.

Данная теорема будет нами использоваться в параграфе 2.2 для описания алгоритма предсказания оценки для новой формулы на основе уже имеющихся оценок. Математическая формулировка данного алгоритма следующая: на входе у нас имеется интервальное частичное означивание $ev: S(\sigma_A) \rightarrow \mathbb{I}$ и предложение $\varphi \in S(\sigma_A)$ такое, что $ev(\varphi) = \emptyset$. Требуется найти интервал $ken_{ev}(\varphi)$.

2. Архитектура программной системы

Для выявления противоречий в оценках, адаптации к новой информации и принятия информированных решений в условиях неопределенности была разработана программная система, которая состоит из двух модулей.

1. **Модуль проверки совместности экспертных оценок** предоставляет возможность проверить согласованность множества экспертных оценок относительно событий. Он создает систему уравнений для определения наличия проблем с согласованностью.

2. **Модуль предсказания оценки для новой формулы** на основе уже имеющихся оценок, чтобы не нарушить согласованность исходной системы.

Рассмотрим основные принципы функционирования каждого из указанных модулей.

2.1. Модуль проверки совместности экспертных оценок

Как было сказано выше, интервальная неопределенность означает, что мы имеем лишь общее представление о значении интересующей нас величины, указывая только диапазон возможных значений или пределы ее изменения. Таким образом, ширина этого интервала является мерой нашей неопределенности или неоднозначности в отношении этой величины.

Для проверки согласованности интервальных экспертных оценок мы будем использовать технику интервального анализа [12]. В частности, будем использовать алгоритмы проверки разрешимости систем интервальных линейных уравнений.

Определение 6. Систему уравнений будем называть интервальной системой линейных уравнений (ИСЛУ), если она имеет вид

$$\left\{ \sum_{j=1}^m \overline{a_{ij}} x_j = \overline{b_i}, \quad i = 1, \dots, n, \right. \quad (2.1)$$

где коэффициенты системы $\overline{a_{ij}}$ и свободные члены системы $\overline{b_i}$ являются интервалами действительных чисел.

Определение 7. ИСЛУ вида (2.1) будем называть **сильно разрешимой (слабо разрешимой)**, если для любых действительных чисел (найдутся действительные числа) $a_{ij} \in \overline{a_{ij}}, b_i \in \overline{b_i}$ таких, что следующая система уравнений совместна, т.е. имеет решение:

$$\left\{ \sum_{j=1}^m a_{ij} x_j = b_i, \quad i = 1, \dots, n, \right.$$

Для проверки согласованности оценочных знаний мы будем проверять слабую разрешимость ИСЛУ

$$\left\{ \begin{array}{l} \sum_{j=1}^m \varepsilon_{ij} x_j = [\alpha_i; \beta_i], \quad i = 1, \dots, n, \\ \sum_{j=1}^m x_j = 1, \end{array} \right. \quad (2.2)$$

где коэффициенты $\varepsilon_{ij} \in \{0; 1\}$, т. е. являются вырожденными (точечными) интервалами, $[\alpha_i; \beta_i] \subseteq [0; 1]$. Параметр m равен количеству формул, согласованность которых необходимо проверить.

Для построения ИСЛУ (2.2) (согласно теореме 1) нам необходимо привести все введенные формулы к совершенной дизъюнктивной нормальной форме (СДНФ). Для этого выделяются все отдельные атомарные высказывания из заданного набора формул и упорядочиваются в лексикографическом порядке.

Коэффициент ε_{ij} равен 1, если конъюнкт под номером j содержится в полном СДНФ-представлении предложения под номером i , и 0 в противном случае. Интервал $[\alpha_i; \beta_i]$ является оценочным значением i -го предложения.

Совместность системы проверяется методом максимизации распознающего функционала [13], реализованном в библиотеке *intervalpy*. Исследование разрешимости интервальной линейной системы уравнений проводится путем решения задачи безусловной максимизации распознающего функционала $Uss(x, A, b)$. При максимизации $U = \max_{x \in \mathbb{R}^n} Uss(x, A, b)$, если $U \geq 0$, то точка максимума τ принадлежит множеству решений $\Xi(A, b)$ системы $Ax = b$, указывая на ее разрешимость; если $U > 0$, то принадлежит внутренности множества решений, и разрешимость системы остается устойчивой к небольшим изменениям в A и b ; если же $U < 0$, то множество решений пусто, что указывает на неразрешимость системы $Ax = b$.

Рассмотрим работу алгоритма на примере. Проверим согласованность формул φ_1 , φ_2 , φ_3 с соответствующими интервальными оценками (табл. 1).

Таблица 1

Входные данные для проверки совместности оценок

Table 1

Input data for checking the consistency of estimates

Формула	Оценка
$\varphi_1 = (P(a) \wedge Q(b)) \vee \overline{R(c)}$	$[0,7; 0,8]$
$\varphi_2 = R(c) \wedge \overline{Q(b)}$	$[0,2; 0,3]$
$\varphi_3 = \overline{P(a)} \vee Q(b)$	$[0,8; 0,9]$

Приведем введенные формулы к виду СДНФ:

$$\varphi_1 = (\overline{P(a)} \wedge \overline{Q(b)} \wedge \overline{R(c)}) \vee (\overline{P(a)} \wedge Q(b) \wedge \overline{R(c)}) \vee (P(a) \wedge \overline{Q(b)} \wedge \overline{R(c)}) \vee$$

$$\begin{aligned}
& \vee (P(a) \wedge Q(b) \wedge \overline{R(c)}) \vee (P(a) \wedge Q(b) \wedge R(c)) \\
\varphi_2 = & (\overline{P(a)} \wedge \overline{Q(b)} \wedge R(c)) \vee (P(a) \wedge \overline{Q(b)} \wedge R(c)) \\
\varphi_3 = & (\overline{P(a)} \wedge \overline{Q(b)} \wedge \overline{R(c)}) \vee (\overline{P(a)} \wedge \overline{Q(b)} \wedge R(c)) \vee (\overline{P(a)} \wedge Q(b) \wedge \overline{R(c)}) \\
& \vee (\overline{P(a)} \wedge Q(b) \wedge R(c)) \vee (P(a) \wedge \overline{Q(b)} \wedge \overline{R(c)}) \vee (P(a) \wedge \overline{Q(b)} \wedge R(c))
\end{aligned}$$

Этим данным соответствует следующая система линейных интервальных уравнений:

$$\begin{cases}
x_1 + x_3 + x_5 + x_7 + x_8 = [0, 7; 0, 8] \\
x_2 + x_6 = [0, 2; 0, 3] \\
x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = [0, 8; 0, 9] \\
x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1 \\
x_i \in [0, 1], \quad i = 1, \dots, 8
\end{cases}$$

Эта система является совместной, а значит и входные данные непротиворечивы.

2.2. Модуль предсказания оценки для новой формулы на основе уже имеющихся оценок

Экспертная оценка в предметной области может быть ограничена огромным разнообразием факторов, воздействующих на конкретные события. Из-за этой сложности и объема информации эксперты могут оценивать только ограниченное число возможных событий. Разработанный нами алгоритм предсказания оценки для новой формулы на основе уже имеющихся оценок становится ключевым инструментом в этом контексте. Он позволяет адаптироваться к новой информации, предоставляя метод оценки новых событий на основе имеющихся экспертных оценок и данных. Такой инструмент позволяет компенсировать ограничения в экспертных оценках и обеспечивает более полное представление о возможных сценариях развития событий. В конечном итоге это способствует более обоснованным и информированным решениям в условиях неопределенности и комплексности предметных областей.

Опишем этот алгоритм. Пусть у нас есть слабо разрешимая ИСЛУ вида (2.2), описывающая уже имеющиеся знания о предметной области. Исходя из этих знаний, требуется получить максимальную согласованную оценку некоторого нового предложения ψ данной сигнатуры.

Под максимальной согласованной оценкой мы будем понимать такое множество значений $A \subseteq [0; 1]$, что любое его надмножество (т. е. $A \subset B \subseteq [0; 1]$) B уже является несогласованной оценкой. Согласно теореме 2, максимальная согласованная оценка предложения ψ будет представлять собой замкнутый интервал $[\alpha; \beta] \subseteq [0; 1]$. Таким образом, целью нашего алгоритма является нахождение левой и правой границ максимального оценочного интервала.

Так как для определения точных границ оценочного интервала нам потребуется сделать счетное число шагов, то на практике будем искать эти границы некоторой заданной точностью Δ .

На **первом шаге** алгоритма мы приводим предложение к виду СДНФ и записываем его в виде многочлена

$$\sum_{j=1}^m \gamma_j x_j,$$

где коэффициент γ_j равен 1, если конъюнкт под номером j содержится в СДНФ-представлении предложения ψ , и 0 в противном случае.

Зададим начальный диапазон поиска границ интервальных оценок: $[\underline{\alpha}; \bar{\alpha}]$ для левой границы оценки и $[\underline{\beta}; \bar{\beta}]$ для правой границы оценки. Очевидно, что как левая граница α , так и правая граница интервальной оценки предложения ψ лежат в интервале $[0; 1]$. Поэтому положим

$$\underline{\alpha} = \underline{\beta} = 0 \text{ и } \bar{\alpha} = \bar{\beta} = 1.$$

На следующих шагах алгоритма будет происходить пошаговое уточнение границ. Алгоритм будет завершаться, когда диапазон возможных значений границ (т. е. длина отрезков $[\underline{\alpha}; \bar{\alpha}]$ и $[\underline{\beta}; \bar{\beta}]$ будет меньше заданного параметра Δ .

Поиск левой границы оценки.

На **втором шаге** ИСЛУ (2.2) пополняется новым уравнением и принимает вид:

$$\begin{cases} \sum_{j=1}^m \varepsilon_{ij} x_j = [\alpha_i; \beta_i], & i = 1, \dots, n, \\ \sum_{j=1}^m x_j = 1, \\ \sum_{j=1}^m \gamma_j x_j = \left[\underline{\alpha}; \frac{\underline{\alpha} + \bar{\alpha}}{2} \right]. \end{cases} \quad (2.3)$$

Далее проверяем, является ли система (2.3) слабо разрешимой.

На **третьем шаге** происходит сужение диапазона $[\underline{\alpha}; \bar{\alpha}]$ левой границы оценки предложения ψ . Если система (2.3) является слабо разрешимой, то определяем новый диапазон:

$$\underline{\alpha} \Leftarrow \underline{\alpha} \text{ и } \bar{\alpha} \Leftarrow \frac{\underline{\alpha} + \bar{\alpha}}{2}.$$

Если же система (2.3) не является слабо разрешимой, то переопределяем диапазон левой границы следующим образом:

$$\underline{\alpha} \Leftarrow \frac{\underline{\alpha} + \bar{\alpha}}{2} \text{ и } \bar{\alpha} \Leftarrow \bar{\alpha}.$$

На **четвертом шаге** проверяем, достигли ли мы заданной точности вычислений. Если выполняется условие $\bar{\alpha} - \underline{\alpha} \leq \Delta$, то определяем значение левой границы оценочного интервала $\alpha^* = \bar{\alpha}$. В противном случае переходим на шаг 2.

Поиск правой границы оценки.

На **втором шаге** ИСЛУ (2.2) пополняется новым уравнением и принимает вид:

$$\begin{cases} \sum_{j=1}^m \varepsilon_{ij} x_j = [\alpha_i; \beta_i], & i = 1, \dots, n, \\ \sum_{j=1}^m x_j = 1, \\ \sum_{j=1}^m \gamma_j x_j = \left[\frac{\underline{\beta} + \bar{\beta}}{2}; \bar{\beta} \right]. \end{cases}$$

Далее проверяем, является ли система (2.4) слабо разрешимой.

На **третьем шаге** производим сужение диапазона $[\underline{\beta}; \bar{\beta}]$ правой границы оценки предложения ψ . Если система (2.4) является слабо разрешимой, то определяем новый диапазон:

$$\underline{\beta} \Leftarrow \frac{\underline{\beta} + \bar{\beta}}{2} \text{ и } \bar{\beta} \Leftarrow \bar{\beta}.$$

Если же система (2.4) не является разрешимой, то переопределяем диапазон левой границы следующим образом:

$$\underline{\beta} \Leftarrow \underline{\beta} \text{ и } \bar{\beta} \Leftarrow \frac{\underline{\beta} + \bar{\beta}}{2}.$$

На **четвертом шаге** проверяем, достигли ли мы заданной точности вычислений. Если выполняется условие $\bar{\beta} - \underline{\beta} \leq \Delta$, то определяем значение левой границы оценочного интервала $\beta^* = \underline{\beta}$. В противном случае переходим на шаг 2.

Рассмотрим работу алгоритма на примере. Имея интервальные оценки для предложений $\varphi_1, \varphi_2, \varphi_3$, найдем максимальную интервальную оценку для предложения φ_4 с точностью вычислений $\Delta = 0,02$ (табл. 2).

Таблица 2

Входные данные для предсказания оценки для новой формулы

Table 2

Input data for predicting the estimate for the new formula

Формула	Оценка
$\varphi_1 = (P(a) \wedge Q(b)) \vee \overline{R(c)}$	$[0,7; 0,8]$
$\varphi_2 = R(c) \wedge \overline{Q(b)}$	$[0,2; 0,3]$
$\varphi_3 = \overline{P(a)} \vee Q(b)$	$[0,8; 0,9]$
$\varphi_4 = R(c)$	?

Для поиска левой границы оценочного интервала строим следующую ИСЛУ:

$$\begin{cases} x_2 + x_4 + x_6 + x_8 = [0; 0,5] \\ x_1 + x_3 + x_5 + x_7 + x_8 = [0,7; 0,8] \\ x_2 + x_6 = [0,2; 0,3] \\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = [0,8; 0,9] \\ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1 \\ x_i \in [0,1], \quad i = 1, \dots, 8 \end{cases}$$

В табл. 3 показаны шаги работы алгоритма. На шестой итерации выполнения алгоритма точность вычисления станет меньше заданного порога $\Delta = 0,02$.

Таблица 3

Шаги алгоритма предсказания оценки для новой формулы

Table 3

Steps of the estimation prediction algorithm for the new formula

Шаг	Диапазон левой оценки формулы φ_4 , $[\underline{\alpha}; \bar{\alpha}]$	Результат проверки разрешимости ИСЛУ	Точность вычисления, $\bar{\alpha} - \underline{\alpha}$
1	[0; 0,5]	True	0,5
2	[0; 0,25]	True	0,25
3	[0; 0,125]	False	0,125
3	[0,125; 0,25]	True	0,125
4	[0,125; 0,1875]	False	0,0625
4	[0,1875; 0,25]	True	0,0625
5	[0,1875; 0,21875]	True	0,03125
6	[0,1875; 0,203125]	True	0,015625

Таким образом, определим левую границу оценочного интервала $\alpha^* = 0,203125$. Поиск правой границы осуществляется аналогично, точный ответ находится в интервале $[0,984375; 1]$. Следовательно, определим правую границу оценочного интервала $\beta^* = 0,984375$.

3. Интерфейс программной системы

Интерфейс программной системы был реализован с использованием библиотеки Tkinter, которая является стандартной библиотекой Python для создания графического пользовательского интерфейса (GUI).

Пользователь имеет возможность описывать события, используя диалоговое окно, где каждое событие представлено в виде логической формулы. Пользователь может задавать субъективные вероятности наступления этих событий, представленные подинтервалами отрезка от 0 до 1. В диалоговом окне пользователь может ввести любое количество таких формул и соответствующих им оценочных значений.

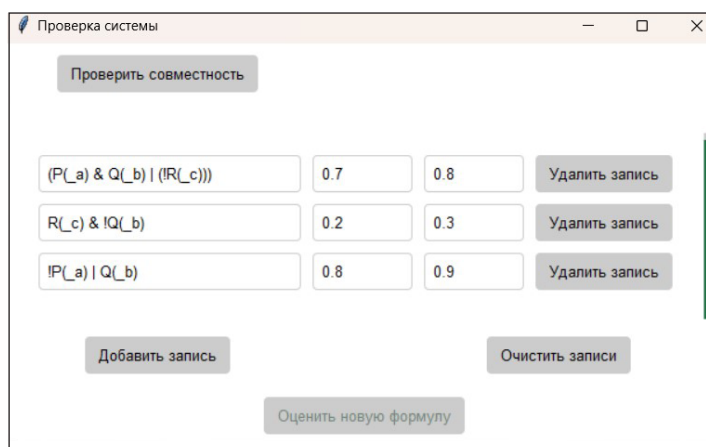
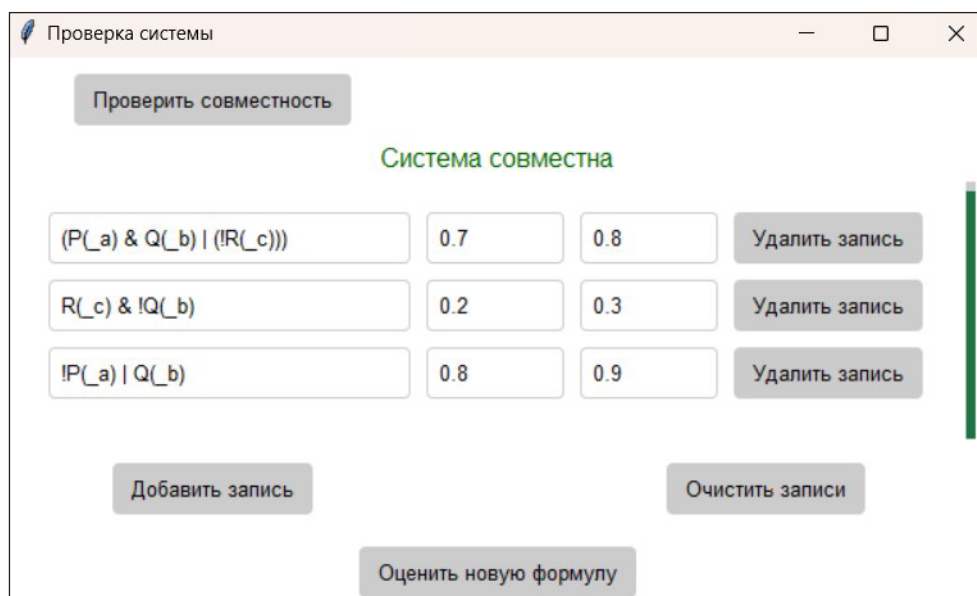


Рис. 1. Интерфейс

Fig. 1. Interface

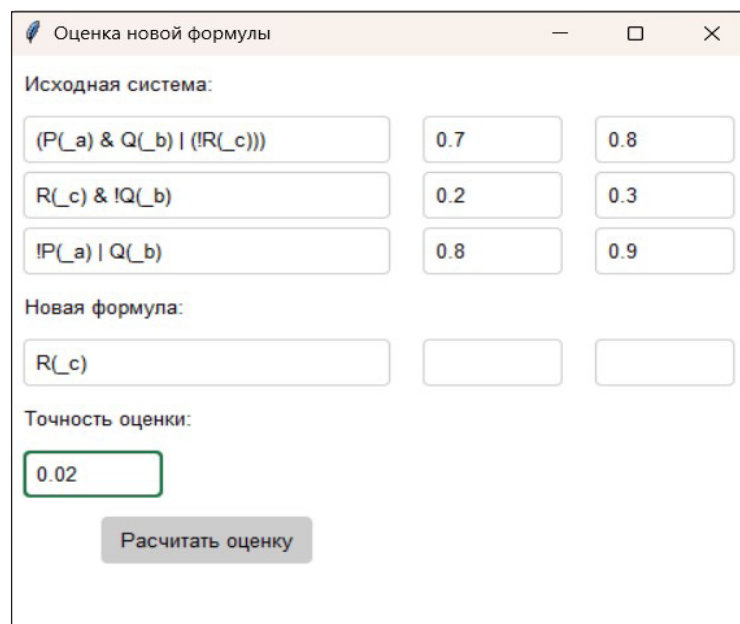
После нажатия кнопки «Проверить совместность» введенные данные передаются на вход алгоритму проверки совместности экспертных оценок (п. 2). Результат проверки появляется на экране.



Формула	Значение 1	Значение 2	Действие
$(P_a) \& Q_b \mid (!R_c)$	0.7	0.8	Удалить запись
$R_c \& !Q_b$	0.2	0.3	Удалить запись
$!P_a \mid Q_b$	0.8	0.9	Удалить запись

Рис. 2. Результат проверки
Fig. 2. The result of the check

В случае если введенная система оказалась совместной, пользователь может ввести новую формулу и получить оценку для нее по алгоритму из п. 3, которая не противоречит исходной системе.



Исходная система:

$(P_a) \& Q_b \mid (!R_c)$	0.7	0.8
$R_c \& !Q_b$	0.2	0.3
$!P_a \mid Q_b$	0.8	0.9

Новая формула:

R_c

Точность оценки:

0.02

Расчитать оценку

Рис. 3. Интерфейс для оценки новой формулы
Fig. 3. Interface for evaluating a new formula

Рис. 4. Результат оценки новой формулы
 Fig. 4. The result of evaluating the new formula

Заключение

В работе представлена программная система, предназначенная для проверки согласованности множества субъективных оценок событий в различных областях. События формализуются с использованием языка логики предикатов, при этом их оценки выражаются подинтервалами в диапазоне от 0 до 1. Для получения вывода о согласованности оценок строится система линейных интервальных уравнений. Кроме того, представлен инструмент для оценки нового события, используя имеющиеся оценки, что необходимо из-за ограниченной возможности экспертов оценить все возможные события в предметной области.

Список литературы

1. Ramsey F. P. Truth and probability // Gärdenfors P., Sahlin N.-E., eds. Decision, Probability and Utility: Selected Readings. Cambridge University Press, 1988. P. 19–47. DOI: <https://doi.org/10.1017/CBO9780511609220.003>
2. Palchunov D. E., Yakhyaeva G. E. Integration of Fuzzy Model Theory and FCA for Big Data Mining // SIBIRCON 2019 – International Multi-Conference on Engineering, Computer and Information Sciences, Proceedings. 2019. P. 961–966.. DOI: 10.1109/SIBIRCON48586.2019.8958216
3. Карасев О. И., Китаев А. Е., Миронова И. И., Шинкаренко Т. В. Экспертные процедуры в форсайте: особенности взаимодействия с экспертами в проектах по долгосрочному прогнозированию // Вестник Санкт-Петерб. ун-та. Серия: Социология. 2017. № 2. С. 169–184. DOI: 10.21638/11701/spbu12.2017.203
4. Yakhyaeva G., Skokova V. Subjective Expert Evaluations in the Model-Theoretic Representation of Object Domain Knowledge // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2021. 12948 LNAI. P. 152–165. DOI: 10.1007/978-3-030-86855-0_11

5. **Малаева Е. Д., Яхьяева Г. Э.** Программная система визуализации и проверки согласованности оценочных знаний экспертов // Вестник НГУ. Серия: Информационные технологии. 2023. Т. 21, № 1. С. 32–45.
6. **Yakhyaeva G. E., Palchunova O. D.** Fuzzy Models as a Formalization of Expert's Evaluative Knowledge // Pattern Recognition and Image Analysis, 2023. Vol. 33, no. 3. P. 529–535. DOI <https://doi.org/10.1134/S105466182303046X>
7. **Palchunov D., Yakhyaeva G.** Representation of Knowledge Using Different Structures of Concepts // CEUR Workshop Proceedings. 2020. Vol. 2729. P. 69–74.
8. **Пальчунов Д. Е., Яхьяева Г. Э.** Нечеткие алгебраические системы // Вестник НГУ. Серия: Математика, механика, информатика. 2010. Т. 10, вып. 3. С. 76–93.
9. **Yakhyaeva G.** Method for Verifying the Logical Correctness of Experts' Evaluative Knowledge // 2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). Yekaterinburg, Russian Federation, 2022. P. 850–854. DOI: 10.1109/SIBIRCON56155.2022.10016972
10. **Yakhyaeva G. E.** Separable Fuzzy Models // 2023 IEEE 16th International conference of actual problems of electronic instrument engineering (APEIE). P. 1480–1483. DOI: 10.1109/APEIE59731.2023.10347792
11. **Yakhyaeva G. E.** On the Local Coordination of Fuzzy Valuations // The Bulletin of Irkutsk State University. Series Mathematics. 2023. Vol. 46. P. 130–144. DOI: 10.26516/1997-7670.2023.46.130
12. **Шарый С. П.** Конечномерный интервальный анализ. Новосибирск: XYZ, 2022. 654 с.
13. **Шарый С. П., Шарая И. А.** Распознавание разрешимости интервальных уравнений и его приложения к анализу данных // Вычислительные технологии. 2013. Т. 18, № 3. С. 80–109.

References

1. **Ramsey F. P.** Truth and probability. In: *Gärdenfors P, Sahlin N-E, eds. Decision, Probability and Utility: Selected Readings.* Cambridge University Press, 1988, pp. 19–47. DOI: <https://doi.org/10.1017/CBO9780511609220.003>
2. **Palchunov D. E., Yakhyaeva G. E.** Integration of Fuzzy Model Theory and FCA for Big Data Mining. *SIBIRCON 2019 – International Multi-Conference on Engineering, Computer and Information Sciences, Proceedings*, 2019, pp. 961–966. DOI: 10.1109/SIBIRCON48586.2019.8958216
3. **Karasev O. I., Kitaev A. E., Mironova I. I., Shinkarenko T. V.** Expert procedures in foresight: interaction with expert professionals during long-term forecasting research. *Vestnik SPbSU. Sociology*, 2017, vol. 10, iss. 2, pp. 169–184. (in Russ.) DOI: 10.21638/11701/spbu12.2017.203
4. **Yakhyaeva G., Skokova V.** Subjective Expert Evaluations in the Model-Theoretic Representation of Object Domain Knowledge. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2021, 12948 LNAI, pp. 152–165. DOI: 10.1007/978-3-030-86855-0_11
5. **Malaeva E. D., Yakhyaeva G. E.** Software System for Visualization and Checking the Consistency of Experts' Evaluative Knowledge. *Vestnik NSU. Series: Information Technologies*, 2023, vol. 21, no. 1, pp. 32–45. (in Russ.) DOI 10.25205/1818-7900-2023-21-1-32-45
6. **Yakhyaeva G. E., Palchunova O. D.** Fuzzy Models as a Formalization of Expert's Evaluative Knowledge. *Pattern Recognition and Image Analysis*, 2023, vol. 33, no. 3, pp. 529–535. (in Russ.), <https://doi.org/10.1134/S105466182303046X>
7. **Palchunov D., Yakhyaeva G.** Representation of Knowledge Using Different Structures of Concepts. *CEUR Workshop Proceedings*, 2020, vol. 2729, pp. 69–74.

