

УДК 519.179.2

## ГРАФ СОЦИАЛЬНЫХ СВЯЗЕЙ РУССКИХ ПИСАТЕЛЕЙ: СЕТЕВОЙ АНАЛИЗ ЛИТЕРАТУРНОГО СООБЩЕСТВА НА ОСНОВЕ КОДИФИКАТОРА ЕГЭ

А. В. Бородина

*Институт прикладных математических исследований КарНЦ РАН,  
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,  
Республика Карелия, Россия, 185910)*

В статье предлагается подход к количественному анализу социальной структуры русских литературных сообществ на основе построения графа личных взаимодействий писателей. Актуальность исследования обусловлена дефицитом структурированных датасетов в данной предметной области и возрастающей ролью методов цифровых гуманитарных наук (Digital Humanities) в изучении социодинамики литературы. Эмпирической базой для построения математической модели сообщества послужил перечень авторов из официального «Кодификатора ЕГЭ по литературе» (2026 г.), что исключает преднамеренную связность авторов. С помощью нейросетевых моделей (DeepSeek, YandexGPT) и последующей верификации экспертом были собраны данные о социальных связях (личное общение, переписка, членство в союзах) для 102 русских писателей 1757–2003 гг. Проведенный анализ сетевой структуры социальных взаимосвязей между писателями выявил типичные для социальных сетей свойства графа. Для ранжирования писателей предложена комбинированная метрика, сочетающая локальную и глобальные центральности, показана ее корреляция с алгоритмом PageRank. Предложенный подход к построению графа социальных взаимодействий дает результаты, согласованные с экспертными источниками, и может быть использован для социодинамического анализа литературных сообществ. Методы анализа социальных сетей хорошо применимы для изучения литературных сообществ и позволяют предложить альтернативный подход к оценке статуса писателя через структуру личных связей.

**Ключевые слова:** социальные сети; случайные графы; нейросеть; сообщество писателей; коэффициент кластеризации; модулярность; ранжирование

Для цитирования: Бородина А. В. Граф социальных связей русских писателей: сетевой анализ литературного сообщества на основе кодификатора ЕГЭ // Труды Карельского научного центра РАН. 2026. № 6. С. 5–19. doi: 10.17076/mat2368

**Финансирование.** Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

# A. V. Borodina. A GRAPH OF SOCIAL CONNECTIONS OF RUSSIAN WRITERS: A NETWORK ANALYSIS OF THE LITERARY COMMUNITY BASED ON THE UNIFIED STATE EXAM CODIFIER

*Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)*

This paper proposes an approach to the quantitative analysis of the social structure of Russian literary communities by constructing a graph of personal connections among writers. The relevance of the study arises from the lack of structured datasets in this subject domain and the growing significance of digital humanities methodologies for investigating the sociodynamics of literature. The empirical basis for constructing a mathematical community model was the list of authors from the official “Unified State Exam Codifier in Literature” (2026). This source ensures representativeness while avoiding artificial connectedness among authors. Social connections data, including personal communication, correspondence, and membership in professional unions, were collected for 102 Russian writers active between 1757 and 2003. The data collection process employed neural network models (DeepSeek, YandexGPT) followed by expert verification. The results showed that the proposed method for constructing the social interactions graph yields results consistent with expert data and can be used for sociodynamic analysis of literary communities. Social network analysis methods are well suited for studying writer communities and offer an alternative approach to assessing a writer’s status through the structure of personal connections.

**Key words:** social networks; random graph; neural network; writers community; clustering coefficient; modularity; ranking

**For citation:** Borodina A. V. A graph of social connections of Russian writers: a network analysis of the literary community based on the Unified State Exam (USE) codifier. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 5–19. doi: 10.17076/mat2368

**Funding.** The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

---

## ВВЕДЕНИЕ

Литературоведы и историки при анализе литературных сообществ оперируют скорее качественными, а не количественными оценками. Исследование свойств сообщества, выделение в нем значимых групп (кластеров) требует выявления социальных связей между каждой парой авторов. Подобные задачи решаются достаточно успешно, если речь идет, например, о структуре социальной сети либо о научном сообществе, где связи – это ссылки цитирования или соавторство в статьях.

