

Научная статья

УДК 81'33

DOI: <https://doi.org/10.18721/JHSS.16206>

EDN: <https://elibrary/FVLFNN>



ОБЪЯСНИТЕЛЬНЫЕ ВОЗМОЖНОСТИ АГРЕГИРОВАНИЯ ТЕОРЕТИКО-ГРАФОВЫХ МОДЕЛЕЙ СКАЗОЧНЫХ СЮЖЕТОВ В ПРАКТИКЕ ФИЛОЛОГИЧЕСКОГО АНАЛИЗА ТЕКСТА

Н.Д. Москин , А.А. Лебедев 

Петрозаводский государственный университет,
г. Петрозаводск, Российская Федерация

✉ moskin@petrsu.ru

Аннотация. В исследовании устного народного творчества одной из важных задач является классификация и систематизация типов русских сказок. Для ее решения была создана математическая модель, представляющая собой помеченный граф. Она предназначена для выявления центральных событий, которые определяют сюжетную структуру волшебных сказок. Настоящее исследование проведено с использованием алгоритмов агрегирования теоретико-графовых моделей сказочных сюжетов с целью упрощения и обобщения их структуры, а также последующей интерпретации. В рамках апробации предложенной методики были проанализированы 41 сказка (включая повторяющиеся сюжеты) и соответствующие им теоретико-графовые модели. Тексты были взяты из сборника «Народные русские сказки А.Н. Афанасьева». При формировании выборки были приняты во внимание распространенность текста в народной традиции и его типичность для русского фольклора. Каждому тексту ставился в соответствие граф, где вершины обозначают персонажей, объекты и события, а ребра отражают их взаимосвязи. К числу таких взаимоотношений относятся: ситуация дарения, ситуация обмена, ситуация повторения событий. На примере текста «Сказка о молодце-удальце, молодильных яблоках и живой воде», чей сюжет включает множество персонажей и их взаимодействий, показана процедура агрегирования, которая заключается в объединении схожих элементов графа с целью сокращения его сложности без потери информативности. При этом не учитывается упорядоченность вершин и ребер, а также нет заранее заданной структуры простого графа. Метод был реализован в информационной системе «Фольклор». Наилучшие показатели качества достигаются при числе вершин, равном десяти, и пороге значимости показателей связи, равном 0,2.

Ключевые слова: русская сказка, персонажи, теоретико-графовая модель, агрегирование, информационная система «Фольклор».

Для цитирования: Москин Н.Д., Лебедев А.А. Объяснительные возможности агрегирования теоретико-графовых моделей сказочных сюжетов в практике филологического анализа текста // Terra Linguistica. 2025. Т. 16. № 2. С. 85–100. DOI: 10.18721/JHSS.16206



EXPLANATORY POTENTIAL OF AGGREGATING GRAPH-THEORETICAL MODELS OF FAIRY TALE PLOTS IN PHILOLOGICAL TEXT ANALYSIS PRACTICE

N.D. Moskin , A.A. Lebedev 

Petrozavodsk State University, Petrozavodsk, Russian Federation

 moskin@petsu.ru

Abstract. In the study of folklore, one of the important tasks is the classification and systematization of the types of Russian fairy tales. To solve this problem, a mathematical model was created, which is a labeled graph. It is designed to identify the central events that define the plot structure of fairy tales. The present study was conducted using algorithms for aggregating graph-theoretical models of fairy-tale plots in order to simplify and generalize their structure, as well as subsequent interpretation. As part of the testing of the proposed methodology, 41 fairy tales (including recurring plots) and their corresponding graph-theoretical models were analyzed. The texts were taken from “Russian Fairy Tales” book compiled by A. Afanasyev. When forming the sample, the prevalence of the text in folk tradition and its typicality for Russian folklore were taken into account. A graph was assigned to each text, where the vertices denote characters, objects, and events, and the edges reflect their relationships (gifting, exchange, and event repetition). Using the example of the text “The Bold Knight, the Apples of Youth, and the Water of Life”, the plot of which includes many characters and their interactions, shows the aggregation procedure, which is the combining of similar elements of the graph in order to reduce its complexity without loss of information. This does not take into account the ordering of vertices and edges, and there is no predefined structure of a simple graph. The method was implemented in the “Folklore” information system. The best quality indicators are achieved with the number of vertices equal to ten and the significance threshold of communication indicators equal to 0.2.

Keywords: Russian fairy tale, characters, graph-theoretical model, aggregation, “Folklore” information system.

Citation: Moskin N.D., Lebedev A.A., Explanatory potential of aggregating graph-theoretical models of fairy tale plots in philological text analysis practice, *Terra Linguistica*, 16 (2) (2025) 85–100. DOI: 10.18721/JHSS.16206

Введение

Современные подходы к анализу текстов, основанные на искусственном интеллекте и методах автоматической генерации текстов, в сочетании с традиционными методами поиска сходств и различий между разными группами текстов, а также их систематизацией и классификацией, находят широкое применение в гуманитарных науках (например, [1–4]). Одним из перспективных направлений является использование графовых моделей для структурного анализа повествовательных текстов разных типов, включая фольклорные произведения¹.

Одной из ключевых задач в области изучения фольклора является классификация и упорядочивание типов сказок на основе русскоязычных источников. В отечественной науке первой значимой работой в этом направлении стал перевод «Указателя сказочных сюжетов по системе Аарне», выполненный Н.П. Андреевым². Кроме того, важным вкладом в эту область считается издание «Сравнительный указатель сюжетов. Восточнославянская сказка»³.

¹ Москин Н.Д. Теоретико-графовые модели, методы и программные средства интеллектуального анализа текстовой информации на примере фольклорных и литературных произведений: дисс. ... д-ра техн. наук. Петрозаводск, 2022. 370 с.

² Андреев Н.П. Указатель сказочных сюжетов по системе Аарне. Л.: Издательство Государственного русского географического общества, 1929. 120 с.

³ Сравнительный указатель сюжетов. Восточнославянская сказка / сост. Л.Г. Бараг, И.П. Березовский, К.П. Кабашников, Н.В. Новиков; отв. ред. К.В. Чистов. Л.: Наука, 1979. 437 с.



Исследование сказочных сюжетов не ограничивается анализом текстов на одном языке, но также предполагает международный подход, включая сопоставление фольклорных произведений разных культур. Вопросы, связанные с передачей сказок между народами, а также сложности в разработке точных методов их изучения, подробно рассмотрены в работах В.М. Жирмунского [5].