8. **Palchunov D. E., Yakhyaeva G. E.** Fuzzy algebraic systems. *Vestnik NSU. Series: Mathematics, mechanics, computer science*, 2010, vol. 10, no. 3, pp. 76–93. (in Russ.)
9. **Yakhyaeva G.** Method for Verifying the Logical Correctness of Experts' Evaluative Knowledge. *2022 IEEE International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*, Yekaterinburg, Russian Federation, 2022, pp. 850–854. DOI: 10.1109/SIBIRCON56155.2022.10016972
10. **Yakhyaeva G. E.** Separable Fuzzy Models. *2023 IEEE 16th International conference of actual problems of electronic instrument engineering (APEIE)*, pp. 1480–1483. DOI: 10.1109/APEIE59731.2023.10347792
11. **Yakhyaeva G. E.** On the Local Coordination of Fuzzy Valuations. *The Bulletin of Irkutsk State University. Series Mathematics*, 2023, vol. 46, pp. 130–144. DOI: 10.26516/1997-7670.2023.46.130
12. Konechnomernyj interval'nyj analiz [Finite-dimensional interval analysis]. Novosibirsk, XYZ, 2022.
13. **Sharyi S. P., Sharaya I. A.** Raspoznavanie razreshimosti interval'nykh uravnenii i ego prilozheniya k analizu dannykh [Recognizing the solvability of interval equations and its applications to data analysis]. *Vychislitel'nye tekhnologii*, 2013, vol. 18, no. 3, pp. 80–109.

Сведения об авторах

Дымонт Надежда Андреевна, студент НГУ

Малаева Елена Дмитриевна, студент НГУ

Яхьяева Гульнара Эркиновна, кандидат физико-математических наук, доцент кафедры общей информатики НГУ

Information about the Authors

Nadezhda A. Dymont, Student

Elena D. Malaeva, Student

Gulnara E. Yakhyaeva, Candidate of Physical and Mathematical Sciences, Associate Professor

Статья поступила в редакцию 27.04.2024;

одобрена после рецензирования 06.06.2024; принята к публикации 06.06.2024

The article was submitted 27.04.2024;

approved after reviewing 06.06.2024; accepted for publication 06.06.2024

Научная статья

УДК 519.688

DOI 10.25205/1818-7900-2024-22-2-20-32

Информационные технологии для расчета антропогенного потока тепла в условиях городской застройки

Игорь Александрович Евсиков

Санкт-Петербургский государственный архитектурно-строительный университет
Главная геофизическая обсерватория им. А. И. Воейкова,
Санкт-Петербург, Россия

ievsikov@lan.spbgasu.ru; <https://orcid.org/0000-0003-4793-2975>

Аннотация

Современные города сталкиваются с необходимостью эффективного управления энергоресурсами и минимизацией воздействия человеческой деятельности на окружающую среду. Одним из значимых аспектов антропогенного воздействия является тепловой поток, порождаемый зданиями, транспортом и промышленностью. В статье представлен комплекс программ для оценки антропогенного потока тепла (АПТ), вызванного теплопотерями зданий в отопительный период. Исследование основано на пространственной геометрической модели города, построенной с использованием данных OpenStreetMap. Теплофизические свойства ограждающих конструкций и удельные характеристики расхода тепловой энергии получены из строительных норм. Предложенный метод включает этапы моделирования зданий, фильтрации, дополнения информации из Яндекс Карт и ГИС ЖКХ, устранения коллизий, присвоения удельных характеристик и расчета АПТ. Реализация метода осуществлена через скрипты для платформы Rhinoceros, известной своим широким функционалом и визуальной средой программирования Grasshopper. Предложенный подход позволяет эффективно анализировать и визуализировать антропогенный поток тепла в городах, что является ключевым шагом в разработке стратегий устойчивого управления энергоресурсами и снижения негативного воздействия на окружающую среду.

Ключевые слова

модель городской застройки, геоинформационные системы, OpenStreetMap, Яндекс Карты, ГИС ЖКХ, строительные нормы и правила

Финансирование

Исследование выполнено при поддержке важнейшего инновационного проекта государственного значения (ВИПГЗ) «Единая национальная система мониторинга климатически активных веществ» (распоряжение Правительства РФ от 29 октября 2022 г. № 3240–р) в рамках НИР «Расширение системы климатического и экологического мониторинга и прогнозирования на территории России в целях обеспечения адаптационных решений в отраслевом и региональном разрезе, включая борьбу с опустыниванием».

Для цитирования

Евсиков И. А. Информационные технологии для расчета антропогенного потока тепла в условиях городской застройки // Вестник НГУ. Серия: Информационные технологии. 2024. Т. 22, № 2. С. 20–32. DOI 10.25205/1818-7900-2024-22-2-20-32

© Евсиков И. А., 2024

Information Technologies for Calculating Anthropogenic Heat Flux in Urban Areas

Igor A. Evsikov

Saint Petersburg State University of Architecture and Civil Engineering
Voeikov Main Geophysical Observatory, Saint Petersburg, Russian Federation
ievnikov@lan.spbgasu.ru; <https://orcid.org/0000-0003-4793-2975>

Abstract

Modern megalopolis face challenges in effectively managing energy resources and minimizing the impact of human activities on the environment. One of the significant aspects of anthropogenic impact is the heat flow generated by buildings, transport and industry. This article presents a set of programs for assessing anthropogenic heat flow (AHF) caused by heat loss from buildings during the heating season. The study is based on a spatial geometric model of the city constructed using OpenStreetMap data. The thermophysical properties of enclosing structures and specific characteristics of thermal energy consumption are obtained from building codes. The proposed method includes the stages of buildings modeling, filtering, supplementing information from Yandex Maps and State Information System of Housing and Communal Services, eliminating collisions, assigning specific characteristics and calculating AHF. The method is implemented through scripts for the Rhinoceros platform, known for its wide functionality and visual programming environment Grasshopper. The proposed approach allows us to effectively analyze and visualize anthropogenic heat flow in cities, which is a key step in developing strategies for sustainable energy management and reducing negative environmental impacts.

Keywords

Urban development model, geographic information systems, OpenStreetMap, Yandex Maps, State Information System of Housing and Communal Services, building codes and regulations

Funding

The research was carried out with the support of the most important innovative project of state importance “Unified national system for monitoring of climatically active substances” (Russian Federation Government Order No. 3240-p dated October 29, 2022) within the framework of the R&D “Expansion of the system of climatic and environmental monitoring and Expansion of the system of climatic and environmental monitoring and forecasting on the territory of Russia in order to provide adaptation solutions in the sectoral and regional context, including combating desertification”.

For citation

Evsikov I. A. Information technologies for calculating anthropogenic heat flux in urban areas. *Vestnik NSU. Series: Information Technologies*, 2024, vol. 22, no. 2, pp. 20–32 (in Russ.) DOI 10.25205/1818-7900-2024-22-2-20-32

Введение

Современные города сталкиваются с неизбежной потребностью в эффективном управлении энергоресурсами и снижении воздействия человеческой деятельности на окружающую среду. Одним из ключевых аспектов антропогенного воздействия является тепловой поток, создаваемый различными источниками, такими как здания, транспорт и промышленность. Зимой, согласно мезомасштабным атмосферным моделям, в некоторых случаях влияние антропогенных тепловых выбросов на интенсивность городского острова тепла может составлять от 50 до 80 % [1], а разность температур в городе и на окраине может достигать до 10 °C [2; 3]. Непосредственно на здания приходится от 50 до 65 % от общей тепловой эмиссии, в зависимости от сезона и особенностей конкретного города [4].

Использование информационных технологий оказывается полезным для расчета антропогенного потока тепла (АПТ), так как их применение позволяет построить модель городской застройки, на основе которой можно оценить тепловую эмиссию от зданий. В последние годы развитие информационного моделирования зданий и геоинформационных технологий, привели к появлению метода оценки АПТ «снизу вверх». Например, модели городского погранично-го слоя с интегрированными моделями энергопотребления зданий, такими как ВЕР + ВЕМ [4]

и DCEP [5]. В рамках такого подхода рассматриваются некоторые типы зданий, для которых из различных источников собирается информация об изменении их энергетического баланса с течением времени и вычисляется АПТ, а затем предполагается, что однотипные здания генерируют со временем такой же поток. Такие подходы обеспечивают достаточно высокую точность, но требуют детальных данных о зданиях и больших вычислительных ресурсов. Однако в настоящее время такой подход реализован только для городов Бостон и Берлин.

В данной работе предлагается другой вариант расчета АПТ «снизу вверх» на основе действующих строительных норм, для реализации которого необходима информация о геометрии городских зданий, а точнее, о площади ограждающих конструкций и объемах зданий.

Существует множество геоинформационных сервисов, таких как Яндекс Карты¹, Google Maps, 2ГИС, OpenStreetMap² (OSM) и другие, позволяющие визуализировать городскую застройку. Однако перечисленные сервисы за исключением OSM являются закрытыми системами, не разрешающими использовать информацию о геометрии строений, что, в свою очередь, не дает возможность оценить необходимые площади и объемы зданий. Знание площади ограждающей конструкции здания в совокупности с сопротивлением ее теплопередаче или объема здания вместе с удельной характеристикой расхода тепловой энергии на отопление и вентиляцию позволяет рассчитать АПТ. Суммируя потоки тепла от всех зданий, получается полный поток тепла, генерируемый городом, а учитывая пространственное расположение зданий, находим распределение АПТ по территории города.

Для того чтобы оценить требуемые площади и объемы, написана программа (скрипт) на Python, собирающая геометрическую и географическую информацию из OSM. Скрипт использует библиотеку «OSMnx»³, которая преобразует XML-данные OSM в табличную структуру данных GeoDataFrame и позволяет сохранить результат в формате геопространственных векторных данных, таких как shapefile, geojson, georackage, включая сопутствующую информацию об объектах (название, адрес, высоту, тип, материал, и т. п.). На основе shapefile строится геометрическая модель города в программе трехмерного моделирования Rhinoceros с помощью визуальной среды программирования Grasshopper, причем shapefile оказался более удобным для работы в среде Rhinoceros.

В OSM информация о высоте и этажности есть не у всех зданий. Восполнить недостающую информацию можно, например, из сервисов Яндекс Карт и ГИС ЖКХ⁴. В некоторых случаях высота здания доступна напрямую из сервиса Яндекс Карт через сайт «Народная карта»⁵, где для каждого здания доступна карточка с информацией о его адресе, типе и высоте. Однако такую информацию можно получить только визуально, без возможности автоматизировать ее сбор с последующим сохранением. При отсутствии данных о высоте здания на сайте «Народная карта» используется информация ГИС ЖКХ. На сайте ГИС ЖКХ в некоторых случаях имеется информация о количестве этажей, дате ввода в эксплуатацию, серии/типе проекта здания и т. п., но такая информация доступна только на первых 19 страницах результатов поиска, что дает 1900 записей. Чтобы обойти эту проблему, формируется запрос напрямую к серверу ГИС ЖКХ по «Глобальному уникальному идентификатору дома по ФИАС» (ГУИД), что позволяет получить полную информацию из базы данных ГИС ЖКХ. Таким образом, для некоторых зданий удастся определить количество этажей. Знание высоты здания позволяет определить его объем и площадь ограждающих конструкций.

Таким образом, задача заключается в построении геометрической модели города, позволяющей оценить АПТ. Предполагая, что здания соответствуют действующим строительным нормам, разработаны два алгоритма расчета АПТ, описанные в [6; 7]. Для расчета по первому

¹ <https://yandex.ru/maps-api/>

² <https://www.openstreetmap.org>

³ <https://osmnx.readthedocs.io>

⁴ <https://dom.gosuslugi.ru/#!/houses>

⁵ <https://n.maps.yandex.ru/>

алгоритму в модель городской застройки вносится информация о коэффициентах теплопередачи ограждающих конструкций, одинаковая для всех зданий внутри одного города. Второй алгоритм учитывает высоту и тип здания, поэтому в рамках второго подхода необходимо каждому объекту добавить информацию об удельной характеристике расхода тепловой энергии на отопление и вентиляцию. Оценка АПТ происходит с учетом разности температур воздуха снаружи и внутри здания.

Этапы построения модели городской застройки

Построение модели для расчета АПТ состоит из нескольких этапов:

- 1) моделирование контуров зданий и наполнение атрибутов (высота, этажность, тип здания);
- 2) фильтрация по площади и по типу;
- 3) дополнение информации из Яндекс Карт и ГИС ЖКХ;
- 4) устранение коллизий (пересечений);
- 5) добавление теплофизических свойств;
- 6) разделение контуров зданий по расчетной сетке.

Часть этапов выполнена в виде отдельных монофункциональных программ (скриптов) для платформы Rhinoceros на языке программирования Python. Другая часть этапов реализована в визуальной среде программирования Grasshopper, встроенной в Rhinoceros. Grasshopper ускоряет процесс программирования, так как уже содержит готовые функциональные блоки – «ноды». При этом создаваемая программа становится более наглядной, так как состоит из отдельных нодов и связей между ними, подобно блок-схеме. Рассмотрим более подробно каждый этап моделирования.

На **первом этапе** следует преобразовать содержимое shapefile в объекты программы Rhinoceros. Стандартные инструменты не позволяют выполнить такую конвертацию, поэтому используется дополнительный пакет Gismo для Grasshopper, который позволяет автоматически генерировать геометрию городской застройки и территорию, свободную от нее (дороги, парки, водоемы и т. д.). Максимальный радиус моделируемой территории равен 27 км, но зависит от широты места, причем на севере он меньше, чем на юге. Так, для Новосибирска радиус области моделирования не превышает 16 км. Учитывая, что размер Новосибирска составляет 27×45 км, таких областей потребуется как минимум две с последующим удалением дубликатов контуров зданий на их пересечении. Также следует отметить, что пакет Gismo был создан в 2019 г. и с тех пор не обновлялся, а в OSM с 2023 г. появился дополнительный тег `map_made`, позволяющий идентифицировать различные сооружения, такие как цистерны, шахты, насосные станции, антенны и т. п. По этим причинам Gismo используется в проекте только для конвертации shapefile и для трансформации географических координат в прямоугольные и обратно. В частности, используется нод «Gismo_Read SHP», который разделяет содержимое shapefile на линии (контуров зданий) и соответствующие им ключи (теги OSM) со значениями (рис. 1). Некоторые здания состоят из нескольких замкнутых контуров, например, типичные для центра Санкт-Петербурга дома с дворами-колодцами. Для минимизации количества объектов (точка, линия, поверхность) с помощью стандартных нодов Grasshopper контуры зданий преобразуются в поверхности (Surface). В результате одному зданию или части здания (крыло, корпус, пристройка) соответствует один элемент модели – поверхность и свой набор атрибутов (пар ключей и значений). К примеру, модель городской застройки Новосибирска состоит из 77918 элементов (зданий, сооружений, и прочих объектов) и 18 атрибутов в каждом.

(рис. 2). По адресу из OSM находится соответствующая запись с адресом в реестре, по которому определяется ГУИД.

Адрес ОЖФ|Идентификационный код адреса дома в ГИС ЖКХ|Глобальный уникальный идентификатор дома по ФИАС|Код
630001, Новосибирская обл, г. Новосибирск, пер. Обские зори, д. 21||00bf5b28-32ec-4ca3-bf6d-854dcdbf0d49|5
630001, Новосибирская обл, г. Новосибирск, пер. Обские зори, д. 22||09a85b9d-3b0a-4cd7-9d82-853b3e2445d8|5
630001, Новосибирская обл, г. Новосибирск, пер. Обские зори, д. 40||f0edeefe-4982-414e-8963-99140cc6cfa7|5
630001, Новосибирская обл, г. Новосибирск, ул. 1-я Сухарная, д. 10||c96f4684-9698-4359-aca1-5f3c50708ae7|5
630001, Новосибирская обл, г. Новосибирск, ул. 1-я Сухарная, д. 101||24a95ae9-0371-4911-b123-ec6b7a5a41c7|5
630001, Новосибирская обл, г. Новосибирск, ул. 1-я Сухарная, д. 103||69292a7a-3365-4ead-8d05-de0af1260e25|5
630001, Новосибирская обл, г. Новосибирск, ул. 1-я Сухарная, д. 107||1cec3f7b-175f-4fd8-b740-c19d104aa71a|5
630001, Новосибирская обл, г. Новосибирск, ул. 1-я Сухарная, д. 109||636850e5-1c75-41e7-b462-443b0383f655|5
630001, Новосибирская обл, г. Новосибирск, ул. 1-я Сухарная, д. 113||20a9268d-6bf3-4045-b9a8-34a9a188c556|5

Рис. 2. Пример записи в реестре объектов жилищного фонда ГИС ЖКХ

Fig. 2. Example of a record in the housing stock register of the State Information System of Housing and Communal Services

Отметим, что адреса в ГИС ЖКХ и OSM могут незначительно различаться (порядок слов, регистр, сокращения), поэтому сравнивать их нужно не строго. В проекте для этого используется библиотека FuzzyWuzzy в Python, основанная на метрике Левенштейна, измеряющая по модулю разность между двумя последовательностями символов [11]. Допустимой выбрана оценка выше 98 и заранее добавлены известные замены, например ул. – улица, пер. – переулок и т. д. Так, для Новосибирска удалось получить ГУИД для 4269 зданий без высоты, но с адресами. Для таких зданий с сервера ГИС ЖКХ было получено 3862 записи с информацией о количестве этажей.

После такой коррекции из 72 458 зданий, излучающих тепло, остается 6333 (8,7 %) здания без информации о высоте, таким зданиям присваивается минимально допустимое значение, равное 3 м (предполагается, что здания меньшей высоты не существуют).

На **четвертом этапе** устраняются коллизии (пересечения) контуров зданий. В OSM, в некоторых случаях, хранится два типа контуров зданий: здание целиком и его части (соответствующие теги: building и building:part), которые могут пересекаться (рис. 3, а, б).

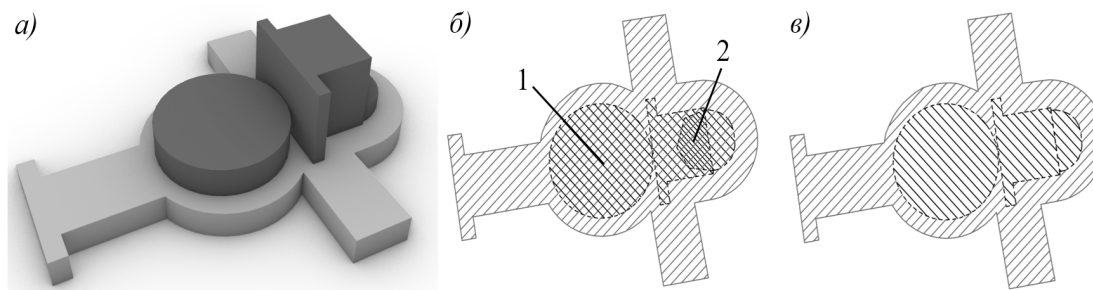


Рис. 3. Здание и его части: а – 3D-вид; б – вид сверху до устранения пересечений

(1 – пересечение здания и его части; 2 – пересечение двух частей здания);

в – вид сверху после устранения пересечений

Fig. 3. The building and its parts: а – 3D view; б – top view before eliminating intersections

(1 – intersection of the building and its part; 2 – intersection of two parts of the building);

в – top view after eliminating intersections

Для устранения таких коллизий требуется учитывать высоту пересекающихся частей зданий, поэтому для дальнейшего корректного расчета АПТ отсекается часть контура того элемента, у которого высота ниже (рис. 3, в).

Для обнаружения коллизий можно применить как минимум два подхода. Самый простой способ заключается в том, что на начальной итерации проверяется пересечение первого кон-

тура здания со всеми остальными, т. е. $N-1$ контуров, где N – количество обрабатываемых контуров. На следующем шаге необходимо проверить пересечение второго контура с оставшимися $N-2$ контурами и т. д. В этом случае скорость обработки будет пропорциональна сумме арифметической прогрессии.

Второй способ позволяет оптимизировать процесс поиска коллизий, что возможно, если проверять на пересечение только контуры, ближайшие к рассматриваемому зданию. Решить такую задачу позволяет упорядочение данных при помощи R-дерева, являющегося древовидной структурой данных для организации доступа к пространственно расположенным объектам [12]. Представление информации о координатах центров контуров зданий в виде R-дерева позволяет быстро найти все здания при заданном радиусе поиска. При этом предполагается, что не существует зданий длиной больше 1 км, поэтому в алгоритме используется поиск с радиусом 500 м. Количество зданий в такой области поиска не превышает 1000 и скорость обработки в худшем случае составит $N \cdot 1000$ итераций. Такой подход увеличивает скорость поиска коллизий контуров зданий в Москве более чем в 122 раза, в Санкт-Петербурге – в 73, в Новосибирске – в 36 раз по сравнению с первым алгоритмом.

На **пятом этапе** каждому объекту добавляется тег «heat», значение которого – «удельная характеристика расхода тепловой энергии на отопление и вентиляцию», учитывающая высоту и тип здания, согласно таблице 14 СП 50.13330.2012 [13]. Также в общие данные заносится информация о коэффициентах теплопередачи ограждающих конструкций (см. табл. 3 СП 50.13330.2012), соответствующая климатическим условиям города.

На **шестом этапе** задается расчетная сетка размером 30×30 м. Если здание попадает на несколько расчетных ячеек, то его контур разбивается на части, попадающие в ту или иную ячейку. Так как расчет внутри каждой ячейки не зависит от соседних, то в алгоритм можно включить распараллеливание, что в разы ускоряет получение результата на современных многопоточных процессорах.

В результате получаются данные, где каждая строка записи хранит информацию об отдельной непустой ячейке, содержащей координаты ее центра, суммарную площадь наружных стен и крыш, внутренний объем и среднюю высоту зданий, а также значения АПТ по 1-му и 2-му алгоритму (см. [6–8; 14]).

Результаты

В настоящее время с помощью представленного комплекса скриптов построены детальные модели 37 городов России с населением более 500 тысяч человек [10]. В отличие от [10] в данной работе модель Новосибирска дополнена информацией из ГИС ЖКХ. Для визуальной оценки качества построенной модели городской застройки плоские контуры зданий преобразованы в трехмерные объекты, что показывает согласованность модели с реальными зданиями города (рис. 4).

Рассматривается Новосибирск в рамках административных границ и его урбанизированная территория – городская застройка со зданиями высотой не менее двух этажей, точнее, не ниже 6 м. Площадь урбанизированной территории определяется суммарной площадью ячеек сетки, содержащих часть хотя бы одного здания с высотой не менее двух этажей. Это связано с тем, что городская застройка занимает лишь незначительную часть площади административной территории, а поэтому рассматривать пространственное распределение плотности АПТ в административных границах не имеет смысла, так как основная часть территории будет генерировать поток менее 5 Вт/м². Например, доля крыш зданий на административной территории Новосибирска занимает всего 6,7 %, а на урбанизированной территории без учета и с учетом данных ГИС ЖКХ – 29,9 и 29,8 %, соответственно. В табл. 1 приведены геометрические характеристики городской застройки Новосибирска. Также для сравнения в табл. 1–2 приводятся данные для Москвы и Санкт-Петербурга.



Рис. 4. Вид на Красный проспект со стороны ул. Державина в Новосибирске, фотография (панорама Яндекс Карт) сверху и трехмерная модель снизу
 Fig. 4. View of Krasny Avenue from Derzhavina Street in Novosibirsk, photo (Yandex Maps panorama) above and 3D model below

Таблица 1

Геометрические характеристики городской застройки

Table 1

Geometric characteristics of urban development

Город	Высота, м		Объем зданий, 10 ⁹ м ³	Площадь стен, 10 ⁶ м ²	Площадь крыш, 10 ⁶ м ²	Площадь террито- рии, 10 ⁶ м ²	Доля крыш в общей площади, %
	Среднее	Медиана					
Административная территория							
Москва	11,0	6,0	2,166	325,8	141,4	2558,1	5,5
Санкт-Петер- бург	10,2	6,0	1,284	183,3	95,07	1451,4	6,6
Новосибирск	6,3	3,0	0,375	55,80	33,86	502,54	6,7
Новосибирск*	6,4	3,0	0,378	56,41	33,86	502,54	6,7
Урбанизированная территория							
Москва	23,4	15,0	1,935	283,0	91,21	307,3	29,7
Санкт-Петер- бург	19,3	15,0	1,144	158,4	65,08	200,2	32,5
Новосибирск	19,9	15,0	0,298	41,12	16,45	54,99	29,9
Новосибирск*	19,8	15,0	0,301	41,70	16,60	55,76	29,8

*Новосибирск с информацией о высоте зданий из ГИС ЖКХ.

В табл. 2 приведено распределение зданий по высоте. Видно, что после уточнения данных по высоте зданий из ГИС ЖКХ скорректировано количество зданий высотой до 9 м включительно.