Как показали результаты поиска, для сообщества писателей достаточно сложно найти готовые датасеты. Задача сбора и описания социальных связей и динамики в писательских сообществах является актуальной на сегодняшний день. В качестве эмпирической базы для изучения русских писателей специалисты-социологи в основном используют фундаментальные издания: «Лексикон русской литера-

туры XX века» В. Казака [1], «Русские писатели. 1800–1917» [6], «Русские писатели XIX века» [7], «Русские писатели XX века» [8]. Основная группа русских авторов советского периода может быть определена по энциклопедическому справочнику «Мир русской культуры» (1997) [4], а профессиональное сообщество последних десятилетий советского периода – по «Справочнику Союза писателей СССР» [11].

На актуальность проблематики количественного описания русских литературных сообществ указывает возросший интерес к исследованиям в сфере цифровых гуманитарных наук (Digital Humanities) и поддержка проектов в этой области ведущими российскими научными фондами.

Большой социодинамический анализ сообщества русских писателей дореволюционного периода и литературной среды в советское время на основе авторской методики С. Я. Суцкого можно найти в работах [12, 13].

Схожие задачи ставят специалисты Института русской литературы РАН (Пушкинский дом) г. Санкт-Петербурга и работают с тем же массивом данных (словарь «Русские писатели. 1800–1917» [6]) в рамках текущего научного проекта «Социальная история русской литературы XIX века. Разработка методологии описания и создание электронного ресурса» (период выполнения 2024–2026 гг.), финансируемого Российским научным фондом [2]. Основная задача проекта – создание электронной базы данных на основе биографического словаря с целью обобщить социальные характеристики авторов (сословие, образование, жанровый выбор, места публикаций и т. д.) и выявить значимые корреляции.

Кроме того, важной открытой задачей является разработка методики оценивания писательского труда и рейтинга писателя в литературном поле. Подробнее вопросы анализа статуса писателей и их борьбы за позиции в современном российском литературном поле можно найти в работах [9, 10]. Критерии оценки, по которым оценивается положение и успех писателей, выявляются эмпирически, на основе тематического анализа интервью с ведущими российскими литературными критиками и публицистами.

Интерес к социологическому портрету писателя и анализу литературного поля активно развивается и требует внедрения новых методов для исследования. Построение структуры социального взаимодействия между писателями в виде графа позволит исследовать свойства данного сообщества и применить известные алгоритмы ранжирования для выявления статуса или веса писателя, основываясь на его личных связях в профессиональном сообществе.

Поскольку задачу описания социально-исторических связей между писателями и построение математической модели всего сообщества за весь временной период развития русской литературы решить не представляется возможным, то далее речь пойдет о выделении некоторого подмножества писателей, для которых было бы возможно установить социальные связи и построить модель такого фрагмента общей (неизвестной) сети в виде неориентированного графа, в котором узлами будут персоналии (авторы, писатели), а ребро между узлами проводится в случае, когда два автора взаимодействовали лично: связаны персональным общением, перепиской, входили в один литературный союз и т. д.

В работе в качестве объекта для исследования и построения математической модели

предлагается выбрать множество авторов, регламентированное в официальном документе «Кодификатор Единого государственного экзамена по литературе» за 2026 г. [3]. Кодификатор – это официальный перечень тем, произведений и навыков, которые выносятся на итоговый экзамен по литературе в старшей школе. Список авторов в кодификаторе ЕГЭ по литературе не является фиксированным и регулярно обновляется, писатели включаются в перечень без учета социальных связей или принадлежности к какому-то одному литературному объединению, таким образом можно исключить связность «по построению». Перечень кодификатора содержит произведения из разных временных периодов. В кодификатор попадают произведения, которые изучаются в средней и старшей школе. Тем не менее следует учитывать, что список составляется из ключевых авторов, чье творчество является традиционной и неотъемлемой основой школьного курса литературы.