Не менее значимой задачей является систематическое описание структуры фольклорного текста. В.Я. Пропп, исследуя волшебные сказки, подчеркивал, что «о происхождении какого бы то ни было явления можно говорить лишь после того, как явление это описано» [6, с. 11]. Он отмечал, что ранее «изучение сказки велось, главным образом, лишь генетически, большей частью без попыток предварительного систематического описания», а «правильная классификация — одна из первых ступеней научного описания» [6, с. 11–12].

Теоретико-графовые методы позволяют формализовать сюжеты сказок, представляя их в виде графов, где вершины соответствуют персонажам, объектам и событиям, а ребра — их взаимосвязям. Такой подход позволяет исследовать сказочные нарративы с точки зрения структурных повторений, архетипических паттернов и закономерностей развития сюжета. На важность подобного структурного описания указывал еще Ю.М. Лотман, который отметил: «Рассматривая содержание художественного текста только на уровне языкового сообщения, мы проходим мимо сложной системы значений, создаваемых собственно художественной структурой» [7, с. 63].

Разработанная графовая модель [8] направлена на выявление ключевых событий, которые формируют сюжетную линию волшебных сказок. Известно, что такие произведения обладают четкой структурой⁴, которую можно формализовать, чтобы сравнить различные тексты между собой и выделить наиболее важные элементы их повествования. Основой данной модели являются персонажи сказки, а также их взаимодействия, которые определяют развитие сюжета. Вершины графа представляют собой персонажей и объекты, разделенные на категории (классификация основана на работе [9]).

Отметим, что при анализе больших по объему текстов теоретико-графовые модели становятся перегруженными, количество вершин и связей многократно возрастает, и их уже сложно исследовать с использованием существующих алгоритмов теории графов и вычислительных мощностей. Поэтому возникает задача агрегирования, т.е. задача нахождения такого простого графа, вершины которого соответствуют подмножествам вершин исходного графа, а ребра — «основным потокам связей» (или «сильным связям») на исходном графе⁵. При этом число вершин в агрегированном графе будет значительно меньше, чем в исходном.

Данная работа посвящена методам агрегирования теоретико-графовых моделей сказочных сюжетов с целью их упрощения и обобщения структуры, а затем их интерпретации.

Теоретико-графовая модель сказки

Для апробации методики построения теоретико-графовых моделей были проанализированы 41 сказка (включая повторяющиеся сюжеты) и их теоретико-графовые модели (список текстов приведен в табл. 1). Для формирования выборки были использованы тексты, входящие в первый том сборника «Народные русские сказки А. Н. Афанасьева» [10], написанные на русском языке. При отборе материала учитывались такие критерии, как распространенность текста в народной традиции и типичность для русского фольклора. Анализу были подвергнуты преимущественно сказки с повторяющимися сюжетами, так как вопрос выявления сходств и различий между вариантами таких сюжетов является объектом интереса исследователей-фольклористов. В частности, подобные исследования позволяют проследить эволюцию сюжетных линий, выявить

⁴ Зубов А.В., Зубова И.И. Основы искусственного интеллекта для лингвистов: учебное пособие. М.: Университетская книга; Логос, 2007. 320 с.

⁵ Москин Н.Д. Теоретико-графовые модели, методы и программные средства интеллектуального анализа текстовой информации на примере фольклорных и литературных произведений: дисс. ... д-ра техн. наук. Петрозаводск, 2022. 370 с.



Таблица 1. Описание выборки текстов
Table 1. Description of the text selection

Номер текста	Название сказки	Число вершин	Источник
93	Ведьма и Солнцева сестра	16	[10, с. 136–138]
95	Морозко	7	[10, с. 140–143]
96	Морозко	9	[10, с. 143–144]
97	Старуха-говоруха	6	[10, с. 145]
98	Дочь и падчерица	8	[10, с. 146–147]
100	Крошечка-хаврошечка	9	[10, с. 150–151]
101	Буренушка	12	[10, с. 152–154]
102	Баба-яга	8	[10, с. 155–156]
103	Баба-яга	15	[10, с. 156–158]
104	Василиса Прекрасная	10	[10, с. 159–165]
105	Баба-Яга и заморышек	14	[10, с. 166–168]
106	Баба-яга и жихарь	7	[10, с. 169–170]
107	Баба-яга и жихарь	7	[10, с. 170–172]
108	Ивашко и ведьма	10	[10, с. 173–176]
111	Ивашко и ведьма	8	[10, с. 180–182]
112	Терешечка	9	[10, с. 183–184]
113	Гуси-лебеди	10	[10, с. 185–186]
115	Правда и кривда	11	[10, с. 191–195]
116	Правда и кривда	16	[10, с. 195–196]
117	Правда и кривда	8	[10, с. 196–197]
118	Правда и кривда	8	[10, с. 197–198]
119	Правда и кривда	7	[10, с. 199–200]
120	Правда и кривда	10	[10, с. 200–201]
121	Правда и кривда	8	[10, с. 201–203]
122	Правда и кривда	6	[10, с. 203–204]
127	Купеческая дочь и служанка	10	[10, с. 224–227]
131	Фролка-сидень	11	[10, с. 244–246]
140	Зорька, вечерка и полуночка	13	[10, с. 299–302]
148	Никита Кожемяка	6	[10, с. 327–328]
162	Хрустальная гора	11	[10, с. 393–394]
169	Жар-птица и Василиса-царевна	8	[10, с. 424–426]
170	Жар-птица и Василиса-царевна	17	[10, с. 427–430]
171	Сказка о молодце-удальце, молодильных яблоках и живой воде	20	[10, с. 431–435]
172	Сказка о молодце-удальце, молодильных яблоках и живой воде	15	[10, с. 436–438]
173	Сказка о молодце-удальце, молодильных яблоках и живой воде	15	[10, с. 438–442]
174	Сказка о молодце-удальце, молодильных яблоках и живой воде	18	[10, с. 442–444]
175	Сказка о молодце-удальце, молодильных яблоках и живой воде	21	[10, с. 444–448]
176	Сказка о молодце-удальце, молодильных яблоках и живой воде	39	[10, с. 448–455]
177	Сказка о молодце-удальце, молодильных яблоках и живой воде	11	[10, с. 455–457]
178	Сказка о молодце-удальце, молодильных яблоках и живой воде	10	[10, с. 457–460]

региональные особенности и понять механизмы трансформации устного нарратива, в том числе и в аспекте дискурсивной отнесенности (см. [11]). Кроме того, сравнительный анализ вариантов



одного сюжета помогает уточнить их жанровую специфику и определить степень влияния различных факторов на формирование окончательной версии текста.