Таблица 2

Распределение зданий по высоте

Table 2

Distribution of buildings by height

Город	Доля зданий разной высоты (в метрах), %									Всего зданий, шт.
	высота неиз-вестна	(0–3]	(3–6]	(6–9]	(9–12]	(12–15]	(15–30]	(30–45]	> 45	
Москва	17	32	22	8	3	6	6	4	4	244 187
Санкт-Петербург	20	30	18	9	5	6	7	2	2	146 921
Новосибирск	14	49	20	4	2	4	4	1	1	72 458
Новосибирск*	9	53	21	5	2	4	4	1	1	

*То же что в табл. 1

В табл. 3 для административной и урбанизированной территории приведены плотность АПТ (Вт/м^2) и плотность антропогенной энергии ($10^6 \text{ Дж/}(\text{сут}\cdot\text{м}^2)$), выделяемой за сутки, а также энергия, генерируемая за сутки (10^{12} Дж/сут) и за отопительный период (10^{15} Дж), при температуре наружного воздуха $t_{\text{НАР}}$ равной средней температуре отопительного периода $t_{\text{ОТ}} = -7,9^\circ\text{C}$, по первому и второму алгоритмам в Новосибирске. Для сравнения среднее значение АПТ по Москве составляет порядка $55,9 \text{ Вт/м}^2$ [15]. По оценкам, полученным по модели COSMO-CLM, АПТ в разных районах Москвы находится в интервале $27\text{--}75 \text{ Вт/м}^2$ [16].

Таблица 3

Для административной и урбанизированной территории плотность АПТ и антропогенной энергии, а также энергия, выделяемая за сутки и за отопительный период, при $t_{\text{НАР}} = t_{\text{ОТ}}$, по первому и второму алгоритмам в Новосибирске

Table 3

For administrative and urban areas, the density of AHF and anthropogenic energy, as well as the energy released per day and during the heating period, at $t_{\text{НАР}} = t_{\text{ОТ}}$, according to the first and second algorithms in Novosibirsk

Административные границы		Урбанизированная территория	
Алгоритм 1	Алгоритм 2	Алгоритм 1	Алгоритм 2
Плотность АПТ (Вт/м^2)			
3,08	7,34	18,3	51,7
3,10*	7,43*	18,6*	52,5*
Плотность антропогенной энергии ($10^6 \text{ Дж/}(\text{сут}\cdot\text{м}^2)$)			
0,266	0,634	1,58	4,46
0,268*	0,642*	1,60*	4,54*
Энергия (10^{12} Дж/сут), выделяемая за сутки			
4,79	11,4	3,12	8,80
4,83*	11,6*	3,16*	8,94*
Энергия (10^{15} Дж), выделяемая за отопительный период			
29,7	70,7	19,3	54,5
29,9*	71,6*	19,6*	55,4*

*То же что в табл. 1

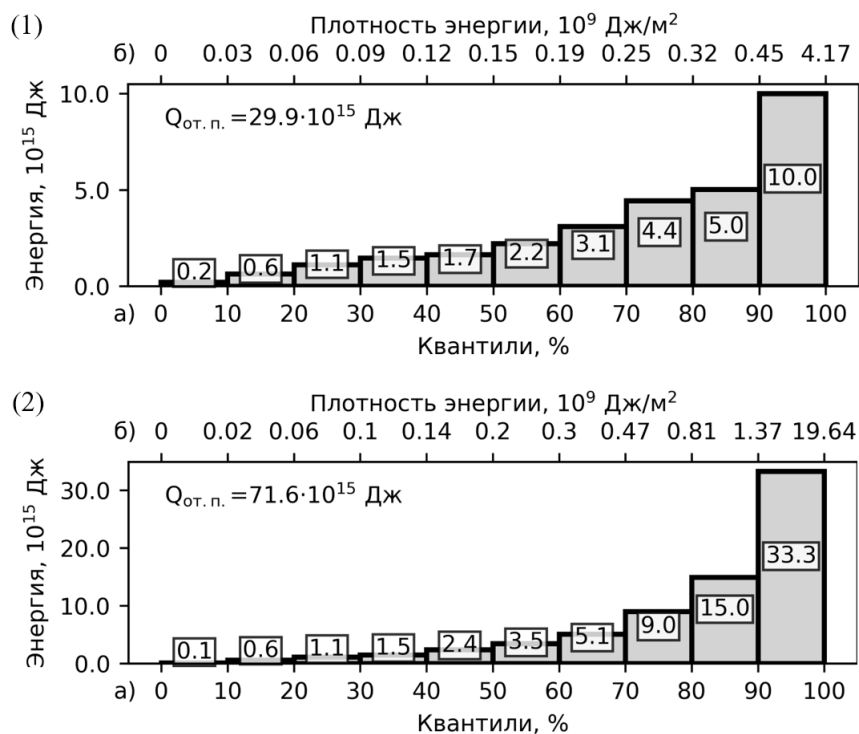


Рис. 5. Гистограммы распределений антропогенной энергии, ПДж, выделяемой за отопительный период зданиями с территории Новосибирска, рассчитанные по первому (1) и второму (2) алгоритмам:

a – уровень квантили; b – распределение плотности антропогенной энергии, ГДж/м²
 Fig. 5. Histograms of anthropogenic energy distributions, PJ, released during the heating period by buildings from the territory of Novosibirsk, calculated using the first (1) and second (2) algorithms:
 a – quantile level; b – anthropogenic energy density distribution, GJ/m²

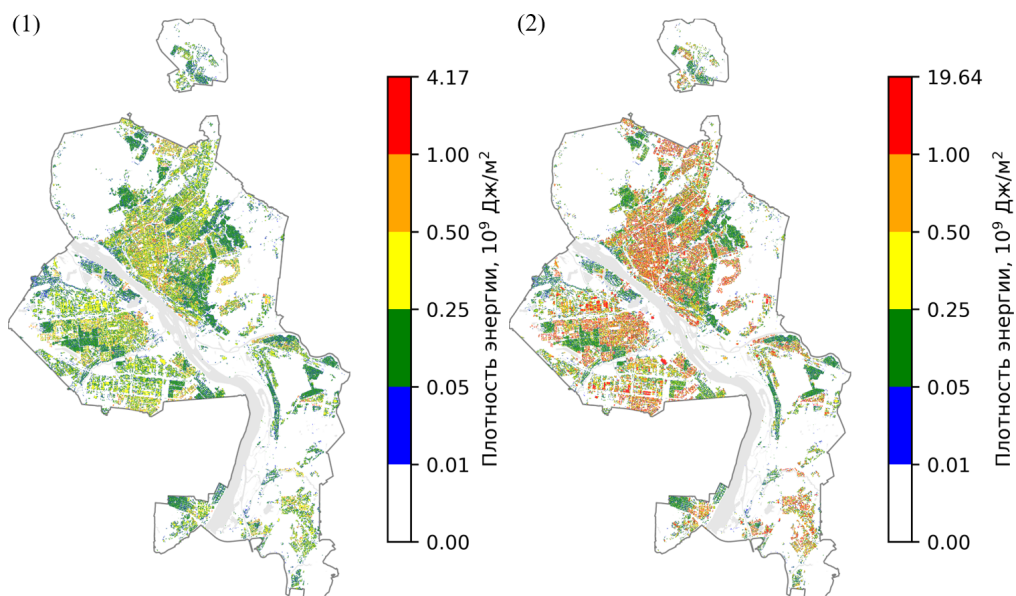


Рис. 6. Карты распределения плотности антропогенной энергии, ГДж/м², выделяемой за отопительный период зданиями с территории Новосибирска, рассчитанные по первому (1) и второму (2) алгоритмам

Fig. 6. Maps of anthropogenic energy density distribution, GJ/m², released during the heating period by buildings on the territory of Novosibirsk, calculated using the first (1) and second (2) algorithms

На рис. 5 представлены гистограммы распределения антропогенной энергии, выделяемой зданиями Новосибирска за отопительный период, рассчитанные по первому и второму алгоритму. Видно, что медиана (уровень квантили – 50 %) соответствует плотности энергии в 0,15 и 0,2 ГДж/м² для первого и второго алгоритма. Распределение АПТ за отопительный период неравномерно: 10 % городской территории со зданиями может генерировать 0,2–10 ПДж тепловой энергии по результатам первого алгоритма и 0,1–33,3 ПДж по результатам второго. Пиковые значения плотности энергии составляют 4,17 и 19,64 ГДж/м² для первого и второго алгоритма соответственно. На рис. 6 показана тепловая карта, позволяющая визуальнo оценить распределения плотности антропогенной энергии, выделяемой за отопительный период зданиями с территории Новосибирска, рассчитанные по двум алгоритмам. Рис. 6 позволяет определить районы с максимальным уровнем АПТ и показать, какие районы нуждаются в проведении мероприятий, позволяющих уменьшить тепловое загрязнение городской окружающей среды.

Заключение

Детально описано создание геометрической модели городской застройки, позволяющее оценить порядок АПТ на основе данных из открытых источников. Отличительной особенностью данного исследования стало использование, помимо данных из OSM и Яндекс Карт, информации из ГИС ЖКХ, что уточнило модель городской застройки, сократив количество зданий с неизвестной высотой. Описанный метод расчета АПТ имеет высокое пространственное разрешение благодаря расчетной сетке 30 × 30 м.

Результаты, полученные с помощью представленного комплекса скриптов, применимы к любой городской агломерации, полезны при планировании и управлении энергопотреблением городского хозяйства [17], а также могут использоваться в качестве входных параметров в региональных прогностических и климатических моделях.

К недостатку алгоритмов расчета АПТ можно отнести предположение, что все здания соответствуют требуемым теплофизическим характеристикам, в действительности это не всегда так. Однако в крупных городах России ведется активная работа по присвоению классов энергоэффективности зданиям⁷, что позволит в рамках 2-го алгоритма улучшить оценку АПТ.

Список литературы

1. **Varentsov M., Konstantinov P., Baklanov A., Esau I., Miles V., Davy R.** Anthropogenic and natural drivers of a strong winter urban heat island in a typical Arctic city // *Atmos. Chem. Phys.* 2018. Vol. 18(23). P. 17573–17587. DOI: 10.5194/acp-18-17573-2018
2. **Varentsov M., Samsonov T., Demuzere M.** Impact of Urban Canopy Parameters on a Megacity's Modelled Thermal Environment // *Atmosphere*. 2020. Vol. 11(12).
3. **Hidalgo J., Masson V., Baklanov A., Pigeon G. and Gimeno L.** Advances in Urban Climate Modeling // *Annals of the New York Academy of Sciences*. 2008. Vol. 1146(1). P. 354–374. DOI: 10.1196/annals.1446.015
4. **Chen W., Zhou Y., Xie Y., Chen G., Ding K. J., & Li D.** Estimating spatial and temporal patterns of urban building anthropogenic heat using a bottom-up city building heat emission model // *Resources, Conservation and Recycling*. 2022. Vol. 177. DOI: 10.1016/j.resconrec.2021.105996
5. **Jin LX., Schubert S., Fenner D., Meier F., & Schneider C.** Integration of a Building Energy Model in an Urban Climate Model and its Application // *Boundary-Layer Meteorology*. 2021. Vol. 178(2). P. 249–281. DOI: 10.1007/s10546-020-00569-y
6. **Фролькис В. А., Гинзбург А. С., Евсиков И. А.** Оценка антропогенного потока тепла, создающего городской покрывающий слой, на основе «OpenStreetMap» // *Атмосферная*

⁷ <https://www.mos.ru/news/item/114275073/>

- радиация и динамика (МСАРД–2021): Сб. тр. Международ. симпозиума, СПб., 29 июня – 02 июля 2021 года. СПб.: ВВМ, 2021. С. 128–134.
7. **Фролькис В. А., Евсиков И. А., Мовсесова Л. В., Мотылев А. Д.** Моделирование антропогенного потока тепла и острова тепла в мегаполисах России // Атмосферная радиация и динамика (МСАРД–2023): Сб. тр. Международ.о симпозиума, СПб., 21–24 июня 2023 года. СПб.: ВВМ, 2023. С. 197–201.
 8. **Гинзбург А. С., Евсиков И. А., Фролькис В. А.** Зависимость антропогенного потока тепла от температуры воздуха (на примере Санкт-Петербурга) // Известия РАН. Сер.: Физика атмосферы и океана. 2021. Т. 57, № 5. С. 526–538. DOI: 10.31857/S0002351521050060
 9. **Фролькис В. А. Евсиков И. А. Гинзбург А. С.** Моделирование антропогенного потока тепла в течение отопительного периода в крупных городах России // Известия РАН. Сер.: Физика атмосферы и океана. 2024. (в печати)
 10. **Фролькис В. А., Евсиков И. А.** Оценка антропогенного потока тепла в течение отопительного периода в городах Российской Федерации с населением не менее полумиллиона человек // Тр. Главной геофизической обсерватории им. А. И. Воейкова. 2024. (в печати)
 11. **Левенштейн В. И.** Двоичные коды с исправлением выпадений, вставок и замещений символов // Докл. АН СССР. 1965. Т. 163 (4). С. 845–848.
 12. **Guttman A.** R-Trees: A Dynamic Index Structure for Spatial Searching // ACM SIGMOD International Conference on Management of Data. ACM. 1984. Vol. 14. P. 47–57.
 13. СП 50.13330.2012 Тепловая защита зданий. Актуализированная редакция СНиП 23-02-2003. М.: Минрегион России, 2012. 96 с.
 14. **Фролькис В. А., Евсиков И. А.** Расчет антропогенного потока тепла за период отопительного сезона в мегаполисе (на примере Санкт-Петербурга) // ENVIROMIS 2022. Томск: Томский центр научно-технической информации, 2022. С. 395–398.
 15. **Гинзбург А. С., Белова И. Н., Расплетина Н. В.** Антропогенные потоки тепла в городских агломерациях // ДАН. 2011. Т. 439, № 2. С. 256–259.
 16. **Гинзбург А. С., Докукин С. А.** Влияние теплового загрязнения атмосферы на климат города (оценки с помощью модели COSMO-CLM) // Известия РАН. Физика атмосферы и океана. 2021. Т. 57, № 1. С. 53–66. DOI 10.31857/S0002351521010053
 17. **Гинзбург А. С., Решетарь О. А., Белова И. Н.** Влияние климатических факторов на энергопотребление в отопительный сезон // Теплоэнергетика. 2016. № 9. С. 20–27. DOI: 10.1134/S0040363616080063

References

1. **Varentsov M., Konstantinov P., Baklanov A., Esau I., Miles V., Davy R.** Anthropogenic and natural drivers of a strong winter urban heat island in a typical Arctic city. *Atmos. Chem. Phys.*, 2018, vol. 18(23), pp. 17573–17587. DOI: 10.5194/acp-18-17573-2018.
2. **Varentsov M., Samsonov T., Demuzere M.** Impact of Urban Canopy Parameters on a Megacity's Modelled Thermal Environment. *Atmosphere*, 2020, vol. 11(12).
3. **Hidalgo J., Masson V., Baklanov A., Pigeon G. and Gimeno, L.** Advances in Urban Climate Modeling. *Annals of the New York Academy of Sciences*, 2008, Vol. 1146(1), pp. 354–374. DOI:10.1196/annals.1446.015.
4. **Chen W., Zhou Y., Xie Y., Chen G., Ding K. J., & Li, D.** Estimating spatial and temporal patterns of urban building anthropogenic heat using a bottom-up city building heat emission model. *Resources, Conservation and Recycling*, 2022, vol. 177. DOI: 10.1016/j.resconrec.2021.105996.
5. **Jin LX., Schubert S., Fenner D., Meier F., & Schneider, C.** Integration of a Building Energy Model in an Urban Climate Model and its Application. *Boundary-Layer Meteorology*, 2021, vol. 178(2), pp. 249–281. DOI: 10.1007/s10546-020-00569-y.
6. **Frolkis V. A., Ginsburg A. S., Evsikov I. A.** Estimation of anthropogenic heat fluxes creating the urban canopy based on “OpenStreetMap”. *International Symposium “Atmospheric Radiation*

- and Dynamics”: *Proceedings of the International Symposium*, St. Petersburg, June 29 – July 2, 2021. St. Petersburg, 2021, pp. 128–134. (in Russ.)
7. **Frolkis V. A., Evsikov I. A., Movsesova L. V., Motylev A. D.** Modeling anthropogenic heat fluxes and heat islands in Russian megacities. *Atmospheric Radiation and Dynamics: Proceedings of the International Symposium*, St. Petersburg, June 21–24, 2023. St. Petersburg., 2023, pp. 197–201. (in Russ.)
 8. **Ginsburg A. S., Evsikov I. A., Frolkis V. A.** Dependence of the anthropogenic heat flux on air temperature (using St. Petersburg as an example). *Izvestiya, Atmospheric and Oceanic Physics*, 2021, vol. 57, no. 5, pp. 461–471. DOI: 10.1134/S0001433821050066
 9. **Frolkis V. A., Evsikov I. A., Ginsburg A. S.** Modeling of anthropogenic heat fluxes during the heating season in major cities of Russia. *Izvestiya, Atmospheric and Oceanic Physics*, 2024. In print. (in Russ.)
 10. **Frolkis V. A., Evsikov I. A.** Estimation of anthropogenic heat fluxes during the heating season in cities of the Russian Federation with a population of at least half a million people. *Proceedings of the Main Geophysical Observatory named after A. I. Voeikov*, 2024. In print. (in Russ.)
 11. **Levenshtein V. I.** Binary codes with correction of deletions, insertions, and substitutions. *Rep. Acad. Sci. USSR*, 1965, vol. 163 (4), pp. 845–848. (in Russ.)
 12. **Guttman A.** R-Trees: A Dynamic Index Structure for Spatial Searching. *ACM SIGMOD International Conference on Management of Data. ACM*, 1984, vol. 14, pp. 47–57.
 13. SP 50.13330.2012 “Thermal Protection of Buildings.” Updated edition of SNiP 23-02-2003. Ministry of Regional Development of Russia. 2012, 96 p. (in Russ.)
 14. **Frolkis V. A., Evsikov I. A.** Calculation of anthropogenic heat flux during the heating season in a metropolis (using St. Petersburg as an example). *ENVIROMIS 2022*. Tomsk, 2022, pp. 395–398. (in Russ.)
 15. **Ginsburg A. S., Belova I. N., Raspeltina N. V.** Anthropogenic heat fluxes in urban agglomerations. *Doklady Earth Sciences*, 2011, vol. 439, no. 2, pp. 256–259. (in Russ.)
 16. **Ginsburg A. S., Dokukin S. A.** Influence of thermal air pollution on the urban climate (estimates using the COSMO-CLM model). *Izvestiya, Atmospheric and Oceanic Physics*, 2021, vol. 57, no. 1, pp. 47–59. DOI: 10.1134/S0001433821010059
 17. **Ginsburg A. S., Reshetar O. A., Belova I. N.** Influence of climatic factors on energy consumption during the heating season. *Thermal Engineering*, 2016, no. 9, pp. 20–27. DOI: 10.1134/S0040363616080063. (in Russ.)

Сведения об авторах

Евсиков Игорь Александрович, старший преподаватель кафедры информатики Санкт-Петербургского архитектурно-строительного университета; сотрудник Главной геофизической обсерватории им. А. И. Воейкова

Information about the Authors

Igor A. Evsikov, Senior Lecturer, Department of Computer Science, Saint Petersburg State University of Architecture and Civil Engineering, Saint Petersburg, Russian Federation

Статья поступила в редакцию 12.12.2023;
одобрена после рецензирования 08.05.2024; принята к публикации 08.05.2024

The article was submitted 12.12.2023;
approved after reviewing 08.05.2024; accepted for publication 08.05.2024

Научная статья

УДК 004.63

DOI 10.25205/1818-7900-2024-22-2-33-43

Метод цифровой подписи изображений без дополнительных файлов, с устойчивостью к JPEG-сжатию и удалению метаданных

Евгений Константинович Мазайшвили¹

Елена Юрьевна Авксентьева²

Университет ИТМО, Санкт-Петербург, Россия

¹evgenij997@yandex.ru

²avksentievaelena@rambler.ru

Аннотация

Рассмотрен метод цифровой подписи изображений, не использующий метаданные или дополнительные файлы. Цифровой подписью подписывается массив блоков 8×8 пикселей с примененным дискретным косинусным преобразованием, из которых состоит JPEG-файл. В качестве алгоритма подписи может быть использован любой известный алгоритм, например, RSA.

Полученная цифровая подпись преобразуется из бинарного вида в изображение. Метод преобразования основан на кодировании каждого нескольких бит подписи одной из 64 базисных функций дискретного косинусного преобразования. Далее базисная функция преобразуется в изображение обратным дискретным косинусным преобразованием. Полученная подпись, закодированная в виде изображения, присоединяется к исходному изображению справа, чтобы сформировать подписанное изображение.

Цифровая подпись обладает устойчивостью к JPEG-сжатию в регулируемых пределах. Для достижения устойчивости к сжатию используется следующий метод. Поскольку значения в блоках могут измениться при сжатии, используется квантизация – уменьшение точности значений. Она устроена таким образом, чтобы при квантизации сжатого и несжатого изображений значения в блоках после квантизации совпадали, что позволяет проверить валидность цифровой подписи. Квантизация также включает в себя шаг проверки четности полученного значения для исключения неверной трактовки значений при сильном сжатии.

Шаг квантизации не применяется к части изображения, содержащей цифровую подпись. Это связано с тем, что эта часть состоит из базисных функций дискретного косинусного преобразования, и даже при сильном сжатии соответствующая базисная функция все равно сохранит большой коэффициент и будет однозначно трактуема.

Ключевые слова

JPEG, дискретное косинусное преобразование, цифровая подпись, сжатие, изображение

Для цитирования

Мазайшвили Е. К., Авксентьева Е. Ю. Метод цифровой подписи изображений без дополнительных файлов, с устойчивостью к JPEG-сжатию и удалению метаданных // Вестник НГУ. Серия: Информационные технологии. 2024. Т. 22, № 2. С. 33–43. DOI 10.25205/1818-7900-2024-22-2-33-43

Digital Signature Method for Images without Additional Files, Robust to JPEG Compression and Metadata Removal

Evgeniy K. Mazaishvili¹, Elena Yu. Avksentieva²

ITMO University, Saint Petersburg, Russian Federation

¹evgenij997@yandex.ru

²avksentievaelena@rambler.ru

Abstract

The article discusses a method of digitally signing images that does not use metadata or additional files. The object for signature is an array of 8x8 pixel blocks with a discrete cosine transformation applied, which make up a JPEG file. Any known algorithm can be used as a signature algorithm, for example, RSA.

The resulting digital signature is converted from a binary form to an image. The conversion method is based on encoding every few bits of the signature of one of the 64 basic functions of the discrete cosine transform. Next, the basic function is transformed into an image by an inverse discrete cosine transformation. The resulting signature, encoded as an image, is attached to the original image on the right to form a signed image.

The digital signature is resistant to JPEG compression within adjustable limits. The following method is used to achieve compression resistance. Since the values in the blocks can change during compression, quantization is used – reducing the accuracy of the values. It is designed in such a way that when quantizing compressed and uncompressed images, the values in the blocks after quantization are becoming the same, which allows you to verify the validity of the digital signature. Quantization also includes a step of checking the parity of the received value to avoid misinterpretation of values when compression is strong.

The quantization step does not apply to the part of the image containing the digital signature. This is due to the fact that this part consists of the basic functions of the discrete cosine transform, and even with strong compression, the corresponding basic function will still retain a large coefficient and will be unambiguously interpreted.

Keywords

JPEG, Discrete cosine transform, Digital signature, Compression, Image

For citation

Mazaishvili E., Avksentieva E. Digital signature method for images without additional files, robust to JPEG compression and metadata removal. *Vestnik NSU. Series: Information Technologies*, 2024, vol. 22, no. 2, pp. 33–43 (in Russ.) DOI 10.25205/1818-7900-2024-22-1-33-43

Введение

Цифровая подпись представляет собой математическую схему, позволяющую верифицировать подлинность и целостность документа, представленного в цифровом виде [1]. Современный криптографический алгоритм цифровой подписи может быть применен к любому цифровому файлу. Большинство реализаций предполагают создание дополнительного файла таким образом, что подписанным документом будет считаться оригинальный файл + файл цифровой подписи. Многие форматы файлов, как, например, PDF, поддерживают встраивание цифровой подписи внутрь исходного файла [2]. В этом случае подписанным документом будет считаться исходный файл со встроеной цифровой подписью.

Однако для цифровых изображений в двух самых популярных форматах – JPEG и PNG – не существует схемы хранения цифровой подписи внутри исходного файла. Таким образом, существует лишь два варианта реализации цифровой подписи изображения.

1. Встраивание подписи в метаданные изображения, например, dSIG-чанки для изображений в формате PNG и метаданные EXIF для изображений в формате JPEG.

2. Сохранение отдельного файла цифровой подписи в дополнение к исходному файлу.

Обе реализации имеют недостаток в виде невозможности передачи подписанных изображений через загрузку в популярные социальные сети, мессенджеры и на популярные хостинги изображений. Причина заключается в том, что вышеперечисленные среды выполняют перекодирование изображений, удаляя метаданные. Многие из них выполняют JPEG-сжатие в дополнение к удалению метаданных.

JPEG (Joint Photographic Experts Group) – это спецификация для сжатия цифровых изображений с потерями, основанная на дискретном косинусном преобразовании – discrete cosine transform, далее DCT [3]. Использование DCT позволяет перевести изображение из пространственного представления (массива пикселей) в частотное представление (сумму косинусных функций разной частоты). Удаление самых высокочастотных функций и обратное преобразование в пространственное представление приводит к большому коэффициенту сжатия без видимого ухудшения качества изображения.

Целью данной работы является создание метода встраивания цифровой подписи в изображение формата JPEG для передачи через среды, выполняющие перекодирование изображений (рис. 1).



Рис. 1. Ожидаемая схема работы цифровой подписи
Fig. 1. Scheme of expected digital signature method

Предлагаемый метод выполняет модификацию самого изображения в виде добавления на его правую грань дополнительных пикселей, содержащих подпись. Подпись должна быть устойчива к JPEG-сжатию в определенных пределах.