В работе проведено исследование структуры и свойств графа социального взаимодействия писательского сообщества на базе списка авторов, представленных в кодификаторе [3]. Предложен подход к проведению социодинамического анализа такого рода сообществ на основе выявления свойств графа. Для двух временных периодов – «XIX – начало XX в.» и «советский период» – проведен сравнительный анализ распределения писателей по жанровой специализации.

## МЕТОДИКА СБОРА ДАННЫХ О СОЦИАЛЬНЫХ СВЯЗЯХ

Специалисты в области социологии и литературы отмечают, что на данный момент литературные персоналии изучены очень выборочно. Кроме того, недостаточно исследованы взаимосвязи между персоналиями и их группами, например между писателями и издателями, писателями и читателями, критикой и читателями, цензурой и издателями [12]. Результаты поиска в сети Интернет также обнаруживают отсутствие в открытом доступе готовых датасетов с такой информацией в русскоязычном сегменте.

Достаточно подробное исследование в области социодинамического анализа русского писательского сообщества можно найти в статьях С. Я. Суцего [12, 13], где отдельно рассмотрены два периода: «XIX – начало XX в.» (1781–1917 гг.) и «советский период» (1886–1985 гг.). В качестве основы для анализа использовались биографические словари русских писателей XIX – начала XX века, кото-

рые содержат сведения примерно о 2,5 тыс. персоналиях, а советский период представлен 760 персоналиями. Эти результаты будем считать эталонными для проведения сравнения с данными, вычисленными на подсети кодификатора, с этой целью из полного списка авторов кодификатора были выбраны только русские писатели (102 автора за 1757–2003 гг.), а границы периодов соответствуют установленным в [12, 13].

С помощью нейросети DeepSeek была построена матрица смежности, представляющая собой таблицу, где 1 указывает на наличие социальной связи: переписка, работа в наставничестве, личное знакомство и т. д.; 0 – если связи не выявлено. Далее для уже имеющихся данных проведен повторный поиск связей с помощью другой нейросети – YandexGPT. На основе поиска был составлен список с подробным описанием каждой связи и проведена верификация полученных данных экспертом в области литературы И. Н. Тихоновой<sup>1</sup>. Благодаря такой верификации найдено еще около 20 связей, которые до этого не были учтены.

Для сравнительного анализа по жанровой принадлежности все писатели были упорядочены по периоду дебюта. Перечень жанров взят в соответствии с классификацией С. Я. Сущего и пронумерован: 1 – проза; 2 – поэзия; 3 – драма; 4 – перевод; 5 – публицистика; 6 – критика; 7 – литературоведение; 8 – сценарии; 9 – журналистика; 10 – мемуары.

*Таблица 1.* Пример расхождения мнений YandexGPT и DeepSeek в жанре проза по писателю Е. А. Евтушенко

*Table 1.* An example of the divergence of opinions between YandexGPT and DeepSeek in the prose for the writer Ye. A. Yevtushenko

DeepSeek	1	«Ягодные места», автобиографическая проза
YandexGPT	0	Поэт и публицист; проза не главная форма творчества

Для поиска жанровой принадлежности снова использовались две нейросетевые языковые модели YandexGPT и DeepSeek. Формат вывода результата включал таблицу, состоящую из трех полей: 1 – ФИО писателя; 2 – ответ о принадлежности писателя к определенному жанру (1 – да, 0 – нет); 3 – обоснование. Данные сохранялись в формате csv, и далее с помощью программы на языке python выявлялись несовпадения по вто-

рому полю. Результаты сравнения показали, что процент несовпадения ответов двух нейросетей в таблицах составляет 1,96–3,92% в зависимости от жанра. В таблице 1 приведен пример несовпадения результатов поиска. В таком случае решение принималось на основе повторного поиска с указанием произведения, а также на основе мнения эксперта.