Рассмотрим методику построения теоретико-графовой модели на примере конкретного сказочного сюжета. В качестве примера проанализируем текст «Сказка о молодце-удальце, молодильных яблоках и живой воде» [10, с. 448–455], чей сюжет включает большое количество персонажей и связей между ними (табл. 2). Рассмотрим более подробно персонажи и их типы:

- Родственник – царь (Т1), первая царевна (Т2), вторая царевна (Т3), третья царевна (Т4), другой царь (Т5), дети третьего сына и Елены Прекрасной (Т6);
- Глупец – старший сын (G1), средний сын (G2);
- Герой – Иван-царевич (Н1);
- Антигерой – Елена Прекрасная (А1), ведьма (А2), осьмiglавый змей (А3), другой змей (А4), двенадцатиглавый змей (А5);
- Награда – лошадь (N1), латы и меч мертвого богатыря (N2), первый сокол (N3), второй сокол (N4), живая и мертвая вода (N5), портрет Елены Прекрасной (N6), молоджавые яблоки (N7), платочек (N8), лошадь, латы и меч (N9), золотое кольцо первой царевны (N10), золотое кольцо второй царевны (N11), золотое кольцо третьей царевны (N12);
- Помощник – старуха (П1), первая прекрасная девушка (П2), вторая прекрасная девушка (П3), красная девица (П4), еще одна прекрасная девушка (П5) еще одна старуха (П6), сокол (П7);
- Даритель – мертвый богатырь (Д1), первая сестра Елены Прекрасной (Д2), вторая сестра Елены Прекрасной (Д3);
- Препятствие – шнуры (R1), море (R2);
- Антипомощник – плешивый Макарка (V1).

Связи между объектами и их краткое описание представлены в табл. 2.

Таблица 2. Связи между вершинами фрагмента текста
«Сказка о молодце-удальце, молодильных яблоках и живой воде»
Table 2. Connections between the vertices of the text fragment
“The Tale of a brave young man, rejuvenating apples and living water”

	Вершина 1	Вершина 2	Событие в тексте
1	Т1	G1	Царь благословляет и отправляет двух старших сыновей
1	Т1	G2	Царь благословляет и отправляет двух старших сыновей
2	Т1	Н1	Царь отправляет третьего сына Ивана-царевича
3	Н1	П1	Иван-царевич беседует со старухой
4	Д1	Н1	Мертвый богатырь беседует с Иваном-царевичем
5	Н1	N1	Иван-царевич забирает лошадь мертвого богатыря
6	Н1	N2	Иван-царевич забирает латы и меч мертвого богатыря
7	П2	Н1	Прекрасная девушка беседует с Иваном-царевичем
8	Н1	N3	Иван-царевич берет летучего сокола у первой прекрасной девушки
8	Н1	П2	Иван-царевич берет летучего сокола у первой прекрасной девушки
8	П2	N3	Иван-царевич берет летучего сокола у первой прекрасной девушки
9	Н1	N1	Иван-царевич оставляет коня первой прекрасной девушке
9	Н1	П2	Иван-царевич оставляет коня первой прекрасной девушке
9	П2	N1	Иван-царевич оставляет коня первой прекрасной девушке
10	Н1	П3	Иван-царевич беседует со второй прекрасной девушкой
11	Н1	N4	Иван-царевич берет второго летучего сокола у второй прекрасной девушки
11	Н1	П3	Иван-царевич берет второго летучего сокола у второй прекрасной девушки



Продолжение таблицы 2

	Вершина 1	Вершина 2	Событие в тексте
11	П3	N4	Иван-царевич берет второго летучего сокола у второй прекрасной девушки
12	H1	N3	Иван-царевич оставляет первого летучего сокола у второй прекрасной девушки
12	H1	П3	Иван-царевич оставляет первого летучего сокола у второй прекрасной девушки
12	П3	N3	Иван-царевич оставляет первого летучего сокола у второй прекрасной девушки
13	H1	A1	Иван-царевич прилетает к Елене Прекрасной
14	H1	N5	Иван-царевич берет живую и мертвую воду и портрет Елены Прекрасной
14	H1	N6	Иван-царевич берет живую и мертвую воду и портрет Елены Прекрасной
15	H1	A1	Иван-царевич насилует Елену Прекрасную
16	H1	N7	Иван-царевич берет молоджавые яблоки
17	H1	N4	Иван-царевич садится на сокола
18	H1	R1	Иван-царевич цепляется за шнуры
19	H1	П3	Иван-царевич прилетает ко второй прекрасной девушке
20	H1	N4	Иван-царевич меняет второго сокола на первого сокола, оставляя второго сокола второй прекрасной девушке
20	H1	N3	Иван-царевич меняет второго сокола на первого сокола, оставляя второго сокола второй прекрасной девушке
20	П3	N3	Иван-царевич меняет второго сокола на первого сокола, оставляя второго сокола второй прекрасной девушке
20	П3	N4	Иван-царевич меняет второго сокола на первого сокола, оставляя второго сокола второй прекрасной девушке
20	H1	П3	Иван-царевич меняет второго сокола на первого сокола, оставляя второго сокола второй прекрасной девушке
21	A1	П3	Елена Прекрасная говорит со второй прекрасной девушкой
22	H1	П2	Иван-царевич прилетает к первой прекрасной девушке
23	H1	N3	Иван-царевич меняет первого сокола на коня, оставляя первого сокола первой прекрасной девушке
23	H1	N2	Иван-царевич меняет первого сокола на коня, оставляя первого сокола первой прекрасной девушке
23	П2	N2	Иван-царевич меняет первого сокола на коня, оставляя первого сокола первой прекрасной девушке
23	П2	N3	Иван-царевич меняет первого сокола на коня, оставляя первого сокола первой прекрасной девушке
23	H1	П2	Иван-царевич меняет первого сокола на коня, оставляя первого сокола первой прекрасной девушке
24	A1	П2	Елена Прекрасная говорит с первой прекрасной девушкой
25	H1	П1	Иван-царевич приезжает к старухе
26	П1	H1	Старуха дает Ивану-царевичу платочек
26	П1	N8	Старуха дает Ивану-царевичу платочек
26	H1	N8	Старуха дает Ивану-царевичу платочек
27	A1	П1	Елена Прекрасная разговаривает со старухой
28	A1	H1	Елена Прекрасная начинает догонять Ивана-царевича
29	H1	N8	Иван-царевич использует платочек
30	H1	R2	Иван-царевич создает море
31	A1	R2	Елена Прекрасная не может преодолеть море
32	A1	H1	Елена Прекрасная кричит Ивану-царевичу
33	П4	H1	Красная девица беседует с Иваном-царевичем
34	A2	H1	Ведьма пытается уложить Ивана-царевича к стенке