1. Предлагаемый метод

Главная проблема при реализации ЭЦП для подписи изображений с устойчивостью к сжатию – это отсутствие гарантии, что после сжатия цвета пикселей изображения останутся без изменений. Соответственно, в качестве объекта подписи нельзя использовать массив всех пикселей. Также нельзя использовать приблизительные значения пикселей как объект цифровой подписи. В этом случае злоумышленник имеет возможность изменять цвета пикселей в определенных пределах, формируя неправильное изображение, не нарушая подписи.

1.1. Подпись изображения

Предлагаемый метод подписи состоит из шести шагов.

1. Вычисление массива DCT-блоков из черно-белого исходного изображения.
2. Квантизация DCT-блоков на фиксированную величину.
3. Вычисление таблицы четности для каждого DCT-блока. Эта таблица позволит цифровой подписи остаться валидной при небольших изменениях в DCT-блоке, например, при сжатии.
4. Цифровая подпись массива DCT-блоков алгоритмом RSA.
5. Кодирование полученной подписи и таблиц четности в отдельное изображение (далее по тексту: изображение-подпись, см. рис. 8, справа) методом, дающим изображению-подписи устойчивость к JPEG-сжатию.
6. Присоединение изображения-подписи к исходному изображению.

1.2. Верификация подписанного изображения

Метод верификации подписи состоит также из шести шагов.

1. Разделение подписанного изображения на исходное изображение и изображение-подпись.
2. Расшифровка данных из изображения-подписи – получение подписи в готовом для верификации виде и таблиц четности.
3. Вычисление массива DCT-блоков из исходного изображения.
4. Квантизация DCT-блоков на фиксированную величину
5. Внесение поправок в DCT-блоки из таблиц четности.
6. Проверка: совпадает ли прочитанная на шаге 2 цифровая подпись с хешем массива DCT-блоков.

Предлагаемый метод имеет ограничение в виде возможности верифицировать только черно-белые изображения.

Метод имеет некоторую схожесть со стеганографией, основанной на сокрытии данных в DCT-коэффициентах [4]. Коэффициенты анализируются, и результат этого анализа выявляет скрытую изначально информацию – является подпись валидной или нет. Однако данный метод не предназначен для сокрытия информации внутри JPEG-изображений. При взгляде на итоговое изображение будет видно, что оно состоит из двух частей – само изображение и его подпись.

Второстепенной проблемой в данном методе является то, что цифровая подпись должна сама являться изображением и быть присоединенной к подписываемому изображению. Это сделает ее подверженной JPEG-сжатию, которое нарушит подпись. Соответственно, подпись должна устойчиво читаться при сжатии изображения в определенных пределах.

2. Создание подписи

Исходное изображение может быть как сжатым без потерь, например, в формате PNG, так и сжатым с потерями в формате JPEG. Далее будет описан процесс для изображения, сжатого без потерь. Процесс сжатия файла формата JPEG отличается лишь отсутствием необходимости вычислять массив DCT-блоков, так как он уже содержится в файле.

2.1. Вычисление массива DCT-блоков из исходного изображения

Дискретное косинусное преобразование (DCT) – алгоритм, преобразующий изображение из пространства пикселей в частотное пространство, где каждому блоку 8×8 пикселей ставится в соответствие блок 8×8 коэффициентов [5]. Пример DCT-блока представлен на рис 2.

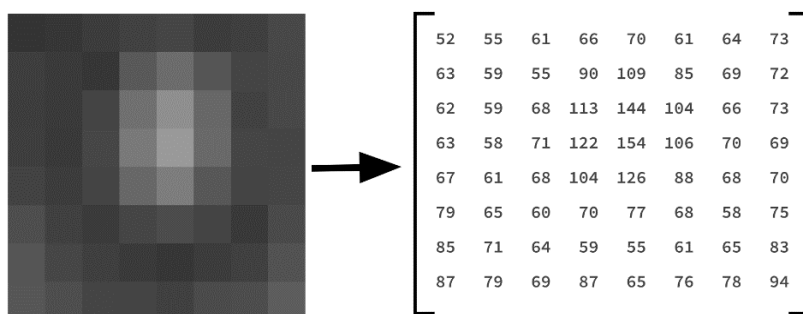
Рис. 2. Дискретное косинусное преобразование участка 8×8 пикселей

Fig. 2. Discrete cosine transform of block 8 by 8 pixels

Значения в DCT-блоке имеют диапазон от -1023 до 1023 с шагом 1. Большому значению соответствует большая амплитуда соответствующей косинусной функции и большая яркость определенных пикселей, зависящих от функции.

С точностью до ошибок округления DCT является алгоритмом преобразования без потерь – таблицу можно преобразовать обратно в исходный массив пикселей.

На этом шаге метода исходное изображение преобразуется в массив DCT-блоков.

2.2. Квантизация DCT-блоков на заданную величину

Квантизация является вторым шагом метода. Этот шаг делает значения в DCT-блоках более грубыми, выполняя квантизацию с определенным шагом. Квантизация представляет собой преобразование, где каждое значение внутри DCT-блока преобразуется по формуле:

$$value_i = \text{floor}\left(\frac{value_i}{Q_i}\right) * Q_i, \quad (1)$$

где Q_i – значение шага квантизации из таблицы шагов квантизации.

В методе используется переменное значение шага в зависимости от места значения в таблице: шаг увеличивается к нижнему правому углу таблицы. Это сделано по причине того, что алгоритм сжатия JPEG вносит более сильные искажения в значения ближе к нижнему правому углу DCT-блока, где представлены более мелкие детали изображения. Поскольку мы хотим точнее сохранить более явные детали, то для значений ближе к левому верхнему углу таблицы используются меньшие шаги квантизации. Значение в левом верхнем углу имеет больший шаг, так как его изменение приводит лишь к изменению общей яркости блока в итоговом изображении. Используемые значения шага обозначены на рис. 3.

40	25	25	25	25	40	40	80
25	25	25	25	40	40	80	80
25	25	25	40	40	80	80	90
25	25	40	40	80	80	90	90
25	40	40	80	80	90	90	90
40	40	80	80	90	90	90	90
40	80	80	90	90	90	90	90
80	80	90	90	90	90	90	90

Рис. 3. Значение шага квантизации в зависимости от места значения в блоке

Fig. 3. Quantization step value for different places in a block

Смысл квантизации можно представить как уход от точных значений в блоке и работу с диапазонами значений. Например, значение 529 в блоке при 60%-м сжатии исказится до 521, но при шаге квантизации 10, они оба будут попадать в интервал 520–530. Аналогично при шаге 50: оба значения будут попадать в интервал 500–550. Нижнее значение диапазона является результатом квантизации с определенным шагом, как на рис. 4.

До квантизации:

[	-98	-227	118	-45	72	-68	48	-24]
[	-197	-165	-42	41	-2	-2	-4	-4]
[	96	81	3	-24	-15	4	-1	6]
[	-28	-13	34	37	11	5	4	0]
[	-8	-19	-26	-13	-1	-3	-6	2]
[	11	14	4	-7	-12	0	5	0]
[	-5	-2	10	19	15	4	-1	2]
[	2	-4	-11	-12	-9	0	3	2]]

После квантизации:

[	-120	-250	100	-50	50	-80	40	-80]
[	-200	-175	-50	25	-40	-40	-80	-80]
[	75	75	0	-40	-40	0	-80	0]
[	-50	-25	0	0	0	0	0	0]
[	-25	-40	-40	-80	-80	-90	-90	0]
[	0	0	0	-80	-90	0	0	0]
[	-40	-80	0	0	0	0	-90	0]
[	0	-80	-90	-90	-90	0	0	0]]

Рис. 4. Значения в одном из DCT-блоков до и после квантизации
(используются шаги квантизации с рис. 3)

Fig. 4. Values in one of DCT blocks before and after quantization
(using quantization steps from fig. 3)

Квантизованные блоки не переводятся обратно в пиксели и не сохраняются в итоговом подписанном изображении. Они нужны только для этапа вычисления их хеша. Именно хеш от всех квантизованных блоков является объектом цифровой подписи.

Если бы в качестве объекта подписи вместо квантизованных, т. е. приблизительных значений DCT-блоков, использовались приблизительные значения пикселей с заданной погрешностью, злоумышленник мог бы изменять их в пределах этой погрешности, вырисовывая нужные детали на изображении. В данной работе используются приблизительные значения, но не пикселей, а DCT-блоков, изменение которых злоумышленником приведет лишь к изменению малозаметных деталей на изображении.

2.3. Вычисление таблицы четности для каждого DCT-блока

При квантизации блока может оказаться, что некоторые значения находятся на границе интервала квантизации или близко к границе. Например, для квантизации с шагом 10 значение 50 будет находиться на границе интервалов 40–50 и 50–60. В этом случае даже небольшое искажение при сжатии может сместить значение в неправильный интервал, делая подпись невалидной.

Для того чтобы избежать этой проблемы, каждому из интервалов, определяемых шагом квантизации, присваивается четность как на рис. 5.

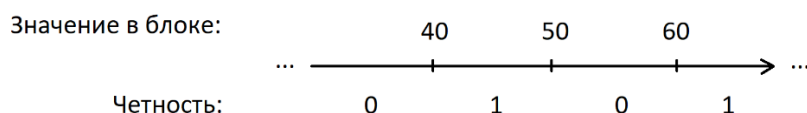


Рис. 5. Вычисление четности для каждого значения в квантованном блоке

Fig. 5. Parity calculation for every value in quantized block

Вместе с цифровой подписью сохраняется четность каждого значения для каждого блока в виде одного бита. Например, при шаге квантизации 10 значение 51 находится в нечетном интервале 50–60, а значение 49 – в четном интервале 40–50.

Таким образом, к каждому квантованному блоку рассчитывается соответствующая ему таблица четности. Эта таблица позволяет на этапе расшифровки делать поправки в значения блока, если значение при сжатии случайно преодолест границу интервала. Пример таблицы четности приведен на рис. 6.

[1 0 0 0 0 0 1 1]
[0 1 0 1 1 1 1 1]
[1 1 0 1 1 0 1 0]
[0 1 0 0 0 0 0 0]
[1 1 1 1 1 1 1 0]
[0 0 0 1 1 0 0 0]
[1 1 0 0 0 0 1 0]
[0 1 1 1 1 0 0 0]

Рис. 6. Таблица четности для значений из нижнего блока на рис. 4

Fig. 6. Parity table for values in bottom block from fig. 4.

Таблица занимает 64 бита на каждый участок изображения 8×8 пикселей.

2.4. Цифровая подпись хеша массива DCT-блоков алгоритмом RSA

Массив всех квантованных блоков + массив таблиц четности хешируются алгоритмом SHA-256, и получаемый хеш подписывается закрытым ключом согласно алгоритму RSA.

В результате четвертого шага получается всего два артефакта, необходимых для верификации подписи: подписанный хеш размером 2048 бит (при использовании RSA-2048) и массив таблиц четности размером 64 бит $\times$ количество блоков 8×8 пикселей в изображении.

2.5. Кодирование полученной подписи и таблиц четности в изображение-подпись

Подписанный хеш и таблицы четности должны храниться непосредственно в изображении, и для этого они должны быть устойчивыми к JPEG-сжатию.

Для того чтобы достичь устойчивости к сжатию, была исследована одна из особенностей работы дискретного косинусного преобразования. DCT-блок представляет собой блок 8×8 значений, от -1023 до 1023 . Блок, состоящий только из максимальных и минимальных значений, будет демонстрировать устойчивость к сжатию, так как сжатие такого блока приведет к тому, что в значения будут внесены лишь незначительные искажения. Например, значение -1023

может стать значением -990 после сжатия. Однако близость его к минимальному значению позволит однозначно считать его минимальным значением. Таким образом, изображение, составленное из DCT-блоков с максимальным и минимальным значением, будет устойчиво к очень высокому уровню сжатия.

В данной работе используется три значения: -1023 , 0 и 1023 . Таким образом, в одном значении DCT-блока в изображении-подписи хранится 2,5 бит полезной информации. Пример такого блока представлен на рис. 7.



Рис. 7. Искусственно созданный DCT-блок (слева) и соответствующий ему блок пикселей (справа), который будет включен в финальное изображение-подпись

Fig. 7. Artificially crafted DCT block (left) and corresponding pixel block (right), that will be included into the final image-signature

Таким образом, кодируя $64 \times 2,5$ бита в картинку 8×8 пикселей, все таблицы четности уместятся в зону, занимающую не более $1/3$ изображения.

2.6. Присоединение изображения-подписи к исходному изображению

После кодирования изображение-подпись, содержащее подписанный хеш и таблицы четности, присоединяется к исходному изображению (см. рис. 8).



Рис. 8. Слева направо: исходное изображение, изображение с цифровой подписью

Fig. 8. Left to right: original image, image with digital signature

Размер области с подписью всегда составляет фиксированную часть изображения, так как количество таблиц четности линейно зависит от размера изображения, а подписанный хеш имеет фиксированную длину.

На этом шаге процесс подписи завершается.

3. Верификация ЭЦП

Входные данные для этапа верификации – изображение с цифровой подписью. Изображение не содержит метаданных из исходного файла и, вероятно, было сжато алгоритмом JPEG.

3.1. Разделение подписанного изображения на исходное и изображение-подпись

Поскольку размер цифровой подписи известен, разделение изображений на исходное и подпись является тривиальным.

3.2. Расшифровка данных из изображения-подписи – получение подписи в готовом для верификации виде и таблиц четности

Для расшифровки изображения-подписи каждый участок 8×8 пикселей переводится в DCT-блок, значения которого округляются до ближайшего: 1023; 0 или –1023.

Таким образом, восстанавливаются подписанный хеш и таблицы четности.

3.3. Вычисление массива DCT-блоков из исходного изображения.

Шаг аналогичен первому шагу при подписывании изображения.

Вычисление DCT не требуется, если файл уже находится в формате JPEG, в таком случае достаточно лишь прочитать готовые DCT-блоки из файла.

3.4. Квантизация DCT-блоков на фиксированную величину

Шаг квантизации полностью повторяет таковой при подписи изображения. Цель этого шага – получить такие же квантизованные блоки, какие были в изображении на момент подписи, чтобы подпись была валидной.

3.5. Внесение поправок в DCT-блоки из таблиц четности

На этом шаге происходит расчет таблиц четности для квантизованных блоков, аналогично шагу подписи. Однако, в отличие от аналогичного шага при подписи, при верификации происходит сравнение таблиц, и уже по результатам сравнения происходит корректировка значений блоков: те значения, которые могли выйти за границу своего интервала, поскольку находились близко к ней, возвращаются в свой интервал.

Если рассчитанная четность значения в блоке не совпадает с соответствующей четностью, записанной в таблице четности в подписи, выполняется коррекция.

Алгоритм данной коррекции описывается формулой:

$$v_i = \begin{cases} |\hat{v}_i - v_i| > \frac{Q_i}{2}: & v_i + \frac{Q_i}{2} \\ \text{else:} & v_i - \frac{Q_i}{2} \end{cases}, \quad (2)$$

где Q_i – таблица шагов квантизации, v – квантизованное значение в блоке, $\hat{v}$ – значение в блоке до квантизации.

Формула описывает следующее преобразование. Если значение ближе к нижней границе интервала, то оно считается ошибочно попавшим в него из нижестоящего интервала. Интервал в квантизованной таблице меняется на нижестоящий. Если же значение ближе к верхней границе интервала, то оно считается ошибочно попавшим в него из вышестоящего интервала. Интервал в квантизованной таблице меняется на вышестоящий.

3.6. Проверка – совпадает ли прочитанная на шаге 2 цифровая подпись с рассчитанным хешем

Последним шагом происходит хеширование массива DCT-таблиц и таблиц четности из подписи по той же схеме, что и при подписывании изображения, после чего полученный хеш сравнивается с тем, что находится в цифровой подписи. Совпадение хешей указывает на то, что подпись верна. Несовпадение указывает либо на слишком сильное сжатие, либо на намеренное искажение изображения.

Заключение

В работе рассмотрен метод цифровой подписи изображений в формате JPEG, а также в любом формате сжатия без потерь, например, PNG. Метод обладает устойчивостью к JPEG-сжатию, благодаря использованию в качестве объекта подписи квантизованных DCT-блоков. Устойчивость к сжатию достигается за счет хранения данных цифровой подписи в виде изображения, состоящего из базовых функций DCT максимальной амплитуды.

Метод может использоваться как упрощенная схема верификации изображений, не требующая ничего, кроме самого контента изображения.

В качестве перспективы развития метода можно увеличить плотность данных в изображении-подписи, чтобы оно занимало меньшую часть основного изображения.

Список литературы

1. **Bellare M., Goldwasser S.** Lecture Notes on Cryptography, 2008, с. 168
2. PDF (Portable Document Format), версия 1.7, ISO 32000-1, 2008 // Библиотека Конгресса. URL: <https://www.loc.gov/preservation/digital/formats/fdd/fdd000277.shtml> (дата обращения: 11.01.2024).
3. **Buchanan W. J.** DCT (Дискретное косинусное преобразование) // Теоретические и практические аспекты криптографии. URL: <https://asecuritysite.com/comms/dct2> (дата обращения: 24.12.2023).
4. **Abdelhamid A. A., Mona F. M., Mursi A., Alsammak A. K.** Data hiding inside JPEG images with high resistance to steganalysis using a novel technique: DCT-M3 // Инженерный журнал Ain Shams. 2018.

5. **Griffin J.** Полное руководство JPEG, включая сжатие и кодирование // Веб-мастер блог. URL: <https://www.thewebmaster.com/jpeg-definitive-guide/> (дата обращения: 10.12.2023).

References

1. **Bellare M., Goldwasser S.** *Lecture Notes on Cryptography*, 2008, p. 168
2. PDF (Portable Document Format), version 1.7, ISO 32000-1, 2008. *Library of Congress*. URL: <https://www.loc.gov/preservation/digital/formats/fdd/fdd000277.shtml> (accessed: 11.01.2024).
3. **Buchanan W. J.** DCT (Discrete Cosine Transform). *Theoretical and practical aspects of cryptography*. URL: <https://asecuritysite.com/comms/dct2> (accessed: 24.12.2023).
4. **Abdelhamid A. A., Mona F. M., Mursi A., Alsammak A. K.** Data hiding inside JPEG images with high resistance to steganalysis using a novel technique: DCT-M3. *Ain Shams Engineering Journal*, 2018.
5. **Griffin J.** The Ultimate Guide to JPEG Including JPEG Compression & Encoding. URL: <https://www.thewebmaster.com/jpeg-definitive-guide/> (accessed: 10.12.2023)

Сведения об авторах

Мазайшвили Евгений Константинович, аспирант

Авксентьева Елена Юрьевна, PhD

Information about the Authors

Evgeniy K. Mazaishvili, PhD Student

Elena Yu. Avksentieva, PhD

Статья поступила в редакцию 10.04.2024;

одобрена после рецензирования 16.04.2024; принята к публикации 16.04.2024

The article was submitted 10.04.2024;

approved after reviewing 16.04.2024; accepted for publication 16.04.2024

Научная статья

УДК 004.89

DOI 10.25205/1818-7900-2024-22-2-44-56

Разработка метода для оптимизации процесса получения значений числовых характеристик компонентов нефтяного сырья из графических данных хроматограмм

**Петр Андреевич Пылов¹, Роман Вячеславович Майтак¹
Даниил Евгеньевич Копылов²**

¹Кузбасский государственный технический университет им. Т. Ф. Горбачева
Кемерово, Россия

²Иркутский государственный университет
Иркутск, Россия

pylovpa@kuzstu.ru; <https://orcid.org/0009-0009-4353-6925>
maytak.roman@mail.ru; <https://orcid.org/0009-0009-4353-6925>
it-daniil@yandex.ru; <https://orcid.org/0009-0005-0554-8934>

Аннотация

Разработан вариант алгоритма (и метода на его основе) для выполнения процедуры автоматизированного восстановления числовых значений графически представленной функции сигнала хроматографа, исследующей компонентный состав проб нефтяного сырья. Проблема, на решение которой направлен разработанный метод, состоит в слабой приспособленности хроматографов к нефтяной промышленности: нефть является природным сырьем, которое не является химически чистым, поэтому в рамках хроматографического исследования происходит фиксация не всех числовых характеристик компонентов, входящих в состав исследуемого образца. Значения незафиксированных координат углеводородных соединений регистрируются специалистами нефтехимических лабораторий вручную, что существенно усложняет процесс анализа. Разработанный метод в качестве входных данных принимает полученные в лаборатории изображения хроматограмм нефти, представленные в оригинальной черно-белой цветовой гамме. Выходными данными метода является массив числовых значений координат, восстановленных с шагом в один пиксель. Величина погрешности при восстановлении значений методом значительно меньше порога, установленного нефтехимической лабораторией. Кроме автоматизации обозначенной задачи, массив полученных значений координат был векторизован с целью применения вектора в качестве входных данных в модели трансформера (вместо изображений хроматограмм) для решения задачи прогнозирования перераспределения углеводородных компонентов тяжелой нефти под воздействием катализаторов. Результатом замены типа представления входной информации стало многократное снижение времени, требуемого для получения прогноза, а также времени обучения, при этом снизилась величина усредненной ошибки предсказания.

Ключевые слова

анализ данных, предварительная обработка графических данных, обработка хроматограмм тяжелой нефти, восстановление числовых координат углеводородных компонентов

Для цитирования

Пылов П. А., Майтак Р. В., Копылов Д. Е. Разработка метода для оптимизации процесса получения значений числовых характеристик компонентов нефтяного сырья из графических данных хроматограмм // Вестник НГУ. Серия: Информационные технологии. 2024. Т. 22, № 2. С. 44–56. DOI 10.25205/1818-7900-2024-22-2-44-56

© Пылов П. А., Майтак Р. В., Копылов Д. Е., 2024

Development of a Method for Optimization of the Process of Extraction of Values of Numerical Characteristics of Oil Feedstock Components from Graphical Data of Chromatographic Analysis

Petr A. Pylov¹, Roman V. Maitak¹,
Daniil E. Kopylov²

¹Gorbachev Kuzbass State Technical University, Kemerovo, Russian Federation

²Irkutsk State University, Irkutsk, Russian Federation

pylovpa@kuzstu.ru; <https://orcid.org/0009-0009-4353-6925>

maytak.roman@mail.ru; <https://orcid.org/0009-0009-4353-6925>

it-daniil@yandex.ru; <https://orcid.org/0009-0005-0554-8934>

Abstract

A variant of the algorithm has been developed to perform the procedure of automated recovery of numerical values of graphically represented chromatograph signal function, studying the component composition of heavy oil feedstock samples. The problem, which the developed method aims to solve, consists in the poor adaptation of chromatographs to the oil industry: oil is a natural raw material, which is not chemically pure, therefore not all numerical characteristics of the components contained in the investigated sample are fixed within the chromatographic study. In the current configuration of the method, the values from the chromatogram are recorded manually. The developed method takes as input data the images of oil chromatograms obtained in the laboratory, presented in the original black and white colour scheme. The output data of the method is an array of numerical values of coordinates reconstructed with a step of one pixel. The size of the error in the reconstruction of the values by the method is much smaller than the threshold set by the petrochemical laboratory. In addition to automating the indicated task, the array of obtained coordinate values was vectorized in order to use the vector as input data in the Transformer model to solve the problem of predicting the redistribution of hydrocarbon components of heavy oil under the influence of catalysts. As a result of the change in input data representation, the time required to obtain a prediction and the training time were reduced by a multiple, while the value of the average prediction error decreased.

Keywords

data analysis, preliminary processing of graphical data, processing of heavy oil chromatograms, recovery of numerical coordinates of hydrocarbon components

For citation

Pylov P. A., Maitak R. V., Kopylov D. E., Evsikov I. A. Development of a method for optimization of the process of extraction of values of numerical characteristics of oil feedstock components from graphical data of chromatographic analysis. Vestnik NSU. Series: Information Technologies, 2024, vol. 22, no. 2, pp. 44–56 (in Russ.) DOI 10.25205/1818-7900-2024-22-2-44-56

Введение

Нефть, как энергетическое сырье, активно используется человечеством на протяжении длительного времени. За это время в предметной области нефтяного дела устоялась классификация всего природного сырья на два условных класса: легкую и тяжелую нефть [1]. Технологически добыча и переработка легкой нефти за несколько последних десятилетий была настолько хорошо налажена промышленностью, что серьезных трудностей в этой подобласти не возникает [2]. По сравнению с легкой, тяжелая нефть является гораздо более сложнокомпонентной смесью углеводородов, поэтому работать с ней значительно сложнее: затрудняется как процесс добычи, так и транспортировки (с дальнейшими ограничениями переработки, вызванными повышенной плотностью и вязкостью сырья) [3]. Для того чтобы нивелировать большинство сложностей на всех этапах жизненного цикла (от добычи сырья тяжелой нефти до превращения его в конечный энергетически ценный продукт), специалисты нефтеперерабатывающей области нашли один из способов решения проблемы в подходе, подразумевающим применение катализаторов в облагораживающей химической реакции, которая позволяет безвозвратно снизить вязкость и плотность тяжелой нефти уже на этапе ее нахождения во внутрипластовом пространстве [4].

Сложностью, которую порождает данный подход, является объемное перераспределение различных углеводородных компонентов между собой. Этот фактор в дальнейшем напрямую влияет на объемный выход конечного продукта переработки нефти и, как следствие, – на уровень коммерциализации нефтеперерабатывающего предприятия в целом.

Наиболее прецизионный способ для выбора оптимального катализатора основан на проведение хроматографического исследования [5], которое позволяет получить графическое представление компонентного состава тяжелой нефти – хроматограмму (рис. 1).