### СВОЙСТВА ГРАФА СОЦИАЛЬНОГО ВЗАИМОДЕЙСТВИЯ ПИСАТЕЛЕЙ

Обозначим  $G = (V, N)$  неориентированный граф социального взаимодействия писателей русского сегмента кодификатора ЕГЭ по литературе [3]. Вершины графа  $G$  – это писатели, их множество фиксировано  $|V| = n$ . Множество ребер графа  $|E| = m$  отражает взаимодействия в сообществе писателей. Ребро между двумя вершинами проводится, если имеется личная связь между парой авторов, а именно: личное знакомство, переписка, соавторство, наставничество, принадлежность к одному литературному сообществу и т. д.

Многие сложные социальные сети реального мира обладают рядом общих базовых свойств. Проведем анализ типичных свойств для построенной структуры. Вычисления реализованы на языке python с использованием библиотеки NetworkX.

### Разреженная структура графа

Известно, что граф является разреженным [5], если при наличии  $n$  вершин количество ребер  $k \cdot n$  для константы  $k > 1$  значительно меньше максимального числа возможных ребер в полном графе с тем же числом вершин. Данное свойство присуще многим социальным сетям и впервые было описано как ключевая особенность веб-графов в исследованиях Barabasi и Albert [15].

Для графа  $G$  число вершин  $n = 102$ , число ребер  $m = 288$ , что существенно меньше, чем число ребер в полном графе, которое равно  $C_n^2 = n(n-1)/2 = 5151$ . Отношение числа ребер графа  $G$  к максимально возможному числу ребер называют плотностью графа

$$D = \frac{2m}{n(n-1)} = 0,0198,$$

у разреженного графа плотность близка к 0.

Кроме того, средняя степень узлов

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = 5,65,$$

где  $d_i$  – число ребер, инцидентных вершине с номером  $i$ , на порядки меньше числа узлов

<sup>1</sup>И. Н. Тихонова – учитель русского языка и литературы МОУ «Лицей № 40», г. Петрозаводск.

графа  $G$ , что также указывает на его разреженность.

### Свойство связности и закон «шести рукопожатий»

Граф  $G$  состоит из четырех компонент связности, наибольшая по размеру (гигантская) компонента связности  $G_{giant} = (V_{giant}, E_{giant})$  состоит из 97 вершин. Три остальных компонента содержат 1, 1 и 3 узла соответственно.

Закон «шести рукопожатий» («шести кликов», «мир тесен») – это свойство, присущее веб-графам и некоторым социальным графам, состоит в том, что среднее расстояние между любыми двумя узлами сети составляет около 6 звеньев. В классическом исследовании Jeffrey Travers и Stanley Milgram [25] средняя длина цепочки составляет около 5–6. Данные исследования по Facebook [17] дают среднее расстояние 4,57 для всех пользователей и 3,46 для жителей США.

Диаметр в подграфе  $G_{giant}$  равен 9, среднее расстояние 3,4, что говорит о тесноте связей и выполнении принципа «мир тесен».

### Степенной закон распределения степеней вершин

Для структуры социальных сетей характерно, что существует небольшое количество узлов с большими степенями (влиятельные лица в социальной сети), а большинство узлов имеют очень малые степени. Распределение степеней вершин в таких сетях моделируется с помощью степенного закона распределения:

$$P(\text{deg}(v) = d) \sim c \cdot d^{-\alpha},$$

где  $\text{deg}(\cdot)$  обозначает степень вершины  $v \in V$ ,  $d$  – значение степени, а константа  $c$  определяется из условия нормировки (сумма вероятностей равна 1),  $\alpha > 1$  [5, 15].