Продолжение таблицы 2

	Вершина 1	Вершина 2	Событие в тексте
35	H1	A2	Иван-царевич сбрасывает ведьму в погреб
36	H1	П4	Иван-царевич берет с собой красную девицу
37	H1	G1	Иван-царевич беседует с братьями
37	H1	G2	Иван-царевич беседует с братьями
38	H1	G1	Иван-царевич передает братьям живую и мертвую воду и молоджавые яблоки
38	H1	G2	Иван-царевич передает братьям живую и мертвую воду и молоджавые яблоки
38	H1	N5	Иван-царевич передает братьям живую и мертвую воду и молоджавые яблоки
38	H1	N7	Иван-царевич передает братьям живую и мертвую воду и молоджавые яблоки
39	G1	H1	Братья опускают Ивана-царевича в яму
39	G2	H1	Братья опускают Ивана-царевича в яму
40	H1	П5	Иван-царевич добывает прекрасную девушку
41	G1	П5	Братья вытаскивают девушку
41	G2	П5	Братья вытаскивают девушку
42	H1	П6	Иван-царевич беседует со старой старухой
43	П6	H1	Старая старуха дает Ивану-царевичу лошадь, латы и меч
43	П6	N9	Старая старуха дает Ивану-царевичу лошадь, латы и меч
43	H1	N9	Старая старуха дает Ивану-царевичу лошадь, латы и меч
44	H1	T2	Иван-царевич приезжает к первой царевне
45	T2	H1	Первая царевна будит Ивана-царевича
46	A3	H1	Осьмиглавый змей разговаривает с Иваном-царевичем
47	H1	A3	Иван-царевич дерется с осьмиглавым змеем
48	H1	T2	Иван-царевич прощается с царевной
49	T2	H1	Первая царевна дает Ивану-царевичу золотое кольцо
49	T2	N10	Первая царевна дает Ивану-царевичу золотое кольцо
49	H1	N10	Первая царевна дает Ивану-царевичу золотое кольцо
50	V1	T2	Макарка беседует с первой царевной
51	V1	T2	Макарка приводит первую царевну к другому царю
51	V1	T5	Макарка приводит первую царевну к другому царю
51	T2	T5	Макарка приводит первую царевну к другому царю
52	V1	T5	Макарка беседует с другим царем
53	V1	T3	Макарка отвозит вторую царевну
54	H1	T3	Иван-царевич приезжает ко второй царевне
55	T3	H1	Вторая царевна будит Ивана-царевича
56	A4	H1	Другой змей разговаривает с Иваном-царевичем
57	H1	A4	Иван-царевич дерется с другим змеем
58	T3	H1	Вторая царевна дает Ивану-царевичу золотое кольцо
58	T3	N11	Вторая царевна дает Ивану-царевичу золотое кольцо
58	H1	N11	Вторая царевна дает Ивану-царевичу золотое кольцо
59	H1	П6	Иван-царевич уезжает к старухе
60	V1	T3	Макарка приводит вторую царевну к другому царю
60	V1	T5	Макарка приводит вторую царевну к другому царю
60	T3	T5	Макарка приводит вторую царевну к другому царю
61	V1	T5	Макарка беседует с другим царем
62	V1	T4	Макарка отвозит третью царевну
63	H1	T4	Иван-царевич приезжает к третьей царевне



Окончание таблицы 2

	Вершина 1	Вершина 2	Событие в тексте
64	T4	H1	Третья царевна будит Ивана-царевича
65	A5	H1	Двенадцатиглавый змей разговаривает с Иваном-царевичем
66	H1	A5	Иван-царевич дерется с двенадцатиглавым змеем
67	N9	H1	Конь дает подсказки Ивану-царевичу
68	H1	A5	Иван-царевич убивает двенадцатиглавого змея
69	T4	H1	Третья царевна дает Ивану-царевичу золотое кольцо
69	T4	N12	Третья царевна дает Ивану-царевичу золотое кольцо
69	H1	N12	Третья царевна дает Ивану-царевичу золотое кольцо
70	H1	P6	Иван-царевич уезжает к старухе
71	V1	T4	Макарка приводит третью царевну к другому царю
71	V1	T5	Макарка приводит третью царевну к другому царю
71	T4	T5	Макарка приводит третью царевну к другому царю
72	V1	T5	Макарка беседует с другим царем
73	H1	T2	Иван-царевич пьет водку у первой царевны
74	H1	N10	Иван-царевич бросает первое золотое кольцо
75	H1	T3	Иван-царевич пьет водку у второй царевны
76	H1	N11	Иван-царевич бросает второе золотое кольцо
77	H1	T4	Иван-царевич пьет водку у третьей царевны
78	H1	N12	Иван-царевич бросает третье золотое кольцо
79	T2	H1	Царевны признают Ивана-царевича
79	T3	H1	Царевны признают Ивана-царевича
79	T4	H1	Царевны признают Ивана-царевича
80	T5	V1	Другой царь приказывает расстрелять Макарку
81	H1	T5	Иван-царевич договаривается с другим царем о возвращении наверх
82	T5	P7	Другой царь зовет сокола
83	P7	H1	Сокол поднимает Ивана-царевича
84	H1	T1	Иван-царевич приходит к царю
85	T1	G1	Царь отправляет старших сыновей к Елене Прекрасной
85	T1	G2	Царь отправляет старших сыновей к Елене Прекрасной
86	T6	G1	Сыновья Елены Прекрасной избивают старших сыновей
86	T6	G2	Сыновья Елены Прекрасной избивают старших сыновей
87	T1	H1	Царь отправляет Ивана-Царевича к Елене Прекрасной
88	T6	H1	Сыновья Елены Прекрасной приводят Ивана-Царевича к Елене Прекрасной
89	A1	H1	Елена Прекрасная выходит замуж за Ивана-царевича
90	H1	T1	Иван-царевич рассказывает отцу о событиях
91	T1	G1	Царь приказывает расстрелять старших братьев
91	T1	G2	Царь приказывает расстрелять старших братьев
92	H1	A1	Иван-царевич счастливо живет с Еленой Прекрасной

На рис. 1 представлена теоретико-графовая модель фрагмента сказки, начиная с фразы «Вот старшие братья пошли к своему отцу домой; а Иван-царевич в той яме так слезьми и обливается...» и заканчивая словами «...Иван-царевич оторвал свои икры и бросил ему, сокол съел и вылетел с царевичем на верхний свет; тут сокол кашлянул и выкинул его икры и платье».