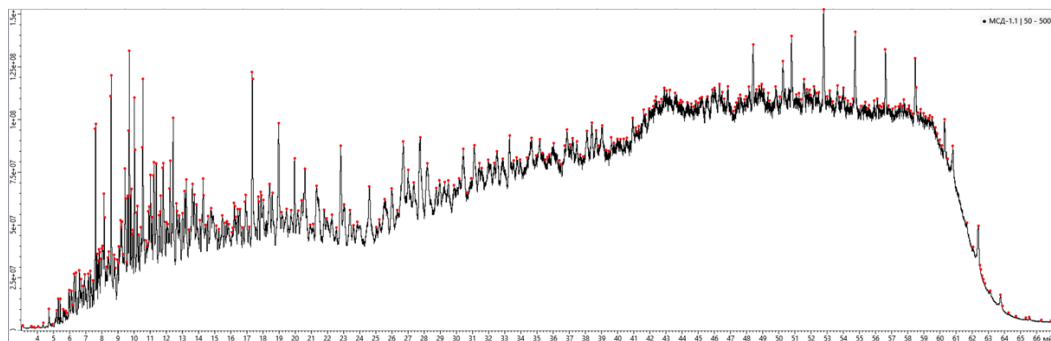


Рис. 1. Пример хроматограммы образца тяжелой нефти
Fig. 1. Example chromatogram of a heavy oil sample

Пики на хроматограмме (красные точки на рис. 1) фиксируют присутствие обособленного углеводородного компонента в исследуемом образце нефти, а значения координат пиков (время выхода по оси абсцисс и энергия выхода по оси ординат) позволяют однозначно идентифицировать название углеводородного соединения и его объемное содержание в образце. Таким образом, после проведения хроматографического исследования образца сырой тяжелой нефти в дальнейшем можно будет выполнить анализ для копии этого же образца с введенным в него катализатором, определив тем самым объемное перераспределение углеводородов на основе сопоставления результатов двух лабораторных исследований.

Однако фиксация соответствующих числовых координатных значений пиков с хроматограммы происходит автоматически только в том случае, когда органическое соединение уже есть в базе данных хроматографа. Поскольку базы данных являются собственностью производителей хроматографов (устройств, выполняющих хроматографический анализ) и не могут быть расширены в лабораториях [6], а сами устройства ориентированы на фармацевтическую промышленность, в которой специалисты оперируют химически чистыми соединениями, то сотрудники нефтехимических лабораторий для большинства пиков вынуждены вручную выполнять фиксацию значений через аппроксимацию числовых данных с графика хроматограммы [7].

В работе [8] задача прогнозирования перераспределения углеводородных компонентов (при вводе катализатора) была автоматизирована моделью мультимодального трансформера. Ключом к решению стал анализ всего изображения хроматограммы путем загрузки изображения в модель трансформера. Несмотря на то что такой подход с высокой точностью решает поставленную задачу, стоит заметить, что процесс является очень ресурсозатратным, так как со всего изображения (рис. 1) вычленяется для анализа график, фиксирующий сигнал хроматографа, при этом большая площадь анализируемого изображения является белой заливкой, которая не несет в себе смысловой информации.

Исходя из этого, было решено провести дополнительное исследование, посвященное оптимизации процесса обработки входных данных с целью упрощения пайплайна разработанного прикладного решения на базе интеллектуальной модели.

Исходные данные

В качестве набора исходных данных выступают лабораторные хроматограммы тяжелой нефти, полученные с помощью хроматографов «Хроматэк» серии «Кристалл» на основе образцов из Ашальчинского, Мордово-Кармальского и Пионерского месторождений высоковязкой нефти. Территориально все обозначенные месторождения находятся в Республике Татарстан. Отметим, что данные из Ашальчинского и Мордово-Кармальского месторождения аналогичны тем, которые были использованы в работе [8].

Каждая хроматограмма является файлом изображения размером 1302×558 пикселей, представленном в одноканальном режиме, имеющем два уровня градации серого цвета: черный и белый. Размер выборки данных составил 220 484 хроматограмм, из которых 210 000 принадлежат к Ашальчинскому месторождению, 2310 – Мордово-Кармальскому, 8174 – Пионерскому. На отличающихся месторождениях присутствуют свои особенности нефтяного сырья, которые будут проявляться в изменяющемся представлении графика на хроматограмме [9].

Отметим, что на всех представленных графических данных существует общая черта, которая должна быть учтена в процессе построения метода для фиксации числовых значений координат: в правом верхнем углу каждой хроматограммы присутствует отметка хроматографа, которая регистрирует порядковый номер исследования (рис. 2).

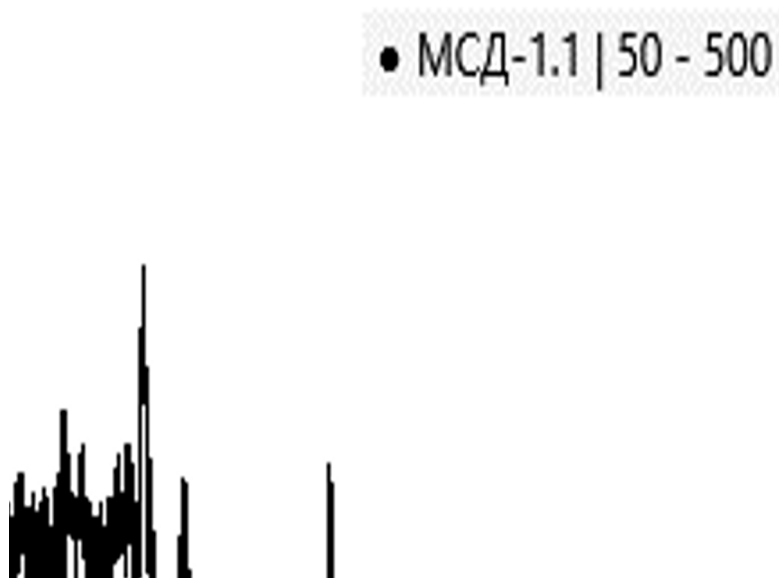


Рис. 2. Регистрационная отметка номера исследования, присваиваемая хроматографом каждой хроматограмме
Fig. 2. Identifies the run number assigned to each chromatogram by the chromatographer

Регистрационная отметка устанавливает однозначное соответствие между конкретным лабораторным исследованием пробы нефти и его хроматограммой, поэтому в рамках решения задачи фиксации числовых значений координат она не содержит в себе полезной информации.

Методы

Для решения задачи восстановления значений временного ряда (зависимости времени выхода углеводородного компонента от его энергии выхода) на основе его графического

представления была учтена важная особенность изображений, представленных исключительно в черном и белом цвете. Реализованный программный метод позволяет восстановить временной ряд благодаря последовательному решению нескольких операций над хроматограммой.

1. Ретуширование белыми пикселями прямоугольного участка регистрационной отметки хроматографа, находящегося в правом верхнем углу хроматограммы.

2. Ограничение зоны исследования метода только на область графика (рис. 3). Обоснованием этому является постоянство значений единиц измерения, находящихся на осях координат (так как характеристики всех проб нефти изменяются в одном диапазоне значений, а изменяется только форма графика).

3. Выполнение условного «разрезания» графика на вертикальные полосы шириной в 1 пиксель, начиная от начала системы координат.

4. Поиск интенсивности пикселя, соответствующей черному цвету на каждой из полученных вертикальных полос.

5. Фиксирование значения удаленности (в пикселях) обнаруженной «черной точки» относительно начала отсчета, которому соответствует верхний левый угол. Величина удаленности рассчитывается от верха графика, а затем инвертируется для получения значения по оси ординат. В том случае, если черных пикселей на одной вертикальной полосе оказывается несколько, то для «серии пикселей» фиксируется наибольшее значение (случай представлен на рис. 4) по оси ординат (по такому принципу функционирует и сам хроматограф, так как для отдельных компонентов нефти энергия выхода возрастает не мгновенно, а инерциально, поэтому из исследования графика может сложиться ошибочное мнение, что одному дискретному значению времени соответствует несколько значений энергии: истинным значением будет максимальная величина зарегистрированной энергии). Именно по этой причине величина удаленности рассчитывается от верха графика, так как такой способ позволяет существенно экономить временные затраты на фиксацию значения энергии в вертикальной полосе.

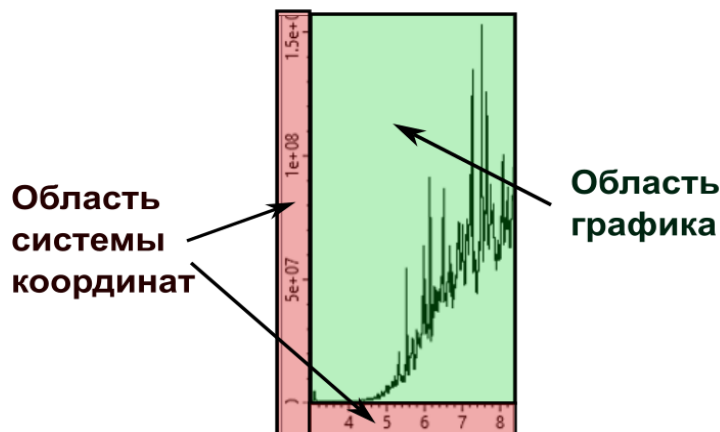


Рис. 3. Графическое представление области графика, исследуемой методом
Fig. 3. Graphical representation of the area of the graph under study by the method

Предлагаемый метод позволяет выполнить конвертацию данных хроматограмм из графической формы представления информации в числовую, которая теперь формализована в виде пар значений координат. Для расчета величины погрешности метода была использована метрика среднеквадратического отклонения, которая позволяет определить разницу между известными значениями координат углеводородных компонентов, зафиксированными хроматографом, и восстановленными при помощи метода.

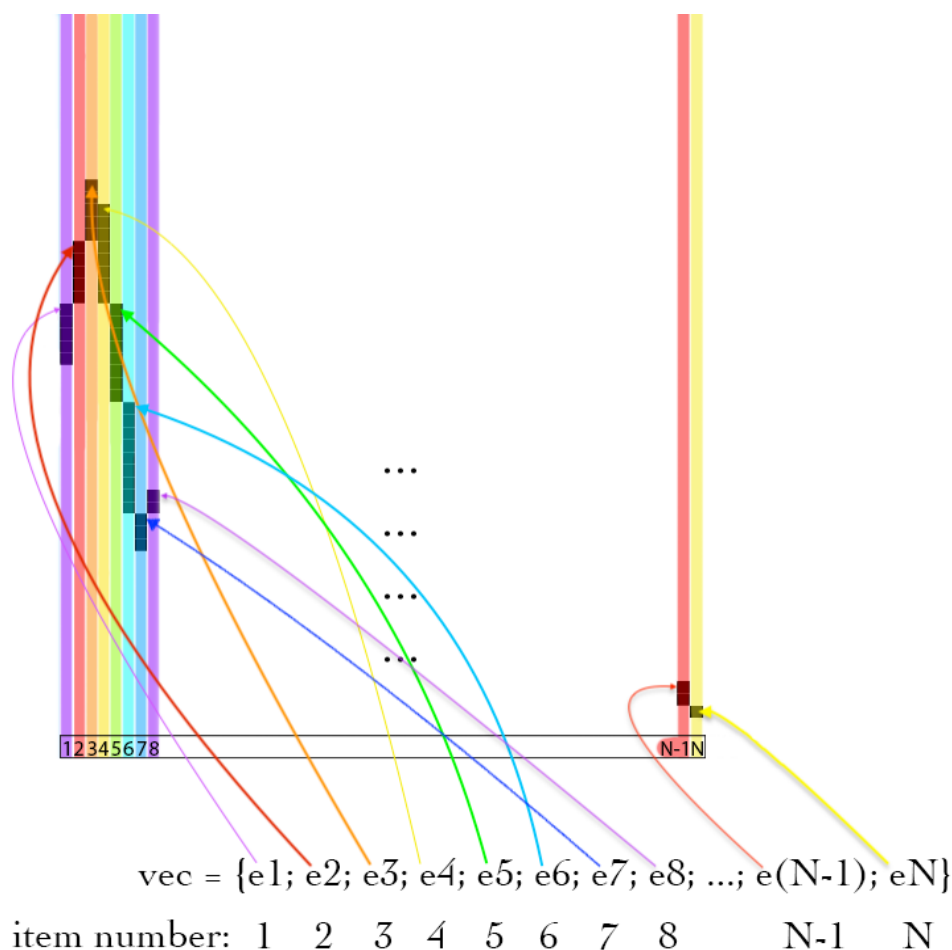


Рис. 4. Графическая схема составления вектора из пар координат восстановленных точек
 Fig. 4. Graphical scheme of vector construction from coordinate pairs of reconstructed points

Отметим, что восстановление числовых значений координат углеводородных компонентов на основе графического представления функции сигнала хроматографа является вторичной задачей. Основная задача, для решения которой был реализован метод, – упрощение пайплайна уже существующей интеллектуальной модели [8], поэтому для оценки эффективности от интеграции метода в конфигурацию модели были выделены следующие критерии:

- период времени, требуемый для обучения интеллектуальной модели;
- период времени, необходимый для предсказания выходных параметров углеводородных компонентов [8] (фактически является общим временем, требуемым для получения прогноза от модели);
- величина ошибки интеллектуальной модели при выполнении предсказаний в рамках решения прикладной задачи.

Поскольку исходные данные теперь представляются в формате числовых координат, возникла необходимость модификации пайплайна базовой модели [8].

1. Модель трансформера теперь не должна быть мультимодальной, так как отсутствует необходимость обработки разнородных данных (изображение и табличная информация); вся необходимая информация представляется в формате чисел.

2. Входные данные, представленные в виде пар координат (значения по двум осям для каждой восстановленной точки), можно преобразовать в вектор значений (рис. 4), в кото-

ром номер элемента будет характеризовать время выхода углеводородного компонента, а само значение – энергию. Так как вектор является обучающими данными для модели трансформера, было решено выполнить нормировку его значений, потому что в таком случае функция активации softmax позволит сформировать более качественную обобщающую способность модели глубокого обучения.

3. В модели трансформера остается только один энкодер, который используется в качестве входа и на который подаются векторизованные представления данных. Для повышения обобщающей способности предварительно все значения вектора (восстановленные координаты по оси ординат) нормализуются.

Результаты

В процессе применения метода для решения задачи конвертации данных из изображений в массив числовых значений координат были получены визуально идентичные графики: на рис. 5 представлено изображение оригинальной хроматограммы, а на рис. 6 – график, построенный на основе восстановленных числовых координат. Отметим, что график, изображенный на основе восстановленных значений, получен путем соединения соседних точек прямыми линиями.

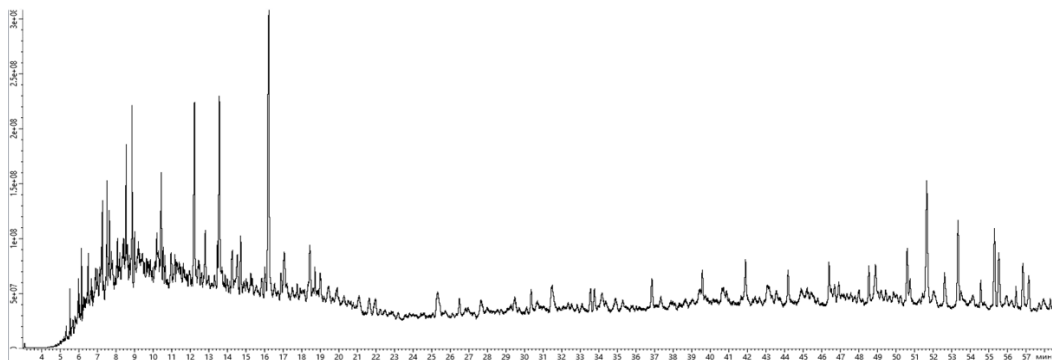


Рис. 5. Оригинальное изображение хроматограммы для пробы нефти
(с заретушированной отметкой хроматографа)

Fig. 5. Original image of the chromatogram for the oil sample
(with the mark of the chromatograph masked out)

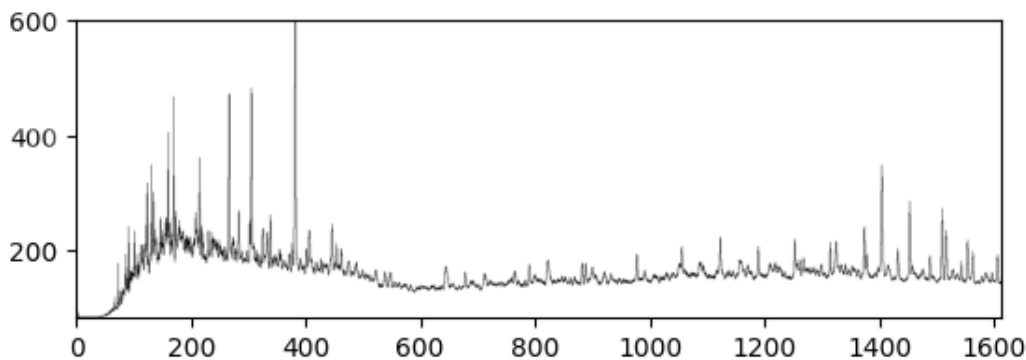


Рис. 6. График, построенный на основе восстановленных значений координат

Fig. 6. Graph based on the reconstructed coordinate values

В табл. 1 представлены числовые значения координат локальных максимумов точек хроматограммы, соответствующей рис. 5. Во втором столбце находятся значения, зарегистрированные хроматографом, в третьем – восстановленные с помощью метода (на рис. 6 значения по осям координат представлены в пикселях; для единства представления в табл. 1 шкала значений приведена к оригинальным единицам измерения).

Таблица 1

Сравнение полноты зарегистрированных значений координат точек
на хроматограмме (представлены только локальные максимумы)
на основе хроматографа и разработанного метода

Table 1

Comparison of the completeness of the recorded values of point coordinates
on the chromatogram (only local maxima are displayed)
based on the chromatograph and the developed method

Порядковый номер	Зафиксированные значения хроматографа		Восстановленные значения		Модуль разницы значений	
	Время выхода	Энергия выхода	Время выхода	Энергия выхода	Время выхода	Энергия выхода
1	2	3	4	5	6	7
1	3,072	0,0507	3,072	0,0495	0,0002	0,0012
2	4,770	0,0244	4,771	0,0254	0,0011	0,0010
3	4,878	0,0454	4,876	0,0460	0,0016	0,0006
4	4,996	0,0804	4,995	0,0810	0,0006	0,0006
5	5,318	0,2055	5,317	0,2045	0,0012	0,0010
6	5,526	0,5466	5,526	0,5480	0,0000	0,0014
7	5,691	0,2837	5,693	0,2827	0,0017	0,0010
8	5,798	0,3076	5,798	0,3084	0,0002	0,0008
9	5,949	0,4439	5,950	0,4440	0,0009	0,0001
10	5,995	0,6377	5,995	0,6382	0,0005	0,0005
11	6,028	0,5072	6,029	0,5072	0,0014	0,0000
12	6,146	0,9121	6,145	0,9120	0,0014	0,0001
13	6,264	0,7644	6,263	0,7632	0,0007	0,0012
14	6,508	0,7779	6,508	0,7789	0,0002	0,0010
15	6,669	0,8835	6,668	0,8832	0,0007	0,0003
16	6,920	0,7346	6,920	0,7349	0,0003	0,0003
17	7,020	1,2453	7,019	1,2463	0,0008	0,0010
18	7,210	1,3642	7,208	1,3634	0,0017	0,0008
19	7,275	0,9231	7,276	0,9238	0,0011	0,0007
20	7,518	1,5379	7,519	1,5382	0,0012	0,0003
21	7,644	1,3382	7,643	1,3396	0,0006	0,0014
22	8,081	1,0128	8,080	1,0136	0,0006	0,0008

Продолжение табл. 1

1	2	3	4	5	6	7
23	8,217	0,8746	8,216	0,8743	0,0008	0,0003
24	8,550	1,8678	8,549	1,8665	0,0007	0,0013
25	8,865	2,2132	8,867	2,2129	0,0016	0,0003
26	9,005	1,0749	9,004	1,0755	0,0010	0,0006
27	9,220	0,9872	9,221	0,9876	0,0012	0,0004
28	10,202	1,7535	10,202	1,7522	0,0004	0,0013
29	10,442	1,6286	10,441	1,6284	0,0007	0,0002
30	10,542	0,9479	10,541	0,9481	0,0009	0,0002
31	10,639	0,8754	10,637	0,8769	0,0015	0,0014
32	10,954	0,8985	10,956	0,8981	0,0017	0,0004
33	11,184	0,8721	11,185	0,8719	0,0008	0,0002
34	12,216	0,8236	12,216	0,8249	0,0001	0,0013
35	12,810	1,1026	12,810	1,1024	0,0005	0,0002
36	13,466	2,3214	13,466	2,3206	0,0001	0,0008
37	13,566	2,2675	13,565	2,2680	0,0009	0,0005
38	14,269	0,9342	14,267	0,9348	0,0017	0,0006
39	14,541	0,8779	14,539	0,8784	0,0015	0,0005
40	14,717	1,0573	14,718	1,0565	0,0011	0,0008
41	16,017	3,0878	16,017	3,0877	0,0004	0,0001
42	16,229	1,1961	16,229	1,1952	0,0002	0,0009
43	16,881	0,8943	16,882	0,8937	0,0011	0,0006
44	17,067	0,9077	17,067	0,9068	0,0001	0,0009
45	18,443	0,9723	18,442	0,9728	0,0009	0,0005
46	18,715	0,7817	18,715	0,7819	0,0004	0,0002
47	18,995	0,7022	18,997	0,7035	0,0017	0,0013
48	19,394	0,5717	19,394	0,5709	0,0002	0,0008
49	19,872	0,4985	19,872	0,4988	0,0002	0,0003
50	21,091	0,5179	21,089	0,5183	0,0017	0,0004
51	21,628	0,4871	21,628	0,4877	0,0004	0,0006
52	21,961	0,4722	21,961	0,4723	0,0003	0,0001
53	25,301	0,5411	25,301	0,5399	0,0005	0,0012
54	26,501	0,4688	26,502	0,4680	0,0005	0,0008
55	27,648	0,4524	27,649	0,4525	0,0015	0,0001
56	30,363	0,5069	30,363	0,5074	0,0000	0,0005
57	31,496	0,5128	31,495	0,5142	0,0013	0,0014
58	33,563	0,5781	33,562	0,5789	0,0012	0,0008
59	33,764	0,5649	33,766	0,5640	0,0016	0,0009
60	34,190	0,5039	34,190	0,5035	0,0002	0,0004
61	36,852	0,5123	36,851	0,5120	0,0009	0,0003

Окончание табл. 1

1	2	3	4	5	6	7
62	39,590	0,5266	39,592	0,5260	0,0016	0,0006
63	41,894	0,8642	41,894	0,8633	0,0004	0,0009
64	44,180	0,7923	44,179	0,7918	0,0009	0,0005
65	46,397	0,6574	46,398	0,6568	0,0005	0,0006
66	48,642	0,7848	48,643	0,7859	0,0015	0,0011
67	48,898	0,8236	48,898	0,8232	0,0005	0,0004
68	49,192	0,5041	49,193	0,5032	0,0011	0,0009
69	49,461	0,5744	49,462	0,5755	0,0010	0,0011
70	50,542	0,9589	50,541	0,9586	0,0006	0,0003
71	51,661	1,6237	51,661	1,6238	0,0003	0,0001
72	52,598	0,7642	52,597	0,7635	0,0007	0,0007
73	53,341	1,2418	53,342	1,2405	0,0006	0,0013
74	54,428	0,6598	54,427	0,6589	0,0014	0,0009
75	55,305	1,0829	55,304	1,0841	0,0006	0,0012
76	55,541	0,9412	55,540	0,9422	0,0014	0,0010
77	56,324	0,6344	56,324	0,6353	0,0004	0,0009
78	56,845	0,7829	56,846	0,7842	0,0012	0,0013
79	57,157	0,6744	57,156	0,6754	0,0007	0,0010
80	58,021	0,4721	58,022	0,4722	0,0014	0,0001
81	58,644	0,6871	58,645	0,6856	0,0006	0,0015

Из табл. 1 следует, что максимальная величина модуля разницы между истинными и восстановленными значениями составила 0,0017 для времени выхода и 0,0015 для энергии выхода – справедливо будет заметить, что эти значения погрешности остаются максимальными только для выборки данных, соответствующей набору известных «истинных» значений. Однако получить числовые значения координат для незарегистрированных хроматографом локальных максимумов можно только с помощью аппроксимации, поэтому величину среднеквадратического отклонения для всех значений на данном этапе вычислить не представляется возможным без привлечения стороннего метода аппроксимации.

В рамках предметной области нефтехимического анализа минимальным шагом времени выхода (параметра оси абсцисс) для перехода от одного химического соединения к другому является пороговое значение 0,025 [11]: погрешность разработанного метода в 1,47 раза меньше обозначенной величины, поэтому он может применяться для решения задачи конвертации данных изображения в формат числового представления координат.

Использование векторного представления числовых восстановленных данных позволило получить существенный прирост в скорости обучения (табл. 2) интеллектуальной модели. В качестве персонального компьютера, на базе которого производилось обучение модели и тестирование ее точности, был использован ноутбук Apple MacBook Pro 16" выпуска 2023 г. в своей максимальной комплектации [12] с установленной версией macOS Sonoma 14.2.1 (23C71).

Таблица 2

Значение критериев оценки эффективности применения метода
для решения прикладной задачи

Table 2

Significance of criteria for evaluating the effectiveness of a method
in solving an applied problem

Тип модели	Время обучения (мин.)	Время, требуемое для предсказания (мин.)	Максимальное значение величины средн. отклонения при получении прогноза	
			Время выхода	Энергия выхода
Базовая модель	1230	8,2	0,0940	0,0610
Применение метода	148	4,3	0,0017	0,0015

Из табл. 2 следует, что время, требуемое для обучения модели и получения от нее предсказаний, сократилось в 8,3 и 1,9 раза соответственно.