На рис. 1 показаны результаты вычисления оценки параметров степенного закона распределения степеней вершин графа  $G$  с помощью построения степенной регрессии. Уравнение регрессии статистически значимо (p-value =  $2,68e - 07$  для теста Фишера), коэффициент детерминации  $R^2 = 0,798$ .

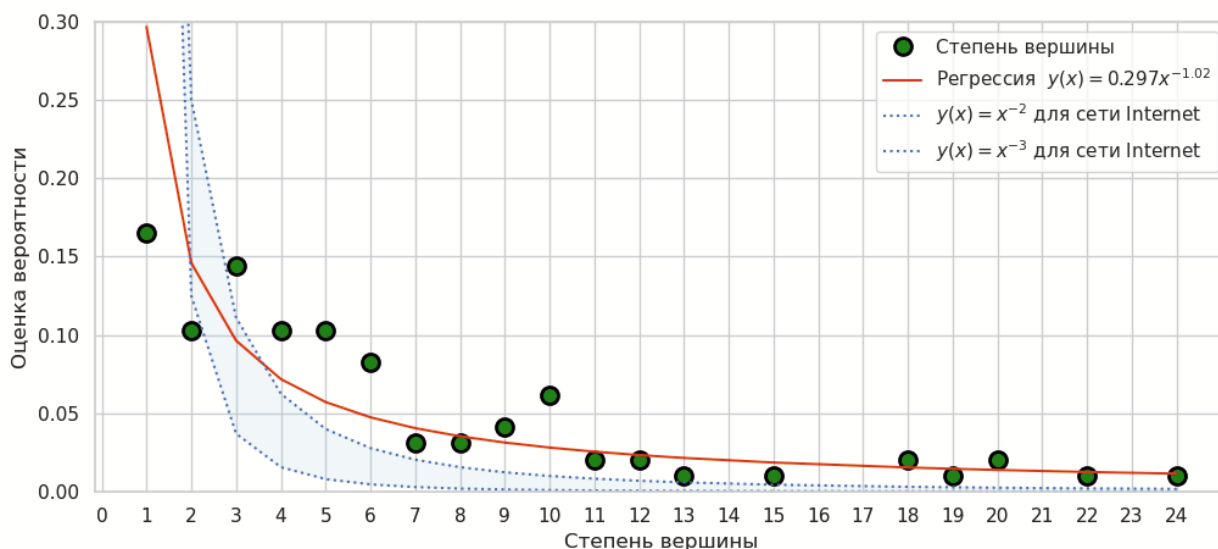


Рис. 1. Параметры степенной регрессионной модели для графа  $G$

Fig. 1. The power-law regression model parameters for the graph  $G$

В общем случае показатель степени  $\alpha$  в распределении степеней вершин для графов социальных сетей чаще всего находится в интервале между 2 и 3. Это наиболее распространенное значение, которое возникает из-за механизма «предпочтительного присоединения» (когда новые участники с большей вероятностью знакомятся с уже популярными людьми). Именно для таких сетей характерно наличие «хабов» – узлов с большим количеством связей. Исследования реальных социальных

сетей показывают, что многие из них могут иметь значительно более низкие показатели масштабирования  $\alpha \sim 1$  (например, [14, 23]).

Для рассматриваемой структуры сообщества писателей параметр  $\alpha = 1,02$  не попал в классический диапазон [2; 3], это может говорить о сильном влиянии принципа взаимности, когда человек общается с теми, кто общается с ним, что характерно для социальных объединений, где связь – это личное общение.

## Коэффициент кластеризации и выделение сообществ

Насколько хорошо граф поддается разбиению на отдельные группы-сообщества, можно оценить на основе модулярности. Для неориентированного графа  $G$  с матрицей смежности  $A_{n \times n} = (A_{ij})_{i,j=1}^n$  и разбиения  $c = \{c_1, \dots, c_k\}$  на  $k$  кластеров модулярность разбиения определяется формулой Ньюмана – Гирвана [21]:

$$Q(c) = \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{d_i d_j}{2m}) I\{c_i = c_j\}, \quad (1)$$

где  $d_i$  – степень узла  $i$ , функция-индикатор  $I\{c_i = c_j\}$  равна 1, если узлы  $i$  и  $j$  принадлежат одному сообществу, и 0, если иначе.