Выбор конкретного фрагмента объясняется общей структурой повествования – в проанализированной части сказки герой путешествует по «нижнему свету», который противопоставляется

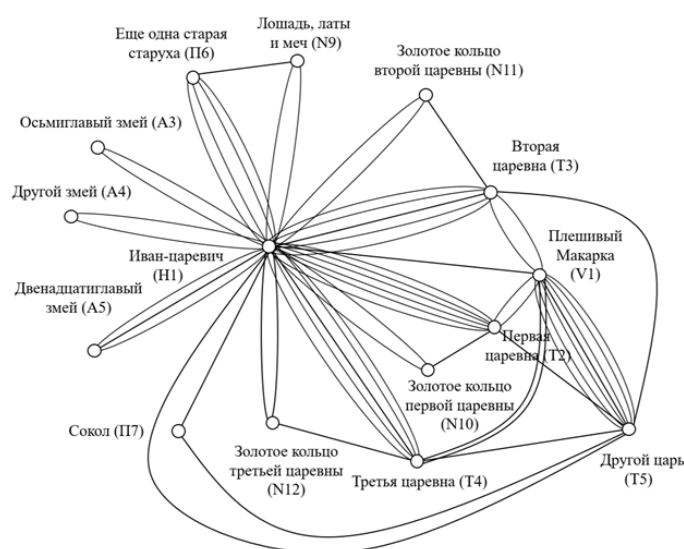


Рис. 1. Теоретико-графовая модель сказочного текста

Fig. 1. Graph-theoretical model of a fairy tale text

«белому свету», в котором проходит остальная часть сказки. Единственным персонажем, который встречается и в сюжете «белого света», и в сюжете «нижнего света», является Иван-царевич, главный герой сказки, в то время как остальные участники повествования находятся только в каком-то одном из миров, что делает правомерным обособленное рассмотрение событий, которые происходят в сказке в пределах «нижнего света».

Задача агрегирования

Алгоритмы агрегирования графовых моделей способны стать эффективным инструментом для лингвистического анализа, позволяя выделить повторяющиеся паттерны взаимодействия персонажей и объектов. Агрегирование теоретико-графовых моделей предполагает объединение схожих элементов графа с целью сокращения его сложности без потери информативности. Некоторые типичные взаимоотношения между персонажами, описываемые в тексте сказки, могут быть унифицированы, так как представляют собой повторяющиеся структуры. К числу таких взаимоотношений относятся:

1. Взаимоотношения дарения. Один персонаж дарит второму персонажу некий предмет. Всегда 3 вершины, каждая из которых связана друг с другом.

Примеры:

Царевна дает Ивану-царевичу золотое кольцо;

Старуха дает Ивану-царевичу платочек.

2. Взаимоотношения обмена. Два персонажа меняются предметами. Всегда 4 вершины и 6 связей; сама операция дробится на два взаимоотношения дарения.

Примеры:

Иван-царевич берет летучего сокола у первой прекрасной девушки. Иван-царевич оставляет коня первой прекрасной девушке;

Иван-царевич берет второго летучего сокола у второй прекрасной девушки. Иван-царевич оставляет первого летучего сокола у второй прекрасной девушки.

3. Взаимоотношения, основанные на повторении событий. Некоторые события могут повторяться с персонажем неоднократно (число связей при этом может быть разным). К примеру, Иван-царевич, поочередно спасая трех царевен, трижды совершает одни и те же действия:



*Иван-царевич приезжает к царевне;
Царевна будит Ивана-царевича перед битвой;
Иван-царевич беседует со змеем;
Иван-царевич побеждает змея;
Иван-царевич берет кольцо;
Иван-царевич прощается с царевной и уезжает.*

При этом меняются царевны (*первая царевна, вторая царевна, третья царевна*), змеи (*осьмиглавый змей, другой змей, двенадцатиглавый змей*) и кольца (*золотое кольцо первой царевны, золотое кольцо второй царевны, золотое кольцо третьей царевны*), но сам порядок действий Ивана-царевича остается неизменным.

Таким образом, в результате агрегирования может получиться более компактная, но по-прежнему информативная модель, которая позволяет анализировать и сопоставлять сказочные сюжеты на более высоком уровне обобщения. Это особенно полезно для сравнительного анализа различных сказок и выявления общих нарративных структур между разными группами текстов.

Группы при агрегировании формируются на основании схожести взаимоотношений (персонажи объединяются, если они участвуют в однотипных взаимодействиях или если с ними происходят одни и те же события), а также с учетом значимости связей (например, если речь идет о дарителе и предмете, преподносимом в дар).

Унификация упомянутых взаимоотношений помогает идентифицировать типичные сюжетные схемы, присутствующие в русских волшебных сказках. Подобный подход упрощает визуализацию сложных текстовых структур (в том случае, когда графовая модель перегружена вершинами и связями, она становится малофункциональной и трудной для восприятия), а также способствует выявлению закономерностей, не являющихся непосредственным объектом рассмотрения традиционного фольклорного анализа (к примеру, иерархия значимости персонажей по их положению в структуре графа, зависимость между количеством вершин-объектов и сложностью сюжетной линии, корреляция между сложностью графа и жанровыми особенностями текста).

Результаты исследования

В качестве эксперимента был выбран алгоритм агрегации, который не учитывает упорядоченность вершин и ребер, а также не имеет заранее заданной структуры простого графа. Он был реализован в информационной системе «Фольклор»⁶. Для фрагмента текста «Сказка о молодце-удальце, молодильных яблоках и живой воде» были построены 30 агрегированных структур, где были заранее заданы число вершин (5, 6, 7, 8, 9, 10, 11, 12, 13 или 14), а также порог значимости показателей связи α (0,1, 0,2 или 0,3). На рис. 2 представлен пример агрегированного графа для $\alpha = 0,2$ (10 вершин).

В табл. 3 представлена матрица смежности агрегированного графа, которая отражает значимость связей между вершинами. Например, между объектами *плешивый Макарка* (V1) и *другой царь* (T5) существует наиболее значимая связь (значение 5,8), что отражает их важное взаимодействие в сюжете: именно Макарка на протяжении всего отрывка текста приводит к царю спасенных Иваном-царевичем царевен, выдавая подвиги главного героя за свои, однако в итоге получает заслуженное наказание.

Связи между объектами *Иван-царевич* (H1) и группами, в которые включаются царевны, их кольца, а также змеи показывают, как главный герой взаимодействует с ключевыми персонажами и объектами.

⁶ Москин Н.Д. Теоретико-графовые модели, методы и программные средства интеллектуального анализа текстовой информации на примере фольклорных и литературных произведений: дисс. ... д-ра техн. наук. Петрозаводск, 2022. 370 с.