Заключение

В рамках проведенного исследования был разработан алгоритм (и метод на его основе) для восстановления значений функции сигнала хроматографа, представленной в графическом виде. Исследована величина погрешности фиксации числовых значений координат для данных с хроматограммы тяжелой нефти. Обоснована достаточная точность восстановления значений для прикладного применения разработанного метода в рамках решения прикладной задачи нефтяной области. Выполнена векторизация значений восстановленных координат и модифицирована базовая интеллектуальная модель для возможности обработки нового представления входной информации. Зафиксировано сокращение величины среднеквадратического отклонения минимум на 0,0923 (по критерию времени выхода) при выполнении прогноза перераспределения углеводородных компонентов (решение прикладной задачи) в случае применения метода, что свидетельствует о формировании лучшей обобщающей способности интеллектуальной модели. Сокращение времени предсказания в 1,9 раза доказало целесообразность предварительной обработки входных данных хроматограмм тяжелой нефти.

Список источников

1. **Шадрина А. В., Крец В. Г.** Основы нефтегазового дела. М.: Нац. открытый ун-т «ИНТУИТ», 2016. 214 с.
2. **Басарыгин Ю. М., Будников В. Ф., Булатов А. И.** Исследование факторов и реализация мер долговременной эксплуатации нефтяных и газовых скважин. М.: Просвещение-Юг, 2004. 242 с.
3. **Шевченко Д. В., Васильева Л. Х.** Математическое моделирование вытеснения тяжелых нефтей горячей водой в тонком пласте. Казань: Познание, 2013. 60 с.
4. **Хуснутдинов И. Ш., Копылов А. Ю., Гончарова И. Н.** Разработка и совершенствование сольвентных технологий переработки тяжелого органического сырья. Казань: КГТУ, 2009. 265 с.

5. **Хадавимогаддам Ф., Мищенко И. Т., Мостаджеран М.** Применение методов искусственного интеллекта в прогнозировании основных свойств нефти // Газовая промышленность. 2019. Вып. 12(794). С. 28–32.
6. **Азиева Р. Х.** Искусственный интеллект в добыче нефти и газа: возможности и сценарный прогноз // Проблемы экономики и управления нефтегазовым комплексом. 2022. Вып. 3(207). С. 38–46. DOI: 10.33285/1999-6942-2022-3(207)-38-46
7. **Байбаров Д. А.** Оценка продуктивности и экономической эффективности технологий искусственного интеллекта для автоматизации процессов разведки и добычи нефти и газа // XXI век: итоги прошлого и проблемы настоящего плюс. 2021. Т. 10, вып. 3. С. 100–105. DOI: 10.46548/21vek-2021-1055-0019
8. **Ломакин Н. И., Дженифер О. Ч., Голодова О. А., Сычева А. В., Кабина В. В.** AI-система «Персептрон2 для прогноза финансового результата деятельности предприятия нефтяной отрасли РФ // Фундаментальные исследования. 2019. Вып. 12-1. С. 98–103. DOI: 10.17513/ft.42629
9. **Овсеенко Г. А., Козелков О. В., Кашаев Р. С.** Использование нейронных сетей в механическом устройстве представительного отбора и анализа проб // Приборостроение и автоматизированный электропривод в топливно-энергетическом комплексе и жилищно-коммунальном хозяйстве: Мат. VII Национальной науч.-практ. конф. 2022. С. 92–96.
10. **Vaswani A., Shazeer N. et al.** Attention is all you need. In Proc. of the 31st Conference on Neural Information Processing Systems (NIPS). 2017. 11 p.
11. **Dosovitskiy A., Beyer L. et al.** An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. In Proc. of the International Conference on Learning Representations (ICLR). 2021. 21 p. DOI: 10.48550/arXiv.2010.11929
12. **Arik S. Ö., Pfister T.** TabNet: Attentive Interpretable Tabular Learning // Proceedings of the AAAI Conference on Artificial Intelligence. 2020. Vol. 35(8). P. 6679–6687. DOI: 10.1609/aaai.v35i8.16826

References

1. **Shadrina A. V., Krets V. G.** Fundamentals of Oil and Gas Engineering. Moscow, National Open University of INTUIT, 2016, 214 p. (in Russ.)
2. **Basarygin Y. M., Budnikov V. F., Bulatov A. I.** Research of factors and implementation of measures for long-term operation of oil and gas wells. Moscow, Prosveshchenie-Yug publ., 2004, 242 p. (in Russ.)
3. **Shevchenko D. V., Vasilieva L. H.** Mathematical modelling of displacement of heavy oils by hot water in a thin reservoir. Kazan, Poznanie Publishing House, 2013, 60 p. (in Russ.)
4. **Khusnutdinov I. Sh., Kopylov A. Yu., Goncharova I. N.** Development and improvement of solvent technologies for processing of heavy organic raw materials. Kazan, KSTU, 2009, 265 p. (in Russ.)
5. **Hadavimoghaddam F., Mishchenko I. T., Mostajeran M.** Application of artificial intelligence methods in predicting basic oil properties. *Gas Industry*, 2019, vol. 12(794), pp. 28–32 (in Russ.)
6. **Azieva R. H.** Artificial intelligence in oil and gas production: opportunities and scenario forecast. *Problems of economics and management of oil and gas complex*, 2022, vol. 3(207), pp. 38–46 (in Russ.). DOI: 10.33285/1999-6942-2022-3(207)-38-46
7. **Baybarov D. A.** Assessment of productivity and economic efficiency of artificial intelligence technologies for automation of oil and gas exploration and production processes. *XXI century: Resumes of the Past and Challenges of the Present plus*, 2021, vol. 10, iss. 3, pp. 100–105 (in Russ.). DOI: 10.46548/21vek-2021-1055-0019
8. **Lomakin N. I., Jennifer O. C., Golodova O. A., Sycheva A. V., Kabina V. V.** AI-system “Perseptron” for forecasting the financial result of the enterprise of the oil industry of the Rus-

- sian Federation. *Fundamental Research*, 2019, vol. 12-1, pp. 98–103 (in Russ.). DOI: 10.17513/fr.42629
9. **Ovseenko G. A., Kozelkov O. V., Kashaev R. S.** Use of neural networks in mechatronic device of representative sampling and analysis. Instrumentation and automated electric drive in fuel and energy complex and housing and communal services. *Proceedings of VII National Scientific and Practical Conference*, 2022, pp. 92–96. (in Russ.)
 10. **Vaswani A., Shazeer N. et al.** Attention is all you need. In *Proc. of the 31st Conference on Neural Information Processing Systems (NIPS)*, 2017, 11 p.
 11. **Dosovitskiy A., Beyer L. et al.** An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. In *Proc. of the International Conference on Learning Representations (ICLR)*, 2021, 21 p. DOI: 10.48550/arXiv.2010.11929
 12. **Arik S. Ö., Pfister T.** TabNet: Attentive Interpretable Tabular Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, vol. 35(8), pp. 6679–6687. DOI: 10.1609/aaai.v35i8.16826

Сведения об авторах

Пылов Петр Андреевич, аспирант Кузбасского государственного технического университета им. Т. Ф. Горбачева; старший разработчик высоконагруженных интеллектуальных систем на позиции Senior Computer Vision Engineer
WoS Researcher ID: JMQ-9360-2023

Майтак Роман Вячеславович, магистрант Кузбасского государственного технического университета им. Т. Ф. Горбачева; Middle+ NLP Data Scientist
WoS Researcher ID: JMQ-9434-2023

Копылов Даниил Евгеньевич, магистрант Иркутского государственного университета

Information about the Authors

Petr A. Pylov, PhD Student at the T.F. Gorbachev Kuzbass State Technical University. He combines his studies with his work as a Senior Computer Vision Engineer

Roman V. Maitak, Master Student at the T.F. Gorbachev Kuzbass State Technical University. He combines his studies with her work as a data scientist at Middle+ NLP

Daniil E. Kopylov, Master Student of Irkutsk State University

*Статья поступила в редакцию 26.01.2024;
одобрена после рецензирования 07.05.2024; принята к публикации 07.05.2024*
*The article was submitted 26.01.2024;
approved after reviewing 07.05.2024; accepted for publication 07.05.2024*

Научная статья

УДК 004.4`2

DOI 10.25205/1818-7900-2024-22-2-57-67

Использование цепей Маркова для автоматического завершения исходного кода программы

Владислав Сергеевич Тимофеев

Новосибирский государственный университет
Новосибирск, Россия

v.timofeev@g.nsu.ru; <https://orcid.org/0009-0008-1060-4613>

Аннотация

В сфере программирования используются разнообразные инструменты с целью оптимизации процесса разработки. Среди них особое место занимают интегрированные среды разработки (IDE), обеспечивающие широкий спектр сервисов, включая текстовый редактор, отладчик и интеллектуальное завершение кода. Настоящая работа посвящена разработке модели, направленной на предсказание вариантов завершения исходного кода программы. Для улучшения точности модели были использованы комбинации цепей Маркова, основанные на различных методах вычисления текущего контекста программы: линейном и с использованием абстрактного синтаксического дерева (AST). Линейный метод анализа контекста представляет собой анализ токенизированного представления исходного кода, в то время как второй метод использует структуру исходного кода в виде AST. Объединение различных моделей позволяет сохранить больше семантической информации о коде и учитывать при автодополнении индивидуальный стиль написания кода. Разработанная модель демонстрирует высокую точность предсказаний при минимальном объеме вычислительных ресурсов, что делает ее применимой в интегрированных средах разработки.

Ключевые слова

автоматическое завершение кода, цепи Маркова, нейронные сети, интегрированные среды разработки, язык программирования Pascal

Для цитирования

Тимофеев В. С. Использование цепей Маркова для автоматического завершения исходного кода программы // Вестник НГУ. Серия: Информационные технологии. 2024. Т. 22, № 2. С. 57–67. DOI 10.25205/1818-7900-2024-22-2-57-67

Code Completion Using Markov Chains

Vladislav S. Timofeev

Novosibirsk State University, Novosibirsk, Russian Federation

v.timofeev@g.nsu.ru; <https://orcid.org/0009-0008-1060-4613>

Abstract

Modern software engineers use many tools to speed up the development process. Many of them use integrated development environments (IDEs), which provide services such as text editors, debuggers and even intelligent code completion. This paper is dedicated to the development of a model for predicting variants of program source code termination. To improve the accuracy of the model, we used combinations of Markov chains constructed using different ways of calculating the current context of the program: linear and with AST. The linear way of computing the context is an analysis of the tokenized representation of the source code. The second method, on the other hand, uses a representation of the

© Тимофеев В. С., 2024

source code in the form of an abstract syntax tree. Combining the different models preserves more semantic information about the code, also adding the ability to support custom code writing style features. In order to compare the different models, a new dataset has been created specifically for the Pascal language. A detailed comparison of the working mechanisms as well as the prediction accuracy on the collected data is given. The proposed model showed high enough accuracy of predictions with minimal computation costs, which allows using it in integrated development environments.

Keywords

code completion, Markov chains, neural networks, integrated development environment, Pascal programming language

For citation

Timofeev V. S. Code completion using Markov chains. *Vestnik NSU. Series: Information Technologies*, 2024, vol. 22, no. 2, pp. 57–67 (in Russ.) DOI 10.25205/1818-7900-2024-22-2-57-67

Введение

Важной частью современных интегрированных сред разработки (англ. integrated development environment, или IDE) является поддержка автоматического завершения исходного кода программы [1]. Данная технология в реальном времени предлагает пользователю подсказки с информацией о возможных вариантах завершения уже написанной части исходного кода. Самые простые реализации используют введенный пользователем префикс и предлагают список доступных сущностей языка (названия функций, классов, переменных и т. д.), предотвращая синтаксические ошибки и избавляя разработчиков от необходимости постоянно обращаться к документации за счет подсказок. Однако особенный интерес представляют модели, учитывающие текущий контекст. Такие модели выдают более корректную информацию, значительно ускоряя процесс разработки и помогая предотвратить семантические ошибки.

В данной статье описана разработка алгоритма автоматического завершения исходного кода программы с использованием комбинации из нескольких цепей Маркова, построенных с помощью различных способов вычисления текущего контекста программы. Последовательно проведены краткий обзор и анализ существующих решений, описаны особенности использования цепей Маркова в контексте рассматриваемой предметной области, представлен разработанный автором алгоритм завершения кода, а также проведен анализ результатов работы алгоритма. Проверка алгоритма осуществлялась на специально собранном наборе данных, состоящем из более чем 20000 программ.

Обзор существующих работ

В настоящее время уже существует достаточно много различных решений. Среди наиболее простых подходов стоит упомянуть алгоритм лексикографической сортировки ответов с заданным префиксом. Более продвинутые методы основаны на статистических подходах, включающих подсчет использования каждой языковой сущности и последующую сортировку ответов в соответствии с их частотой. Также существуют алгоритмы, которые вместо префикса используют различные варианты сокращений, например, аббревиатуры [2]. Тем не менее эти алгоритмы не учитывают текущий контекст пользователя, что приводит к снижению точности предсказаний [3].

С ростом популярности нейронных сетей появились модели, показывающие значительный прирост точности. Например, модель, использующая сети долгой краткосрочной памяти (англ. long short-term memory, или LSTM) [4]. В данном подходе сначала векторизуется токенизированный исходный код, затем применяются рекуррентные нейронные сети, учитывающие текущий контекст. Это значительно повышает точность работы модели. В статье [5] приводится подробное сравнение результатов работы различных подходов. Самую высокую точность показывает модель с использованием LSTM. В более актуальной работе [6] удалось повысить точность данной LSTM-модели за счет расширения набора данных для обучения. Однако

представленные в этих статьях решения направлены на предсказание только вызовов методов из популярных Python-библиотек, тогда как ключевая особенность данной работы заключается в разработке алгоритма, способного предсказывать любые языковые сущности, а также учитывающего особенности пользовательского кода. Помимо этого, LSTM-модель требует наличия значительного набора данных и вычислительных мощностей для обучения. Сбор такого количества данных для Pascal является проблематичным, так как Python является намного более популярным языком программирования. Также следует отметить, что из-за своей архитектуры представленные выше решения не расширяемы на другие языки программирования, а также достаточно чувствительны к пользовательскому стилю написания кода.

Разработка алгоритма автоматического завершения кода

В качестве языка программирования, для которого алгоритм будет предлагать варианты завершения кода, был выбран язык Pascal. Данный язык обладает относительно простой грамматикой, на которой значительно проще проверять гипотезы при построении модели. Предполагается, что решение может быть легко использовано для других языков, использующих грамматику Pascal. Например, многие языки для программируемых логических контроллеров: Structured Text (ST) или Structured Control Language (SCL), poST и т. д. Помимо этого, данное решение является частью разрабатываемой на данный момент IDE для Pascal с помощью инструмента Xtext [7].

Использование цепей Маркова

Для того чтобы точнее предсказывать, что сейчас будет писать пользователь, нам важно учесть контекст, проанализировав уже написанный пользователем код. Использование методов глубокого обучения для работы с исходным кодом программ является довольно сложной задачей, показывающей при этом относительно низкий прирост точности [8], поэтому для решения этой задачи предлагается использовать другой подход, а именно построение модели с использованием цепей Маркова. Цепь Маркова – это стохастическая модель, описывающая последовательность возможных событий, в которой вероятность каждого события зависит только от состояния, достигнутого в предыдущем событии [9]. Таким образом, предлагается накапливать предыдущие языковые сущности, чтобы предсказать следующую с помощью этой модели.

В качестве состояния предлагается использовать не сами языковые сущности, а их обезличенные типы: вызов процедуры, вызов функции, название переменной, название константы, название типа, строковый литерал, ключевое слово языка и т. д. Другими словами, задача цепи Маркова – предсказать следующий тип языковой сущности. Далее результат предсказания может быть использован как параметр других алгоритмов автоматического завершения кода. В данной статье рассмотрено следующее применение модели: список всех доступных пользователю языковых сущностей фильтруется по указанному префиксу и сортируется согласно вероятности их типа, полученной с помощью цепи Маркова. Рассмотрим алгоритм работы модели на конкретном примере (рис. 1).

В блоке (0) объявлены различные языковые сущности, для простоты все они имеют общий префикс «ABC1». Пользователь запрашивает варианты завершения для префикса «AB». Действие (1) обозначает сбор всех вариантов ответа, имеющих заданный префикс. Далее происходит вызов цепи Маркова (2), которая выдает типы возможных ответов и их вероятности. Затем объединяются полученные результаты (3): конкретные языковые сущности сортируются по вероятности их типа, и пользователю предлагается отсортированный список возможных

завершений. Стоит отметить, что для запроса предсказаний из собранного контекста формируется несколько ключей различной длины (рис. 2).

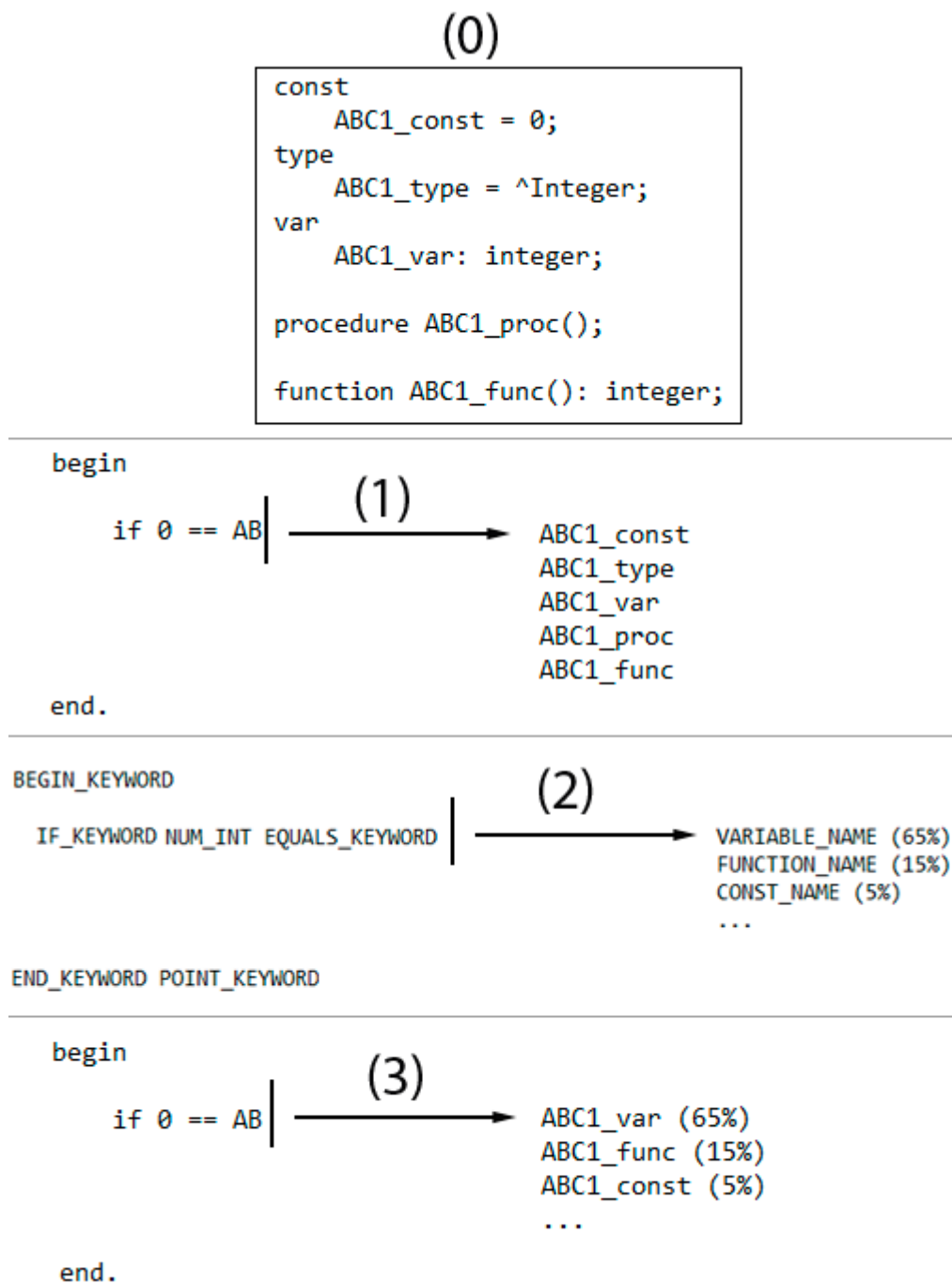
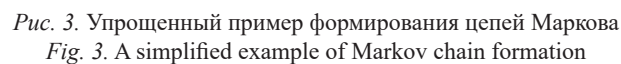


Рис. 1. Пример работы модели
Fig. 1. Example of model operation

Рис. 2. Псевдокод алгоритма получения предсказаний
Fig. 2. Pseudocode of the algorithm for obtaining predictions

Под линейным сбором контекста подразумевается, что исходный код представлен в виде последовательности токенов. Далее это представление используется для получения множества цепей событий, формируя непрерывные подмножества токенов различной длины. Рассмотрим упрощенный пример (рис. 3) формирования цепей Маркова с количеством событий равным трем, в котором, для простоты, не учитываются некоторые токены, состоящие из одного символа.



```

type Token:
    Int startOffset
    Int endOffset
    String text
    String type

String[] getContext(String fileContent, Int offset, Int maxContextSize):
    Token[] tokens = lexer(fileContent)
    for (Int i = 0; i < tokens.size; i++):
        Token token = tokens[i]
        if (token.startOffset <= offset && token.endOffset <= offset):
            String[] tokenTypes = []
            for (Int j = max(0, i - maxContextSize); j < i; j++):
                tokenTypes.add(tokens[j].type)
            return tokenTypes
    return []

```

Рис. 4. Псевдокод алгоритма линейного сбора контекста
 Fig. 4. Pseudocode of the linear context collection algorithm

Составив и обезличив все такие цепочки из набора данных для обучения, мы можем считать для данного набора типов вероятности следующих типов токенов. Все цепочки разбиваются на пары ключ – значение. Ключом выступает набор типов, а значением – список типов, которые могут идти после ключа, с их вероятностями. Когда алгоритму поступает запрос на завершение, достаточно собрать указанное заранее количество токенов слева, сформировать из них ключ и получить список следующих вероятных типов токенов (рис. 4). Стоит отметить, что с использованием хеш-таблиц данное решение работает по времени за $O(1)$ и состоит из набора примитивных операций, что позволяет иметь несколько различных наборов цепочек и агрегировать их результат.

Сбор контекста с использованием AST

Абстрактное синтаксическое дерево (англ. abstract syntax tree, или AST) – конечное помеченное ориентированное дерево, в котором внутренние вершины сопоставлены (помечены) с операторами языка программирования, а листья – с соответствующими операндами.

Если внимательнее изучить пример формирования цепей (рис. 3), то можно заметить, что при линейном сборе контекста довольно быстро (всего лишь за пару токенов) теряется информация о глобальном контексте. Начиная со второй цепи, ключевое слово «while» было полностью утеряно, т. е. модель не понимает, что пользователь, работая на данный момент с условием или телом цикла, хочет получить релевантные варианты завершения этой синтаксической конструкции. Данная проблема приводит к понижению точности предсказаний. Для ее решения предлагается использовать второй набор цепей Маркова, который был собран с использованием AST. Рассмотрим, как выглядит тот же код, но в виде упрощенного синтаксического дерева (рис. 5).

Теперь, когда алгоритму поступает запрос на завершение, он с использованием знания о позициях всех элементов в файле находит ближайшую к позиции курсора вершину, затем ее родительскую вершину, затем ее левого родственника, а после снова родительскую вершину и так далее, пока не наберется необходимое количество вершин (рис. 6). Эта цепочка эквивалентна пути от листа по родителям в LCRS-представлении n -арного дерева (англ. left-child right-sibling, или LCRS) [10]. Далее, как и в линейном варианте, алгоритм формирует ключ, получает из хеш-таблицы предсказание цепи Маркова и формирует ответ. Стоит отметить,

что в IDE зачастую уже имеется построенное AST, поэтому данное решение имеет ту же асимптотику временной сложности, что и линейная версия алгоритма.