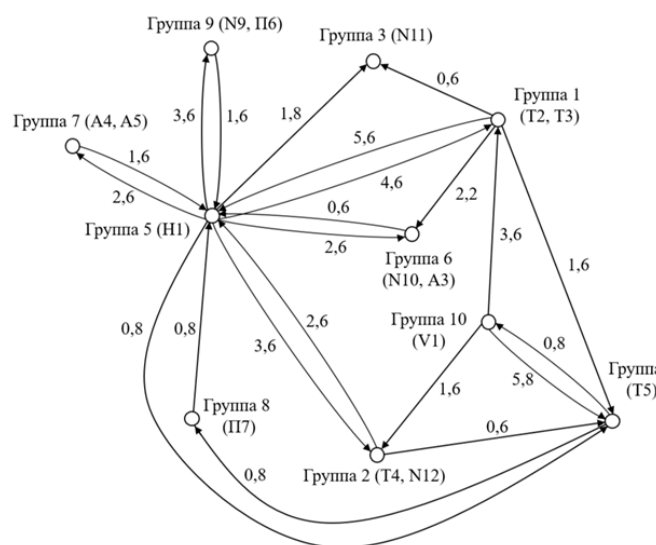


Рис. 2. Агрегированный граф «сильных» связей (10 вершин)

Fig. 2. Aggregated graph of “strong” connections (10 vertices)

Таблица 3. Матрица смежности агрегированного графа (10 вершин)

Table 3. Matrix of adjacency of an aggregated graph (10 vertices)

	T2+T3	T4+N12	N11	T5	H1	A3+N10	A4+A5	П7	N9+П6	V1
T2+T3	1,2	0	0,6	1,6	5,6	2,2	0	0	0	0
T4+N12	0	2,2	0	0,6	2,6	0	0	0	0	0
N11	0	0	0,8	0	0	0	0	0	0	0
T5	0	0	0	0,8	0	0	0	0,8	0	0,8
H1	4,6	3,6	1,8	0,8	0,8	2,6	2,6	0	3,6	0
A3+N10	0	0	0	0	0,6	1,2	0	0	0	0
A4+A5	0	0	0	0	1,6	0	1,2	0	0	0
П7	0	0	0	0	0,8	0	0	0,8	0	0
N9+П6	0	0	0	0	1,6	0	0	0	2,2	0
V1	3,6	1,6	0	5,8	0	0	0	0	0	0,8

Наименее «сильная» связь (значение 0,6) наблюдается между группами, что указывает на меньшую значимость взаимодействий:

- T2 (первая царевна) + T3 (вторая царевна) и N11 (золотое кольцо второй царевны);
- T4 (третья царевна) + N12 (золотое кольцо третьей царевны) и T5 (другой царь);
- A3 (осьмиглавый змей) + N10 (золотое кольцо первой царевны) и H1 (Иван-царевич).

На теплокарте, изображенной на рис. 4, представлены экспертные оценки, насколько тот или иной граф похож на «правильный», построенный специалистом-филологом. Филологическая логика построения агрегированного графа основывается на повторяющихся паттернах изложения, которые можно формализовать и объединить (например, у каждой из трех царевен есть принадлежащее ей кольцо, которое она дарит Ивану-царевичу), функциях персонажей в сюжете, которые можно обобщить (например, в отрывке представлены три змея, каждого из которых побеждает Иван-царевич), и неразрывности связи между отдельными персонажами и объектами, когда это напрямую связано с функцией персонажа (к примеру, наиболее тесна связь между дарителем (Еще одна старуха) и предметом, который даритель преподносит в дар



(Лошадь, латы и меч), потому что именно с целью дарения персонаж и появляется в тексте произведения).

Экспертная оценка выставлялась одним специалистом-филологом с учетом вышеуказанных представлений об объединении вершин на материале фрагмента текста сказки «Сказка о молодце-удальце, молодильных яблоках и живой воде». При этом за ошибку экспертом признавалось: 1) объединение вершин в одну группу, которые не должны находиться в одной группе; 2) отсутствие объединения вершин в одну группу, если они должны в ней находиться. Значения в таблице меняются от 0 до 1 (чем ближе к единице — тем лучше качество, а чем ближе к нулю — тем хуже). Методика оценки заключается в следующем: максимальный 1 балл получает та структура графа, которая соответствует графу, построенному вручную с учетом выделенных ранее и указанных в разделе «3. Задача агрегирования» типичных взаимоотношений дарения, обмена и повторения событий. 0,04 балла снимается за каждый случай, когда в одну группу объединяются те объекты, которые не должны находиться в одной группе. 0,04 балла снимается за каждый случай, когда в одну группу не объединяются те объекты, которые должны быть в одной группе.

Рассмотрим механизм оценивания на примере двух конкретных теоретико-графовых моделей. В первом из них количество вершин равно 10 и $\alpha = 0,2$ (см. рис. 2). Этот граф содержит в себе три неточности:

- 1) вершина T2 (*первая царевна*) ошибочно объединяется с вершиной T3 (*вторая царевна*);
- 2) вершина T2 (*первая царевна*) ошибочно не объединяется с вершиной N10 (*золотое кольцо первой царевны*);
- 3) вершина T3 (*вторая царевна*) ошибочно не объединяется с вершиной N11 (*золотое кольцо второй царевны*).

Таким образом, данный граф получает оценку $1 - 3 \times 0,04 = 0,88$.

Во втором графе (5 вершин, $\alpha = 0,1$), изображенном на рис. 3, обнаруживается значительно большее количество неточностей (матрица смежности представлена в табл. 4):

- 1) вершина N10 (*золотое кольцо первой царевны*) ошибочно объединяется с вершиной N11 (*золотое кольцо второй царевны*);
- 2) вершина N10 (*золотое кольцо первой царевны*) ошибочно не объединяется с вершиной T2 (*первая царевна*);
- 3) вершина N11 (*золотое кольцо второй царевны*) ошибочно не объединяется с вершиной T3 (*вторая царевна*);
- 4) вершина T2 (*первая царевна*) ошибочно объединяется с вершиной T3 (*вторая царевна*);
- 5) вершина A3 (*осьмиглавый змей*) ошибочно объединяется с вершиной N10 (*золотое кольцо первой царевны*);
- 6) вершина A3 (*осьмиглавый змей*) ошибочно объединяется с вершиной N11 (*золотое кольцо второй царевны*);
- 7) вершина A4 (*другой змей*) ошибочно объединяется с вершиной N9 (*лошадь, латы и меч*);
- 8) вершина A4 (*другой змей*) ошибочно объединяется с вершиной П6 (*еще одна старуха*);
- 9) вершина A4 (*другой змей*) ошибочно объединяется с вершиной П7 (*сокол*);
- 10) вершина A5 (*двенадцатиглавый змей*) ошибочно объединяется с вершиной N9 (*лошадь, латы и меч*);
- 11) вершина A5 (*двенадцатиглавый змей*) ошибочно объединяется с вершиной П6 (*еще одна старуха*);
- 12) вершина A5 (*двенадцатиглавый змей*) ошибочно объединяется с вершиной П7 (*сокол*);
- 13) вершина N9 (*лошадь, латы и меч*) ошибочно объединяется с вершиной П7 (*сокол*);
- 14) вершина П6 (*еще одна старуха*) ошибочно объединяется с вершиной П7 (*сокол*);
- 15) вершина T4 (*третья царевна*) ошибочно объединяется с вершиной T5 (*другой царь*);

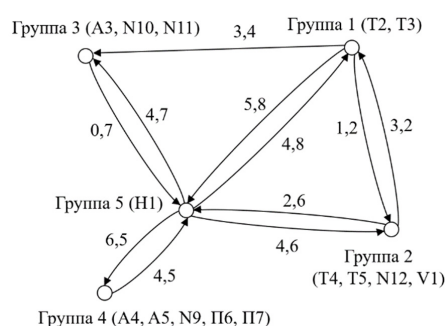


Рис. 3. Агрегированный граф «сильных» связей (5 вершин)

Fig. 3. Aggregated graph of “strong” connections (5 vertices)

Таблица 4. Матрица смежности агрегированного графа (5 вершин)

Table 4. Matrix of adjacency of an aggregated graph (5 vertices)

	T2+T3	T4+T5+N12+V1	A3+N10+N11	A4+A5+N9+П6+П7	H1
T2+T3	1,6	1,2	3,4	0	5,8
T4+T5+N12+V1	3,2	13,4	0	0	2,6
A3+N10+N11	0	0	2,1	0	0,7
A4+A5+N9+П6+П7	0	0	0	3,5	4,5
H1	4,8	4,6	4,7	6,5	0,9

16) вершина T4 (*третья царевна*) ошибочно объединяется с вершиной V1 (*плешивый Макарка*);

17) вершина T5 (*другой царь*) ошибочно объединяется с вершиной N12 (*золотое кольцо третьей царевны*);

18) вершина T5 (*другой царь*) ошибочно объединяется с вершиной V1 (*плешивый Макарка*);

19) вершина N12 (*золотое кольцо третьей царевны*) ошибочно объединяется с вершиной V1 (*плешивый Макарка*).

Сильные связи (13,4) внутри группы «T4+T5+N12+V1» свидетельствуют о тесном взаимодействии этих объектов, но с точки зрения восприятия текста такое объединение некорректно (например, кольцо третьей царевны и плешивый Макарка ни при каких условиях не должны объединяться в одну группу, так как это не следует из содержания текста).

Таким образом, данный граф получает оценку $1 - 19 \times 0,04 = 0,24$. Низкое качество этого графа объясняется избыточным объединением разнородных вершин.

На рис. 4 видно, что наилучшие показатели качества достигаются при числе вершин, равном 10, т.е. 5 вершин могут быть при агрегации объединены с другими. Если говорить о пороге значимости показателей связи, то наилучшие значения получаются при $\alpha = 0,2$ (для любого фиксированного числа вершин). При этом если для 10 вершин изменить это значение на 0,1 в большую или меньшую сторону, то качество меняется незначительно.

Разработанная методика агрегации графовых моделей позволяет не только формализовать сюжетные линии, но и выявить новые аспекты организации фольклорного текста. В частности, анализ «сильных» связей между персонажами и объектами показал, что ключевые нарративные компоненты (например, дарители, помощники, антигерои) могут образовывать устойчивые кластеры. Это способно дополнить существующие лингвистические представления о роли повторяющихся структур в формировании смыслового единства сказки и их влиянии на восприятие фольклорного текста.



0,3	0,04	0,4	0,44	0,68	0,76	0,84	0,8	0,8	0,76	0,72
0,2	0,28	0,52	0,6	0,68	0,84	0,88	0,84	0,8	0,76	0,72
0,1	0,24	0,36	0,6	0,68	0,76	0,84	0,8	0,8	0,76	0,72
	5	6	7	8	9	10	11	12	13	14

Рис. 4. Тепловая карта качества процедуры агрегации для моделей от 5 до 14 вершин
Fig. 4. Heatmap of the quality of the aggregation procedure for models from 5 to 14 vertices

Заключение

Проведенное исследование показало, что структуру сказочных сюжетов можно представлять и анализировать с помощью теоретико-графовых моделей. Для решения задачи формализации сюжетной линии волшебных сказок последовательно реализованы следующие этапы:

1. Выделение персонажей и объектов, которые разбиваются на категории с учетом их роли в тексте; установление взаимоотношений между этими персонажами и объектами;
2. Построение теоретико-графовой модели, где каждый персонаж или объект — это вершина, а взаимодействие — это ребро, соединяющее вершины;
3. Агрегация вершин, позволяющая объединить вершины в группы с учетом значимости связей, представленных в теоретико-графовой модели;
4. Анализ корректности получившихся в ходе агрегации данных.

Подобный подход позволяет перейти от детального описания к обобщенной модели, сохраняя основные сюжетные линии и упрощая сравнительный анализ сказок. Проведенное исследование позволяет продемонстрировать, что применение теоретико-графовых моделей и их агрегации выявляет универсальные структуры, характерные для русских волшебных сказок. Определение и систематизация таких структур расширяют представления о структурных закономерностях фольклорных текстов. В прикладном аспекте рассмотренный в статье метод может быть использован для автоматизированного анализа и классификации сказочных сюжетов.

Алгоритмы агрегации с поиском наилучшей структуры «простого» графа позволяют выявить наиболее значимые («сильные») связи в теоретико-графовой модели, когда некоторые подмножества вершин объединяются в новые обобщенные вершины. На примере текста «Сказка о молодце-удальце, молодильных яблоках и живой воде» получилось, что наилучшее качество достигалось при пороге значимости показателей связи $\alpha = 0,2$. В дальнейшем, чтобы улучшить этот параметр, можно модифицировать алгоритм агрегации, дополнительно учитывая информацию об упорядоченности объектов и связей в сюжете.