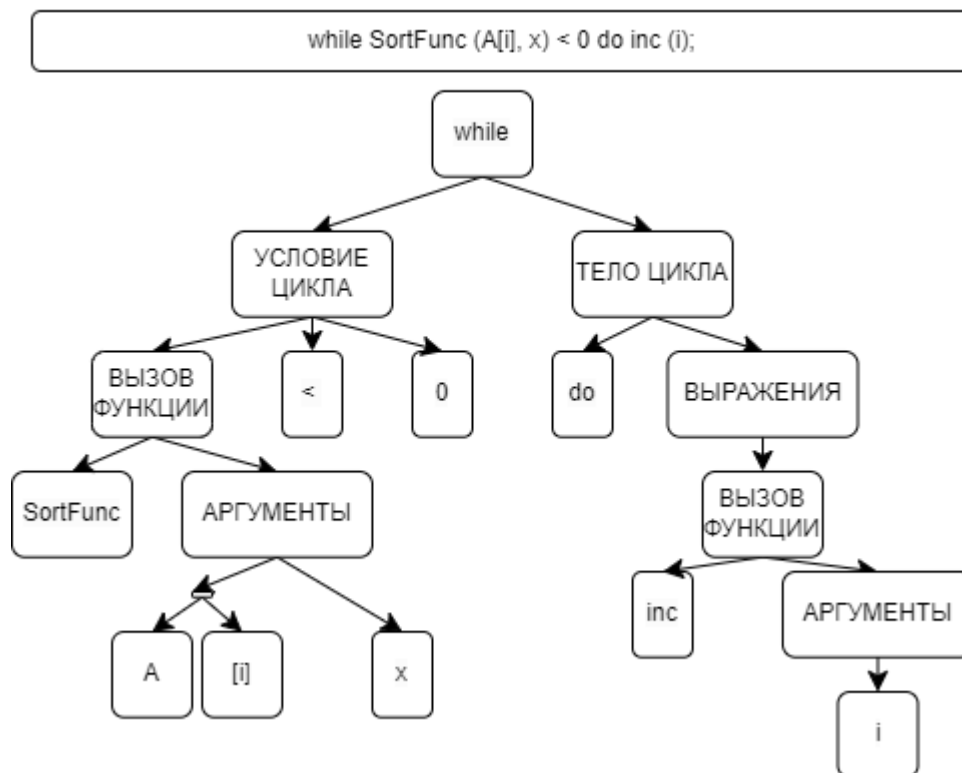


Рис. 5. Упрощенный пример AST
Fig. 5. Simplified AST example

```

type AstNode:
    Int startOffset
    Int endOffset
    String text
    String type
    AstNode parent
    AstNode leftSibling

String[] getContext(String fileContent, Int offset, Int maxContextSize):
    AstNode root = parse(fileContent)
    AstNode node = getDeepestNodeByOffset(root, offset)
    String[] nodeTypes = []
    while (nodeTypes.size < maxContextSize):
        if (node.leftSibling == NULL):
            node = node.parent
            if (node == NULL):
                break
        else:
            node = node.leftSibling
            nodeTypes.add(node.type)
    return nodeTypes
  
```

Рис. 6. Псевдокод алгоритма сбора контекста с использованием AST
Fig. 6. Pseudocode of the context collection algorithm using AST

Комбинирование моделей

Одна объемлющая синтаксическая конструкция обычно включает в себя несколько инструкций или выражений. Так как линейный контекст содержит больше информации о самих выражениях, то ожидается, что он зачастую будет показывать результат лучше, чем цепь, обученная на контексте из AST. Однако знание о структуре кода может повысить точность предсказаний для некоторых специфичных токенов, завершающих синтаксические конструкции [11]. Поэтому итоговый ответ предлагается формировать, объединив результаты цепи с линейным контекстом и цепи с AST.

Обе цепи были получены путем обучения на специально собранном наборе данных из более чем 20000 программ на языке Pascal. Важно отметить, что в статье подробно рассмотрены лишь два способа построения цепей, однако архитектура системы позволяет легко добавлять новые способы сбора контекста. Так, в модель были добавлены еще две цепи: линейная и с AST. Однако эти цепи обучаются на текущем пользовательском проекте, что позволяет учитывать стиль программирования пользователя. Для тестирования был создан проект, состоящий всего из 2000 строк кода, в котором для эмуляции стиля программирования во всех местах, где это возможно, результаты вызова функций передавались в качестве аргументов другой функции напрямую, без размещения в промежуточных переменных. Рассмотрим простой пример кода (рис. 7), где промежуточный результат присваивается в переменную, и тот же пример кода, но отредактированный под данную особенность стиля написания кода (рис. 8).

```
begin
    maxA = Max(a1, a2);
    maxB = Max(b1, b2);
    max = Max(maxA, maxB);
end.
```

Рис. 7. Пример кода

Fig. 7. Code example

```
begin
    max = Max(Max(a1, a2), Max(b1, b2));
end.
```

Рис. 8. Пример конкретной особенности стиля написания кода

Fig. 8. An example of a specific feature of code writing style

Зачастую в обучающем наборе использовался первый стиль написания, поэтому заранее обученные цепи при тестировании на пользовательском проекте после токена «FUNCTION_NAME» (вызов функции) предлагали самый вероятный следующий токен «VARIABLE_NAME» (имя переменной). Однако цепи, обученные на самом проекте, содержащем вышеописанные особенности стиля, в той же ситуации предлагали «FUNCTION_NAME» как более вероятный токен, что подтверждает гипотезу о применимости подхода для учета пользовательского стиля написания кода. Однако дальнейшее тестирование данной идеи осложняется

отсутствием в наборе данных достаточного количества проектов с ярко выраженными особенностями написания кода.

Метрики оценки решения

Для оценки решения использовались наиболее популярные в смежных исследованиях метрики:

1. Точность топ-1. Количество ответов, когда самый вероятный ответ модели являлся правильным.

2. Точность топ-2. Количество ответов, когда среди двух самых вероятных ответов модели один являлся правильным.

3. Степень уверенности модели. Так как модель предоставляет вероятности для каждого ответа, полученные вероятности для правильного ответа складывались. Если правильный ответ совсем отсутствовал в списке результатов, то его вероятность считалась равной нулю. Данная метрика применима только для оценки работы самих моделей, предсказывающих типы токенов.

Все метрики собирались для каждого вида токенов отдельно, что позволило провести анализ полученных данных и внести коррективы в модель.

Собранный набор данных был разделен на две части: 90 % данных использовалось для обучения модели, оставшиеся 10 % – для тестирования.

Результаты работы алгоритма

Для начала рассмотрим наиболее интересные результаты в разрезе типов токенов (рис. 9). Можно заметить, что модель, построенная на контексте, собранном с помощью AST, работает для ключевых слов языка (keywords) значительно лучше, чем модель с линейным контекстом. Тем не менее она плохо справляется в среднем – из-за низкой точности других типов токенов. Именно поэтому за счет комбинирования цепей удалось значительно увеличить точность предсказаний.

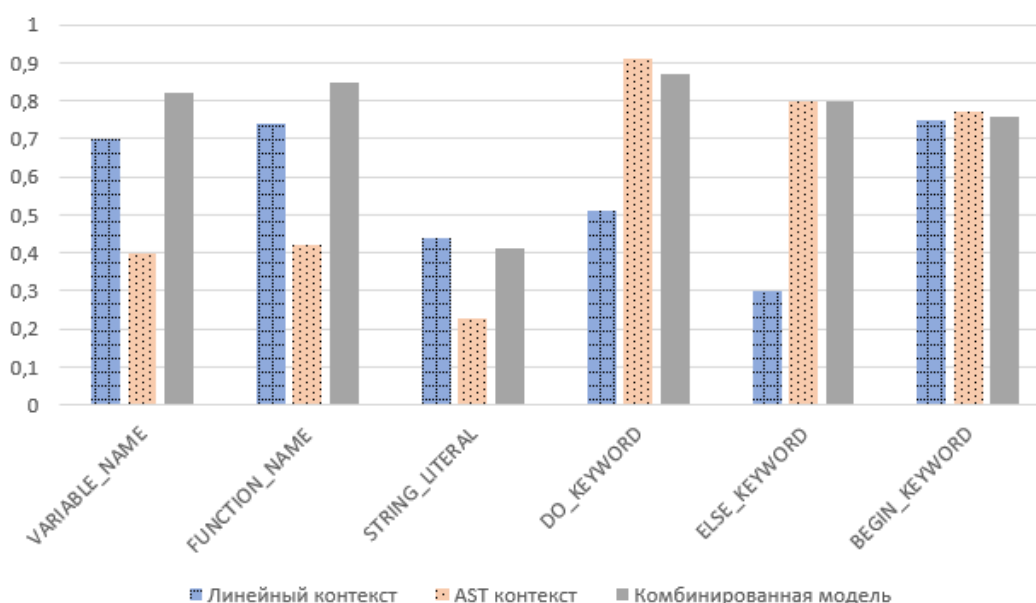


Рис. 9. Точность топ-1 моделей в разрезе типов токенов

Fig. 9. Top-1 model accuracy by token type

В результате тестирования моделей, предсказывающих типы токенов, были получены следующие значения метрик (табл. 1).

Таблица 1

Результаты работы моделей

Table 1

Results of the models

	Линейный контекст	AST-контекст	Комбинированная модель
Точность топ-1	0,49	0,23	0,57
Точность топ-2	0,64	0,37	0,79
Точность топ-5	0,98	0,72	0,98
Степень уверенности	0,84	0,79	0,91

Как и ожидалось, комбинированная модель показывает наилучший результат. Стоит отметить, что в таблице приведена точность предсказания только типа токена. Полученные типы токенов необходимо использовать в качестве дополнительного параметра для повышения точности алгоритма предсказаний самих значений токенов. Рассмотрим влияние работы модели на префиксный алгоритм предсказаний (табл. 2).

Таблица 2

Результаты работы алгоритмов предсказания завершения кода

Table 2

Results of the code completion algorithms

	Префиксный алгоритм	Префиксный алгоритм, с использованием комбинированной модели
Точность топ-1	0,09	0,28
Точность топ-2	0,17	0,41
Точность топ-5	0,43	0,64

Обычный префиксный алгоритм показывает довольно низкую точность топ-5 предсказания самих токенов, равную 0,43. Если отсортировать результаты работы префиксного алгоритма, используя результаты работы комбинированной модели, можно заметить ощутимый прирост точности до 0,64. Данная точность является приемлемой для использования в разрабатываемой IDE для Pascal, поэтому было решено использовать данный алгоритм. Однако предполагается, что, благодаря высокой производительности модели, результаты ее работы могут быть интегрированы в уже существующие более сложные алгоритмы.

Заключение

В статье представлен анализ существующих подходов к решению задачи автоматического завершения исходного кода программы. Разработана архитектура модели, комбинирующая цепи Маркова, обученные на различных контекстах: линейном и с использованием AST. В кратком изложении представлен анализ устойчивости модели к различным пользовательским стилям написания кода за счет включения в модель цепей, обучаемых на текущем пользовательском проекте. В качестве набора данных были собраны более 20000 различных программ на языке Pascal. Представлен сравнительный анализ работы нескольких реализаций моделей. Показано, что комбинированная модель обладает более высокой точностью по сравнению с остальными

моделями. На примере префиксного алгоритма продемонстрировано, что модель позволяет повысить точность предсказаний.

Данное решение является важной частью в разрабатываемой IDE для Pascal и позволит пользователям эффективнее работать со средой программирования за счет более точных подсказок с вариантами завершения кода.

Список литературы / References

1. **Marasoiu M., Church L., Blackwell A.** An empirical investigation of code completion usage by professional software developers. Annual Workshop of the Psychology of Programming Interest Group, 2015.
2. **Han S., Wallace D. R., Miller R. C.** Code Completion from Abbreviated Input. 2009 *IEEE/ACM International Conference on Automated Software Engineering*, Nov. 2009, pp. 332–343. DOI: 10.1109/ase.2009.64
3. **Hou D., Pletcher D. M.** An evaluation of the strategies of sorting, filtering, and grouping API methods for Code Completion. 2011 *27th IEEE International Conference on Software Maintenance (ICSM)*, Sep. 2011, pp. 233–242. DOI: 10.1109/icsm.2011.6080790
4. **Ginzberg A., Kostas L., Balakrishnan T.** Automatic code completion. Stanford, Class Project, 2017.
5. **Svyatkovskiy A., Zhao Y., Fu S., Sundaresan N.** Pythia: AI-assisted Code Completion System. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Jul. 2019, pp. 2727–2735. DOI: 10.1145/3292500.3330699
6. **Buksbaum D.** Increasing Code Completion Accuracy in Pythia Models for Non-Standard Python Libraries. Doctoral dissertation. Nova Southeastern University, 2023, https://nsuworks.nova.edu/gscis_etd/1188
7. **Eysholdt M., Behrens H.** Xtext: implement your language faster than the quick and dirty way. *Proceedings of the ACM international conference companion on Object oriented programming systems languages and applications companion*, Oct. 2010, pp. 307–309. DOI: 10.1145/1869542.1869625
8. **Hellendoorn V. J., Devanbu P.** Are deep neural networks the best choice for modeling source code? *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*, Aug. 2017, pp. 763–773. DOI: 10.1145/3106237.3106290
9. **Chan K. C., Lenard C. T., Mills T. M.** An introduction to Markov chains. *49th Annual Conference of Mathematical Association of Victoria*, 2012, pp. 40–47. DOI: 10.13140/2.1.1833.8248
10. **Cormen T. H., Leiserson C. E., Rivest R. L., Stein C.** Introduction to Algorithms (2nd ed.). MIT Press and McGraw-Hill, 2001, pp. 214–217.
11. **Wu B., Liang B., Zhang X.** Turn tree into graph: Automatic code review via simplified AST driven graph convolutional network. *Knowledge-Based Systems*, 2022, pp. 109450. DOI: 10.1016/j.knosys.2022.109450

Сведения об авторах

Тимофеев Владислав Сергеевич, магистрант

Information about the Author

Vladislav S. Timofeev, Master's Student

*Статья поступила в редакцию 22.04.2024;
одобрена после рецензирования 24.06.2024; принята к публикации 24.06.2024*

*The article was submitted 22.04.2024;
approved after reviewing 24.06.2024; accepted for publication 24.06.2024*

Научная статья

УДК 004.8

DOI 10.25205/1818-7900-2024-22-2-68-78

Исследование архитектур нейронных сетей для определения концентраций газов по спектрам

Александр Евгеньевич Щелоков¹
Константин Иванович Будников²

Новосибирский государственный университет
Институт автоматики и электрометрии СО РАН
Новосибирск, Россия

¹azarion123@gmail.com

²budnikov@iae.nsk.su

Аннотация

Исследованы нейронные сети разных архитектур с целью определения концентраций газов по спектрам, полученным с помощью оптического эмиссионного газоанализатора, который измеряет спектр электромагнитного излучения, испускаемого газами при возбуждении электрическим разрядом. Нейронная сеть обучается на данных из лаборатории по оптической спектроскопии и способна предсказывать концентрации газов по спектрам с высокой скоростью. Исследования касались архитектур глубокой нейронной сети со сверточными и рекуррентными слоями. Сверточные слои выделяют особенности спектров, а рекуррентные слои учитывают последовательную структуру данных. Качество работы нейронной сети оценивали по коэффициенту детерминации R^2 , а сравнение между сетями проводили по показателю RMSE между предсказанными и реальными концентрациями газов.

Ключевые слова

спектральный анализ, нейронные сети, глубокое обучение

Для цитирования

Щелоков А. Е., Будников К. И. Исследование архитектур нейронных сетей для определения концентраций газов по спектрам // Вестник НГУ. Серия: Информационные технологии. 2024. Т. 22, № 2. С. 68–78. DOI 10.25205/1818-7900-2024-22-2-68-78

© Щелоков А. Е., Будников К. И., 2024

Study of Neural Network Architectures for Determining Gas Concentrations by Spectra

Alexander E. Shchelokov¹, Konstantin I. Budnikov²

Novosibirsk State University
Institute of Automation and Electrometry SB RAS
Novosibirsk, Russian Federation

¹azarion123@gmail.com

²budnikov@iae.nsk.su

Abstract

A study of a number of neural networks of different architectures for determining gas concentrations from spectra obtained using an optical emission gas analyzer, which measures the spectrum of electromagnetic radiation emitted by gases when excited by an electric discharge, is presented. The neural network is trained on data from the optical spectroscopy laboratory and is able to predict gas concentrations from spectra at high speed. The research concerned the deep neural network architectures with convolutional and recurrent layers. Convolutional layers highlight the features of the spectra, while recurrent layers take into account the consistent structure of the data. The quality of the neural network is evaluated by the R^2 coefficient of determination, and the comparison between networks by the RMSE indicator between the predicted and real gas concentrations.

Keywords

spectral analysis, neural networks, deep learning

For citation

Shchelokov A. E., Budnikov K. I. Study of neural network architectures for determining gas concentrations by spectra. *Vestnik NSU. Series: Information Technologies*, 2024, vol. 22, no. 2, pp. 68–78 (in Russ.) DOI 10.25205/1818-7900-2024-22-1-68-78

Введение

Определение концентраций компонентов газовой смеси по их спектральным характеристикам представляет собой важную задачу, имеющую широкий спектр практического применения, включая мониторинг загрязнения атмосферы, анализ химических процессов, диагностику состояния плазмы и многие другие области. Одним из ключевых методов, используемых для измерения концентраций газов, является оптическая спектроскопия, основанная на измерении спектра электромагнитного излучения, испускаемого газами при их возбуждении электрическим разрядом. Каждый газ обладает уникальным спектром поглощения и излучения, который определяется его химическим составом и физическими условиями. Анализ спектра позволяет определить, какие газы и в каких количествах присутствуют в смеси.

В зависимости от химического состава спектры компонентов газовой смеси могут быть искажены или перекрыты спектрами других составляющих смесь газов, что значительно осложняет их определение. Кроме того, спектры могут изменяться в зависимости от параметров измерительного прибора, таких как длина волны, интенсивность, разрешение и т. д. В связи с этим необходимы эффективные алгоритмы обработки и анализа спектров, которые могут учитывать эти факторы и обеспечивать точные и быстрые результаты.

В настоящее время в качестве одного из наиболее перспективных подходов к решению этой задачи выделяются нейронные сети. Они способны обучаться на основе предъявленных им образцов и впоследствии выявлять сложные зависимости между входными и выходными данными.

Анализ смесей газов с целью установления их качественного и количественного состава в настоящее время широко применяется и при поиске природных залежей нефти. Основными компонентами попутной газовой смеси являются элементы гомологического ряда алканов: метан, этан, пропан, бутан и пентан с преобладающим содержанием метана [1]. Контроль их

содержания происходит при помощи газоанализаторов. На данный момент наиболее распространены следующие их виды:

- электрохимические, определяющие массовую концентрацию газа в смеси по показателям электрической проводимости электролита, который поглотил этот газ;
- оптико-эмиссионные, основанные на том, что атомы каждого химического элемента могут испускать свет определенных длин волн – спектральные линии, причем эти длины волн разные для разных элементов.

Результат работы электрохимических газоанализаторов легко обработать и интерпретировать, но при этом такой газоанализатор обладает невысокой точностью и избирательностью.

Необходимо получить точный и быстрый результат, что можно достичь с помощью оптического эмиссионного газоанализатора, который позволяет получать динамику изменения газовой смеси в виде набора спектров. Но анализ такого набора классическими методами очень сложен. Поэтому необходим алгоритм, способный упростить и ускорить работу со спектрами.

Цель настоящей работы – исследовать архитектуры нейронных сетей, чтобы определить концентрации газов по их спектрам, полученным с помощью оптического эмиссионного газоанализатора на буровой установке. Обучение нейронной сети происходит на эталонных данных, полученных в лаборатории оптической спектроскопии. Впоследствии она делает предсказания концентраций газов по спектрам входных данных.

Обзор существующих решений

Изучение спектров методами машинного обучения, в том числе при помощи нейронных сетей, в настоящее время получило широкое распространение. Так, в [2] используется метод опорных векторов для классификации трех видов чая по их спектрам. Лучшие результаты точности идентификации были до 90, 99 и 93,33 % соответственно при обучении, а при тестировании – 90, 99 и 95 %.

В работе [3] также исследовали спектры сортов чая. В ходе экспериментов было получено, что сочетание методов спектрометрии и нейронных сетей позволяет эффективно и быстро классифицировать виды чаев.

В исследовании [4] был разработан и применен метод на основе спектроскопии, индуцированной лазерным пробоем и нейронных сетей для идентификации, контроля качества, отслеживаемости и обнаружения подделки оливковых масел с высоким качеством.

Сверточные нейронные сети также показали высокую эффективность при определении калия в почве [5]. Авторы использовали метод, называемый временно разрешенной лазерной спектроскопией индуцированного пробоя (LIBS), в сочетании со сверточными нейронными сетями (CNN). Они обнаружили, что традиционный метод LIBS имеет сложности с точным определением количества калия из-за влияния окружающей среды. Однако когда они применили CNN к данным, полученным с помощью LIBS, они смогли значительно улучшить точность обнаружения калия.

В нефтегазовой отрасли нейронные сети также получили широкое распространение [6]. Поскольку большинство традиционных источников нефти и газа уже обнаружены и добываются в течение многих лет, то разведка, бурение, добыча и управление месторождениями нефти и газа в настоящее время представляют сложную задачу. Поэтому нефтяники используют передовые методы, такие как искусственные нейронные сети для поддержки принятия решений о сокращении непроизводительных финансовых и временных затрат. Так, нейронные сети применяются для обработки данных при сейсморазведке [7; 8], бурении скважин [9], прогнозе добычи газа [10] и нефти [11].

Исследуемые модели нейронных сетей

Проведенные исследования касались архитектур глубокой нейронной сети [12; 13] с одномерными сверточными, рекуррентными и линейными слоями. Исследуемые модели нейронных сетей выбирались из следующих соображений. Предназначение сверточных слоев – выделить особенности исследуемых спектров, а рекуррентные слои должны учитывать последовательную временную структуру данных. Финальным звеном каждой архитектуры выступал полносвязный слой, который формировал результат работы модели (рис. 1).

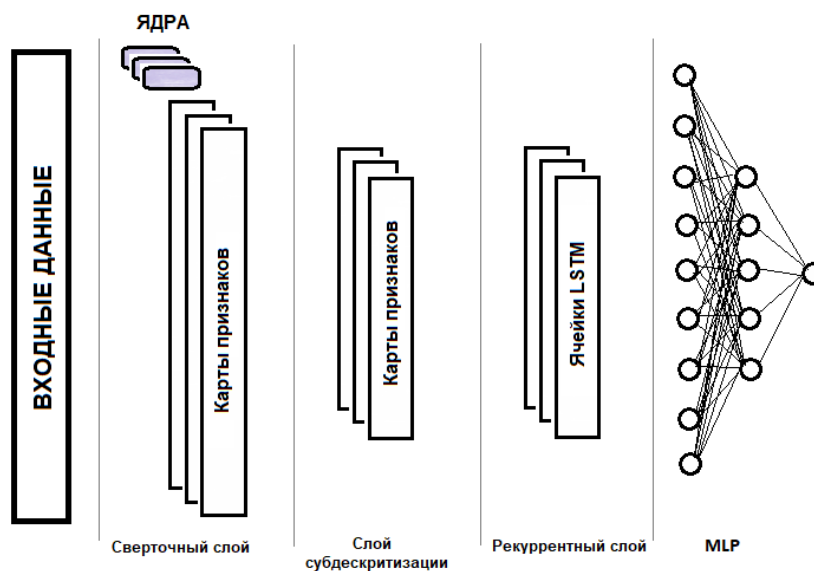


Рис. 1. Архитектура исследуемых моделей нейронных сетей
Fig. 1. Architecture of the studied neural network models

Последовательность слоев можно описать следующим образом.

1. Входной слой. На него подаются данные спектров.
2. Одномерные сверточные слои. Варьировалось количество слоев (от 1 до 3) и размер сверточного ядра (от 1 до 9).
3. Подвыборочные слои. Связаны со сверточными слоями.
4. Рекуррентные слои. Варьировалось количество слоев (от 1 до 3).
5. Выходной полносвязный слой.

Для оценки качества прогнозов нейронной сети использован коэффициент детерминации R^2 ,

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2},$$

а сравнение между сетями по показателю корня из среднеквадратичной ошибки RMSE между предсказанными и реальными концентрациями газов:

$$\text{RMSE} = \sqrt{\frac{1}{N} * \sum_{i=1}^N (y_i - \hat{y}_i)^2}.$$

В качестве функции потерь применялся показатель среднеквадратичной ошибки MSE:

$$MSE = \sum_{i=1}^N \frac{(y_i - \hat{y}_i)^2}{n},$$

где N – количество элементов в выборке, $\hat{y}_i$ – реальное значение, а y_i – предсказанное моделью.

Экспериментальное исследование моделей нейронных сетей

Для исследований были построены модели нейронных сетей с использованием различного количества LSTM, сверточных, линейных слоев и слоев пулинга и их комбинаций с одним из вариантов построения пространства признаков. В процессе исследования решались 2 задачи:

- 1) определение концентраций всех газов одновременно;
- 2) определение концентрации каждого газа в отдельности.

Тактика исследования заключалась в первоначальном исследовании архитектур для первой задачи, выявлении лидера и решении второй задачи.

В качестве набора данных использованы показания оптико-эмиссионных спектрометров, полученных с буровой станции. Показания снимались с периодичностью в 1 минуту. Один блок спектральных данных представляет собой 2048 вещественных чисел, которые записаны в виде текстового файла.

Весь набор представляет собой спектры, полученные в определенные моменты времени, и информацию о концентрации 7 газов – метана, этана, пропана, изобутана, бутана, изопентана, пентана. Для обучения и анализа в качестве одного объекта был представлен набор спектров между двумя последовательными i -й и $i + 1$ минутами (признаки) и набор концентраций за $i + 1$ минуту в качестве результата.

Для построения пространства признаков были подготовлены 4 трансформации данных:

- использовать все спектры за исследуемый промежуток времени;
- использовать последние n спектров, (n – натуральное, $0 < n < \text{число спектров}$);
- использовать спектры с некоторым шагом;
- один спектр на некоторый промежуток времени (выбранный случайным образом).

В процессе определения эффективности трансформаций входных данных наиболее эффективной оказалась подвыборка, которая использует выборку блока спектральных данных с определенным шагом.

Модель с рекуррентными слоями для одновременного определения всех газов

После экспериментов с различными комбинациями сверточных, подвыборочных и рекуррентных слоев и их параметров выявилась десятка лидеров, таблица с метриками RMSE которых представлена в табл. 1. Меньшее значение соответствует лучшему показанию.

В результате сравнения выделены 2 варианта архитектуры с номерами 2 и 10. Для них определены показатели коэффициента дискриминации R^2 , которые показывают качество прогноза каждой из моделей. Они приведены в табл. 2. Наилучшие показатели имеет модель с номером 10. Она хорошо прогнозирует концентрацию всех газов, за исключением метана.