СПИСОК ИСТОЧНИКОВ

1. Jurafsky D., Martin J.H. Speech and Language Processing // An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models. 3rd ed. Online manuscript released January 12, 2025. URL: <https://web.stanford.edu/~jurafsky/slp3> (дата обращения 28.04.2025).
2. Катермина Т.С., Тагиров К.М., Тагиров Т.М. Элементы искусственного интеллекта в решении задач анализа текстов // Computational Nanotechnology. 2022. Т. 9, № 2. С. 35–44. DOI: 10.33693/2313-223X-2022-9-2-35-44
3. Морозов Д.А., Глазкова А.В., Губарькова Я.Н., Гарипов Т.А., Столяров С.С., Власова Н.А., Ляшевская О.Н., Смаль И.А., Козеренко А.Д. Применение инструментов обработки естественного языка на базе машинного обучения при разработке корпусов: опыт Национального корпуса русского языка // Материалы Международной научной конференции, посвященной 20-летию Национального корпуса русского языка. М., 2024. С. 118–121. DOI: 10.31912/nac_corp_20-2024



4. **Фатхулин Т.Д., Климов Н.Ю., Гежин С.А.** Анализ нейросетевых технологий, позволяющих генерировать текст // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. Ростов-на-Дону, 2023. С. 123–127.
5. **Жирмунский В.М.** К вопросу о международных сказочных сюжетах // Человек: Образ и сущность. Гуманитарные аспекты. 2004. № 1 (15). С. 227–236.
6. **Пропп В.** Морфология сказки. Л.: Academia, 1928. 152 с.
7. **Лотман Ю.М.** Структура художественного текста. М.: Искусство, 1970. 384 с.
8. **Москин Н.Д., Лебедев А.А.** Теоретико-графовый подход к анализу вариативности русских сказок // Интеллектуальные системы в производстве. 2024. Т. 22, № 4. С. 90–97. DOI: 10.22213/2410-9304-2024-4-90-97
9. **Гаазе-Рапопорт М.Г.** Поиск вариантов в сочинении сказок // Зарипов Р.Х. Машинный поиск вариантов при моделировании творческого процесса. М.: Наука, 1983. С. 213–223.
10. **Афанасьев А.М.** Народные русские сказки А.Н. Афанасьева: в 3 т. Т. 1. / подготовка текста и примеч. В.Я. Пропп. М.: Гослитиздат, 1957. 514 с.
11. **Сюй Ц., Агеева Ю.В.** Дискурсивная типология сказки (на примере сборника русских сказок А. Н. Афанасьева) // Филологические науки. Вопросы теории и практики. 2020. Т. 13, № 2. С. 134–139. DOI: 10.30853/filnauki.2020.2.26

REFERENCES

- [1] **Jurafsky D., Martin J.H.**, Speech and Language Processing, An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd ed., Online manuscript released January 12, 2025. Available at: <https://web.stanford.edu/~jurafsky/slp3> (accessed 28.04.2025).
- [2] **Katermina T.S., Tagirov K.M., Tagirov T.M.**, Elements of artificial intelligence in solving problems of text analysis, Computational Nanotechnology, 9 (2) (2022) 35–44. DOI: 10.33693/2313-223X-2022-9-2-35-44
- [3] **Morozov D.A., Glazkova A.V., Gubarkova Ya.N., Garipov T.A., Stolyarov S.S., Vlasova N.A., Lya-shevskaya O.N., Smal I.A., Kozerenko A.D.**, Primeneniye instrumentov obrabotki yestestvennogo yazyka na baze mashinnogo obucheniya pri razrabotke korpusov: opyt Natsionalnogo korpusa russkogo yazyka [Application of natural language processing tools based on machine learning in corpus development: the experience of the National Corpus of the Russian Language], Materialy Mezhdunarodnoy nauchnoy konferentsii, posvyashchennoy 20-letiyu Natsionalnogo korpusa russkogo yazyka [Proceedings of the International Scientific Conference dedicated to the 20th anniversary of the National Corpus of the Russian Language], Moscow, 2024, 118–121. DOI: 10.31912/nac_corp_20-2024
- [4] **Fatkhulin T.D., Klimov N.Yu., Gezhin S.A.**, Analysis of Neural Network Techniques for Text Generation, Trudy Severo-Kavkazskogo filiala Moskovskogo tekhnicheskogo universiteta svyazi i informatiki [Proceedings of the North Caucasian branch of the Moscow Technical University of Communications and Informatics], Rostov-on-Don, 2023, pp. 123–127.
- [5] **Zhirmunskiy V.M.**, On the Theme of International Fairy-Tale Plots, Human Being: Image and Essence. Humanitarian Aspects, 1 (15) (2004) 227–236.
- [6] **Propp V.**, Morfologiya skazki [Morphology of the fairy tale], Academia, Leningrad, 1928.
- [7] **Lotman Yu. M.**, Struktura khudozhestvennogo teksta [The structure of the artistic text], Art, Moscow, 1970.
- [8] **Moskin N.D., Lebedev A.A.**, Graph-theoretical Approach to the Analysis of the Variability of Russian Fairy Tales, Intellektual'nye sistemy v proizvodstve, 22 (4) (2024) 90–97. DOI: 10.22213/2410-9304-2024-4-90-97
- [9] **Gaaze-Rapoport M.G.**, Poisk variantov v sochinenii skazok [Search for variants in the composition of fairy tales], Zari-pov R.Kh., Mashinnyy poisk variantov pri modelirovanii tvorcheskogo protsessa [Machine search for variants in modeling the creative process], Nauka, Moscow, 1983, pp. 213–223.
- [10] **Afanasyev A.M.**, Narodnyye russkiye skazki A.N. Afanasyeva [Russian folk tales by A.N. Afanasyev], Vol. 1, Goslitizdat, Moscow, 1957.
- [11] **Xu J., Ageyeva Yu.V.**, Discursive Typology of Fairy Tale (by the Example of A. N. Afanasyev's Collection "Russian Fairy Tales"), Philology. Theory & Practice, 13 (2) (2020) 134–139. DOI: 10.30853/filnauki.2020.2.26



СВЕДЕНИЯ ОБ АВТОРАХ / INFORMATION ABOUT AUTHORS

Москин Николай Дмитриевич

Nikolai D. Moskin

E-mail: moskin@petrsu.ru

ORCID: <https://orcid.org/0000-0001-5556-5349>

Лебедев Александр Александрович

Alexander A. Lebedev

E-mail: perevodchik88@yandex.ru

ORCID: <https://orcid.org/0000-0001-9939-9389>

Поступила: 05.05.2025; Одобрена: 10.06.2025; Принята: 30.06.2025.

Submitted: 05.05.2025; Approved: 10.06.2025; Accepted: 30.06.2025.