Таблица 1

Метрика RMSE для моделей на основе LSTM и сверточных слоев

Table 1

RMSE metric for LSTM-based models and convolutional layers

Газы/ модель	lstm_ model1	lstm_ model2	lstm_ model3	lstm_ model5	lstm_ model7	lstm_ model8	lstm_ model9	lstm_ model10
Метан	0,006	0,00047	0,00027	0,00026	0,00030	0,00030	0,00032	0,001
Этан	0,032	0,034	0,035	0,039	0,032	0,039	0,039	0,024
Пропан	0,152	0,150	0,150	0,190	0,107	0,190	0,190	0,108
Изобутан	0,109	0,097	0,114	0,141	0,107	0,141	0,141	0,066
Бутан	0,341	0,282	0,359	0,209	0,199	0,878	0,411	0,173
Изопентан	0,076	0,060	0,086	0,090	0,042	0,090	0,090	0,045
Пентан	0,027	0,021	0,029	0,031	0,015	0,031	0,031	0,015

Таблица 2

Метрика R^2 для моделей на основе LSTM и сверточных слоев

Table 2

 R^2 metric for LSTM-based models and convolutional layers

Газы/модель	lstm_model10	lstm_model_2
Метан	–3,461	–1,4
Этан	0,586	0,158
Пропан	0,629	0,292
Изобутан	0,71	0,391
Бутан	0,769	0,387
Изопентан	0,691	0,455
Пентан	0,695	0,421

Графики с аппроксимацией концентраций пентана моделью для lstm_model_2 и lstm_model10 представлены на рис. 2 и 3 соответственно. Красным цветом показаны данные, полученные моделью, а зеленым – реальные данные.

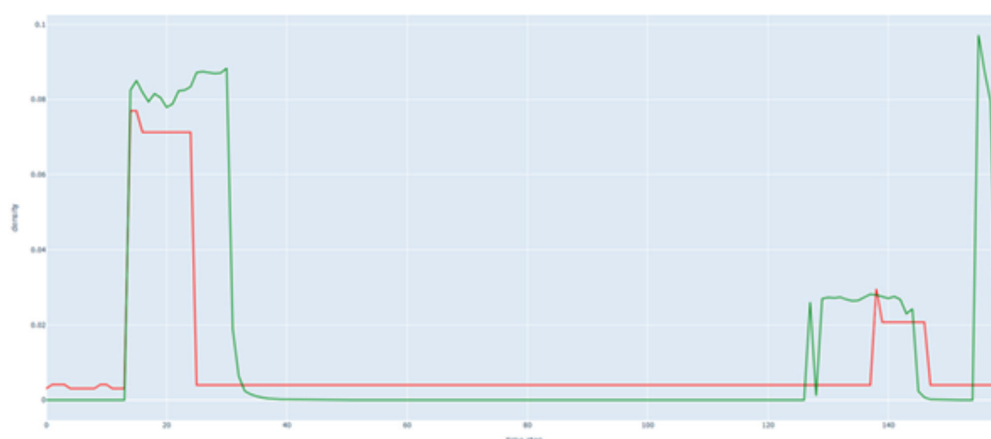


Рис. 2. График аппроксимации пентана моделью lstm_model2

Fig. 2. Graph of pentane approximation lstm_model2 model

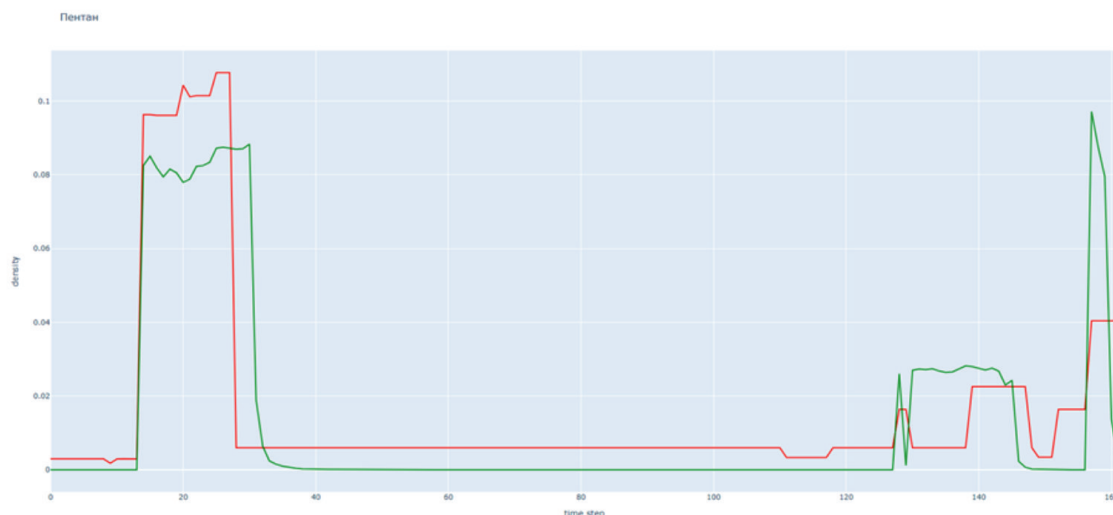


Рис. 3. График аппроксимации пентана моделью lstm_model10

Fig. 3. Graph of the pentane approximation model lstm_model10

Модель с рекуррентными слоями для определения отдельного газа

Для определения концентрации каждого газа по отдельности была создана модель с использованием сверточных, линейных и LSTM-слоев на основе модели 10. Изменению подвергся выход модели, полносвязный слой. В предыдущем исследовании он имел 7 выходов (по 1 на каждый газ), а в данном исследовании оставлен только 1 выход. В итоге обучалось 7 независимых моделей, каждая на свой газ.

Ниже представлены таблицы со значениями RMSE (табл. 3) и R^2 (табл. 4) для всех 7 моделей. Результат получился хуже, чем при применении модели на основе LSTM и сверточных слоев для всех газов сразу, в силу взаимного влияния спектров газов на их концентрации.

Таблица 3

Метрика RMSE для модели с LSTM и сверткой для отдельного газа

Table 3

RMSE metric for a model with LSTM and convolution for a single gas

Газы/модель	7 сетей
Метан	0,00030
Этан	0,031
Пропан	0,149
Изобутан	0,110
Бутан	0,327
Изопентан	0,074
Пентан	0,028

Таблица 4

Метрика R^2 для модели с LSTM и сверткой для отдельного газа

Table 4

 R^2 metric for a model with LSTM and convolution for a single gas

Газы/модель	7 сетей
Метан	0,0375829334
Этан	0,2887644838
Пропан	0,2893148676
Изобутан	0,2056482239
Бутан	0,1737171388
Изопентан	0,1692956415
Пентан	-0,0266482254

Регрессионные модели на основе одномерных сверток

С целью облегчения модели для выделения признаков из одного спектра, без учета серии и, следовательно, без рекуррентных слоев LSTM, была построена архитектура, в которой использовались только одномерные сверточные и полносвязные слои.

Ниже показаны результаты метрик RMSE (табл. 5) и R^2 (табл. 6) после обучения 2 лучших моделей архитектуры cmlp_2 и cmlp_3.

Таблица 5

Метрика RMSE для модели без LSTM на основе одномерных сверток

Table 5

RMSE metric for a non-LSTM model based on one-dimensional convolutions

Газы/модель	cmlp_2	cmlp_3
Метан	0,00080189	0,00100474
Этан	0,03460961	0,03651403
Пропан	0,16876938	0,1771527
Изобутан	0,09626544	0,10771695
Бутан	0,22776742	0,29589342
Изопентан	0,04309401	0,06060036
Пентан	0,01213099	0,02126244

Таблица 6

Метрика R^2 для модели без LSTM на основе одномерных сверток

Table 6

 R^2 metric for a non-LSTM model based on one-dimensional convolutions

Газы/модель	cmlp_2	cmlp_3
Метан	-4,21725655	-7,19075073
Этан	0,32370003	0,24722458
Пропан	0,30557955	0,23487787
Изобутан	0,50529742	0,38059954
Бутан	0,66708366	0,43814724
Изопентан	0,7841424	0,57314152
Пентан	0,83986919	0,50806441

Из табл. 6 видно, что результаты сетей данного вида уступают сетям с рекуррентными слоями.

Заключение

Использование сверточных слоев является ключевым компонентом в исследованных моделях. Это видно из приведенных результатов. Однако без учета временной последовательности результаты получаются существенно хуже. Также следует отметить тот факт, что из массива данных спектров газов, полученных при групповом анализе, не удастся построить модель для определения концентрации газа в отдельности.

Представленное исследование архитектур нейронных сетей, построенных на основе сверточных, рекуррентных и полносвязных слоев, для определения концентраций газов по спектрам, полученным с помощью оптического эмиссионного газоанализатора, показало применимость данного метода для быстрого определения концентраций газов. Нейронная сеть обучается на данных, полученных из лаборатории по оптической спектроскопии, и способна предсказывать концентрации газов по спектрам с высокой скоростью. Качество работы нейронной сети оценивается по коэффициенту детерминации R^2 , который доходил до 77 %, а сравнение между сетями проводится по показателю RMSE между предсказанными и реальными концентрациями газов. Дальнейшее совершенствование методики требует использования спектров газов в отдельности.

Впоследствии планируется провести эксперименты по извлечению признаков из чистых газов с дальнейшим применением их к газовым смесям для повышения качества.

Список литературы

1. **Чумаков В. М.** Определение выбросов одоранта на узле одоризации газораспределительных станций // Газовая промышленность. Сер.: Подготовка, переработка и использование газа. Экспресс-информация. М.: ВНИИГазпром, 1987. С. 15–16.
2. **Chen Q.** et al. Feasibility study on identification of green, black and Oolong teas using near-infrared reflectance spectroscopy based on support vector machine (SVM) // Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 2007. Т. 66, № 3. С. 568–574.
3. **Круглов В. В., Борисов В. В.** Искусственные нейронные сети. Теория и практика. М.: Горячая линия – Телеком, 2001. Т. 382.
4. **Caceres J. O.** et al. Application of laser-induced breakdown spectroscopy (LIBS) and neural networks to olive oils analysis // Applied spectroscopy. 2013. Т. 67, № 9. С. 1064–1072.
5. **Chengxu L. U.** et al. Detection of K in soil using time-resolved laser-induced breakdown spectroscopy based on convolutional neural networks // Plasma Science and Technology. 2018. Т. 21, № 3. С. 034014.
6. **Husam Alkinani, Abo Taleb Tuama** et. al. Applications of Artificial Neural Networks in the Petroleum Industry: A Review // Proc.SPE Middle East Oil and Gas Show and Conference, March 2019. DOI: 10.2118/195072-MS
7. **Karrenbach M., Essenreiter R., Treitel S.** Multiple Attenuation With Attribute-based NeuralNetworks // Society of Exploration Geophysicists. January 2000. DOI: 10.1190/1.1815827
8. **Aminzadeh F., Maity D., Tafti T. A., Brouwer, F.** Artificial Neural Network Based Autopicker ForMicro-earthquake Data // Society of Exploration Geophysicist. Jan. 2011. DOI: 10.1190/1.3627514
9. **Dashevskiy D., Dubinsky V., Macpherson J. D.** Application of Neural Networks for Predictive Control in Drilling Dynamics // Society of Petroleum Engineers. Jan. 1999. DOI:10.2118/56442-MS

10. **Ghahfarokhi P. K., Carr T., Bhattacharya S., Elliott J., Shahkarami A., Martin K.** A Fiber-OpticAssisted Multilayer Perceptron Reservoir Production Modeling: A Machine Learning Approach in Prediction of GasProduction // *Proc. Unconventional Resources Technology Conference*, August 9, 2018. DOI: 10.15530/urtec-2018-2902641
11. **Gholamreza Moradi, Majid Mohadesi, Mohammad Reza Moradi.** Prediction of wax disappearance temperature usingartificial neural networks// *Journal of Petroleum Science and Engineering*. 2013. Vol. 108. P. 74–81. <http://dx.doi.org/10.1016/j.petrol.2013.06.003>
12. **LeCun Y., Bengio Y., Hinton G.** Deep learning // *Nature*. 2015. T. 521, № 7553. С. 436–444.
13. **Аггарвал Ч.** Нейронные сети и глубокое обучение: учебный курс. СПб.: Диалектика, 2020. 752 с.

References

1. **Chumakov B. M.** Determination of odorant emissions at the odorization node of gas distribution stations. *Gas Industry. Ser.: Preparation, processing and use of gas. Express information*. Moscow, VNIIEgazprom publ., 1987, pp. 15–16. (in Russ.)
2. **Chen Q. et al.** Feasibility study on identification of green, black and Oolong teas using near-infrared reflectance spectroscopy based on support vector machine (SVM). *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2007, vol. 66, no. 3, pp. 568–574.
3. **Kruglov V. V., Borisov V. V.** Artificial neural networks. Theory and practice. Moscow, Hotline–telecom, 2001, vol. 382. (in Russ.)
4. **Caceres J. O. et al.** Application of laser-induced breakdown spectroscopy (LIBS) and neural networks to olive oils analysis. *Applied spectroscopy*, 2013, vol. 67, no. 9, pp. 1064–1072.
5. **Chengxu L. U. et al.** Detection of K in soil using time-resolved laser-induced breakdown spectroscopy based on convolutional neural networks. *Plasma Science and Technology*, 2018, vol. 21, no. 3, p. 034014.
6. **Husam Alkinani, Abo Taleb Tuama et. al.** Applications of Artificial Neural Networks in the Petroleum Industry: A Review. *Proc. SPE Middle East Oil and Gas Show and Conference*, March 2019, DOI: 10.2118/195072-MS
7. **Karrenbach M., Essenreiter R., & Treitel S.** Multiple Attenuation With Attribute-based NeuralNetworks. *Society of Exploration Geophysicists*, January 2000, DOI: 10.1190/1.1815827
8. **Aminzadeh F., Maity D., Tafti T. A., & Brouwer F.** Artificial Neural Network Based Autopicker ForMicro-earthquake Data. *Society of Exploration Geophysicist*, January 2011, DOI: 10.1190/1.3627514
9. **Dashevskiy D., Dubinsky V., & Macpherson J. D.** Application of Neural Networks for Predictive Control in Drilling Dynamics. *Society of Petroleum Engineers*, January 1999. DOI:10.2118/56442-MS
10. **Ghahfarokhi P. K., Carr T., Bhattacharya S., Elliott J., Shahkarami A., & Martin K.** A Fiber-OpticAssisted Multilayer Perceptron Reservoir Production Modeling: A Machine Learning Approach in Prediction of GasProduction. *Proc. Unconventional Resources Technology Conference*, August 9, 2018, DOI: 10.15530/urtec-2018-2902641
11. **Gholamreza Moradi, Majid Mohadesi, Mohammad Reza Moradi** Prediction of wax disappearance temperature usingartificial neural networks. *Journal of Petroleum Science and Engineering*, 2013, vol. 108, pp. 74–81, <http://dx.doi.org/10.1016/j.petrol.2013.06.003>
12. **Lecun Y., Bengio Y., Hinton G.** Deep learning. *Nature*, 2015, vol. 521, no. 7553, pp. 436–444. (in Russ.)
13. **Aggarwal Ch.** Neural networks and deep learning: a training course. *Dialectics LLC*, St. Petersburg, 2020, 752 p.

Сведения об авторах

Щелоков Александр Евгеньевич, магистрант кафедры компьютерных технологий факультета информационных технологий НГУ, инженер-программист Института автоматизации и электрометрии СО РАН

Будников Константин Иванович, кандидат технических наук, доцент кафедры компьютерных технологий факультета информационных технологий НГУ, старший научный сотрудник, заведующий тематической группы Института автоматизации и электрометрии СО РАН

Information about the Authors

Alexander E. Shchelokov, Master's Degree Student, Department of Computer Technologies, Faculty of Information Technology, Novosibirsk State University, Software Engineer, Institute of Automation and Electrometry SB RAS (Novosibirsk, Russian Federation)

Konstantin I. Budnikov, PhD in Computer Science, Associate Professor, Department of Computer Technologies, Faculty of Information Technologies, Novosibirsk State University, Senior Researcher, Head of the Thematic Group, Institute of Automation and Electrometry SB RAS (Novosibirsk, Russian Federation)

*Статья поступила в редакцию 25.05.2024;
одобрена после рецензирования 10.06.2024; принята к публикации 10.06.2024*

*The article was submitted 25.05.2024;
approved after reviewing 10.06.2024; accepted for publication 10.06.2024*

Правила оформления текста рукописи

Авторы представляют статьи на русском или английском языке объемом от 0,5 авторского листа (20 тыс. знаков) до 1 авторского листа (40 тыс. знаков), включая иллюстрации (1 иллюстрация форматом 190 × 270 мм = 1/6 авторского листа, или 6,7 тыс. знаков). Публикации, превышающие указанный объем, допускаются к рассмотрению только после индивидуального согласования с редакцией журнала.

Текст рукописи должен быть представлен в редколлегию в виде файла MS Word (.doc, .docx). Гарнитура Times New Roman, размер шрифта 11, межстрочный интервал 1, размеры полей – стандартные значения текстового редактора. Форматирование – выравнивание по ширине страницы, переносы слов включены, каждый новый абзац начинается с красной строки. Не допускается ручное форматирование абзацев (пробелами, лишними переводами строк, разрывами страниц).

Структура статьи

- Индекс УДК (универсальной десятичной классификации). Выравнивание по левому краю
- Название статьи. Выравнивание по центру, полужирный шрифт
- ФИО авторов (полностью). Выравнивание по центру, полужирный шрифт
- Места работы всех авторов. Выравнивание по центру, курсив
- Адреса электронной почты, ORCID авторов
- Аннотация статьи
- Ключевые слова, не более 10
- Благодарности, сведения о финансовой поддержке
- Название статьи **на английском языке**. Выравнивание по центру, полужирный шрифт
- ФИО авторов **на английском языке** (полностью). Выравнивание по центру, полужирный шрифт
- Места работы авторов **на английском языке**. Выравнивание по центру, курсив
- Аннотация статьи **на английском языке (Abstract)**, 200–250 слов
- Ключевые слова **на английском языке (Keywords)**, не более 10
- Благодарности, сведения о финансовой поддержке **на английском языке**, если есть соответствующий раздел на русском языке (**Acknowledgements**)
- Основной текст
- Список литературы / **References**
- Сведения об авторах

Требования к оформлению основного текста и иллюстративных материалов

Основной текст должен быть представлен в структурированном виде, рекомендуется использовать подзаголовки – например: Введение, Методика..., Выводы, Результаты, Заключение.

Подзаголовки отделяются и набираются полужирным шрифтом. В целях выделения частей текста и отдельных слов и словосочетаний допускается использование курсива или полужирного шрифта. Подчеркивание, разрядка, изменение основного кегля и выделение цветом не используются.

Иллюстрации к рукописи статьи должны быть приложены в виде отдельных файлов. При этом в тексте должно содержаться включенное изображение с указанием имени файла. Все иллюстрации, содержащие схемы, графики, алгоритмы и т. п., должны быть представлены в векторном виде (.ai, .eps, .cdr). Скриншоты и другие растровые изображения должны быть представлены в максимально высоком качестве, без каких-либо потерь и искажений (.jpg, .tif). Все иллюстрации должны иметь подрисуночную подпись – свое название. Надписи к таблицам и подписи к иллюстрациям приводятся **на двух языках (русском и английском)**.

Примеры:

Рис. 1. Диаграмма производительности...

Fig. 1. Performance diagram...

Таблица 1

Сравнение алгоритмов...

Table 1

Comparison of algorithms...

Нумерация последовательная и неразрывная от начала статьи. Не допускается использование других наименований, кроме «Рис.» / «Fig.», «Таблица» / «Table», и усложнение нумерации (например, «Рис. 3.2.»). Ссылка на иллюстрацию в тексте должна быть приведена в круглых скобках, например: (рис. 1), (табл. 1).

Формулы должны быть набраны с использованием редактора MathType либо встроенного редактора формул MS Word. Кегль основных символов – 11, греческие символы набираются прямым шрифтом, латинские – курсивом. Нумеруются только те формулы, на которые автор ссылается в тексте.

Abstract

Аннотация статьи на английском языке (Abstract) не должна быть дословным переводом русскоязычной аннотации. Раздел Abstract, как и основной текст, должен быть структурирован, в нем должно содержаться описание цели работы, методов исследования, научной значимости, выводов / результатов. Требуется качественный перевод на английский язык (при необходимости просим авторов обращаться к профессиональным переводчикам). **Объем Abstract 200–250 слов.**

Список литературы / References

Список литературы и список литературы на английском языке (References) размещаются в общем разделе. Рекомендуемое количество цитируемых в статье источников – не менее 10, в список желательно включать ссылки на актуальные работы по теме исследования, особенно в иностранных периодических изданиях.

В тексте статьи ссылки на литературу указываются цифрами в квадратных скобках, при необходимости указываются номера страниц, например: [2; 3. С. 15].

Список литературы нумеруется в порядке цитирования и оформляется в соответствии с ГОСТ Р 7.0.5-2008 на библиографическое описание (знаки тире в описании опускаются). Ссылки на неопубликованные работы, а также на Интернет-ресурсы (кроме электронных изданий, поддающихся библиографическому описанию) оформляются в виде сноски.

В Список литературы ссылки на источники следует включать на оригинальном языке опубликования. Каждый источник должен быть также оформлен на английском языке (References) по международному стандарту для публикаций в области информатики IEEE Style со следующими отличиями:

- инициалы авторов указываются после фамилии;
- название статьи не берется в кавычки, отделяется точкой;

- отсутствует союз «and» перед фамилией последнего автора;
- в диапазоне страниц – удвоенная «р» (например, «pp. 2–9»);
- год издания указывается после места издания (для книг) и сразу после названия журнала (для периодики).
- Перевод источника на английский язык:
- если источник имеет выходные данные на английском языке, то для формирования References **следует использовать именно эти данные**;
- если оригинальная публикация не содержит выходных данных на английском языке, то допускается транслитерация названия материала на латинский алфавит в сочетании с переводом на английский язык в квадратных скобках. В конце описания указывается, на каком языке написана эта работа, например, (in Russ.). При транслитерации можно воспользоваться интернет-ресурсом <http://ru.translit.ru/>, рекомендуется выбрать стандарт BSI. Место издания не транслитерируется, указывается полностью на английском языке, например: Moscow. Название издательства / издателя, как правило, транслитерируется. Для журналов, у которых есть официальное название на английском языке, – использовать его (проверить на сайте журнала, или, например, в библиотеке WorldCat), если названия на английском языке нет, использовать транслитерацию по системе BSI. Не следует самостоятельно переводить названия журналов.

Если у цитируемого источника есть **цифровой идентификатор DOI** (<https://search.crossref.org/>), его требуется обязательно указывать в конце библиографической ссылки.

Примеры оформления ссылок. Каждый источник в том же пункте дублируется на английском языке (References).

Источник на русском языке, перевод на английский доступен в метаданных статьи

1. Журавлев С. С., Рудометов С. В., Окольников В. В., Шакиров С. Р. Применение модельно-ориентированного проектирования к созданию АСУ ТП опасных промышленных объектов // Вестник НГУ. Серия: Информационные технологии. 2018. Т. 16, № 4. С. 56–67. DOI 10.25205/1818-7900-2018-16-4-56-67

Zhuravlev S. S., Rudometov S. V., Okolnishnikov V. V., Shakirov S. R. Model-Based Design Approach for Development Process Control Systems of Hazardous Industrial Facilities. *Vestnik NSU. Series: Information Technologies*, 2018, vol. 16, no. 4, pp. 56–67. (in Russ.) DOI 10.25205/1818-7900-2018-16-4-56-67

Источник на английском языке. Оформляем согласно требованиям для References. Приводим только 1 раз.

2. Telnov V. I. Optimization of the Beam Crossing Angle at the ILC for E + e- and yy Collisions. *Journal of Instrumentation*, 2018, vol. 13, no. 03, pp. P03020–P03020. DOI 10.1088/1748-0221/13/03/p03020

Метаданные источника доступны только на русском языке

3. Жижимов О. Л., Федотов А. М., Шокин Ю. И. Технологическая платформа массовой интеграции гетерогенных данных // Вестник НГУ. Серия: Информационные технологии. 2013. Т. 11, вып. 1. С. 24–41.

Zhizhimov O. L., Fedotov A. M., Shokin Yu. I. Tekhnologicheskaya platforma massovoi integratsii geterogennykh dannykh [Technology Platform for the Mass Integration of Heterogeneous Data]. *Vestnik NSU. Series: Information Technologies*, 2013, vol. 11, no. 1, pp. 24–41. (in Russ.)

Сведения об авторах

Последний раздел статьи – информация об авторе / авторах **на русском и английском языках**:

- ФИО полностью, ученая степень, ученое звание;
- идентификаторы автора, такие как ResearcherID (всем авторам рекомендуется использовать данные сервисы для ведения актуального списка своих публикаций);
- контактный телефон (не публикуется).

Если статья представляется на английском языке, необходимо приложить перевод на русский язык названия, аннотации, ключевых слов, сведений об авторе.

Доставка материалов

Материалы предоставляются в редакцию по электронной почте inftech@vestnik.nsu.ru.

Порядок рецензирования

Все статьи сначала проходят проверку на заимствование и только после этого отправляются на рецензирование. Редакционный совет не допускает к публикации материал, если имеется достаточно оснований полагать, что он является плагиатом.

Тип рецензирования статей – двухуровневое, одностороннее анонимное («слепое»).

Для каждой статьи редколлегией выбираются рецензенты, научная деятельность которых связана с темой представленного материала. Ответственный секретарь журнала обращается к ним с просьбой дать экспертную оценку статье либо помочь организовать рецензирование.

Рецензии для журнала «Вестник НГУ. Серия: Информационные технологии» составляются по единой схеме и подразумевают оценку по следующим критериям: соответствие тематике журнала, оригинальность и значимость результатов, качество изложения материала.

Заполненный бланк рецензии высылается на электронный адрес редакции. В зависимости от экспертных заключений статья может быть принята редакционным советом к опубликованию, рекомендована автору к доработке (с последующим повторным рецензированием либо без него) или отклонена (с предоставлением автору мотивированного отказа). Автору на электронный адрес высылается текст рецензии без указания ФИО рецензента и его контактных данных.

Все рецензии хранятся в редакции журнала не менее 5 лет. Редколлегия журнала обязуется при поступлении соответствующего запроса направлять копии рецензий в Министерство науки и высшего образования Российской Федерации.