Если разбиение на сообщества известно, то модулярность целесообразно вычислять в виде:

$$Q(c) = \sum_{i=1}^k \left[ \frac{L_i}{m} - \left( \frac{D_i}{2m} \right)^2 \right],$$

где  $k$  – общее количество сообществ,  $L_i$  – количество ребер внутри сообщества  $c_i$ ,  $D_i$  – сумма степеней всех узлов, входящих в сообщество  $c_i$ .

### 1. Кластеризация методом Лувена.

Для разбиения на сообщества будем использовать алгоритм Лувена (Louvain algorithm), который имеет практическую эффективность при работе с разреженными графами и максимизирует модулярность  $Q(c)$  [14, 16].

Алгоритм выявляет группы тесно связанных узлов в сети, где связи внутри группы плотнее, чем связи с остальным «миром». Его часто используют в социальных сетях (для поиска групп друзей), в биоинформатике (для анализа взаимодействия белков) и в финансовой сфере (для обнаружения мошеннических транзакций). Метод Лувена удобен тем, что не требует предварительного знания количества кластеров, но на поздних итерациях могут возникать плохо связанные кластеры. Кроме того, для него характерно повышенное число кластеров. В частности, на модельных данных, где разбиение на сообщества известно, алгоритм Лувена может разбивать истинные сообщества на более мелкие сообщества с большей модулярностью (см., например, [14]).

Стандартный алгоритм Лувена максимизирует модулярность, определенную формулой (1). В этом случае алгоритм может не обнаружить сообщества, размер которых меньше некоторого порогового значения (масштаб). Это обстоятельство может быть критично для больших сетей, где алгоритм будет объединять маленькие, но плотные (с точ-

ки зрения модулярности) сообщества в более крупные кластеры.

В работе для прироста модулярности в алгоритме Лувена вместо (1) использована обобщенная формула Райхардта – Борнхольдта [22] в виде:

$$Q(c) = \frac{1}{2m} \sum_{ij} (A_{ij} - r \frac{d_i d_j}{2m}) I\{c_i = c_j\}, \quad (2)$$

где параметр  $r$  (разрешение) управляет масштабом обнаруживаемых сообществ.

При  $r = 1$  формула (2) является частным случаем классической модулярности. Выбор  $r < 1$  способствует поиску крупных структур, на практике, как правило, выбирают  $r \in [0, 5; 0, 8]$ . Для обнаружения более мелких сообществ нужно выбирать  $r > 1$ , чем выше число, тем больше будет итоговых сообществ. Как правило, достаточно протестировать алгоритм для значений из диапазона  $r \in [1, 2; 2, 0]$ .

Для выбора оптимального значения параметра разрешения  $r$  для графа  $G$  используется эмпирический подход по методу «локтя» (Elbow method). На рис. 2 приведены результаты вычисления для  $r \in [0, 5; 2, 1]$  с шагом 0,1. Оптимальное значение модулярности  $Q = 0,6055$  достигается при  $r \in [0, 9; 1, 2]$ , число кластеров  $k = 8$ , что говорит о высокой модулярности и четко выраженной структуре сообществ.

Для верификации полученных результатов проведена кластеризация методом Лейдена.

### 2. Кластеризация методом Лейдена.

Поскольку алгоритм Лувена имеет недостатки и иногда может создавать несвязные сообщества, проведем кластеризацию графа  $G$  с использованием алгоритма Лейдена (Leiden algorithm), который как раз гарантирует, что все обнаруженные сообщества будут хорошо связаны внутри.

Метод был предложен в работе [24], он считается улучшенной версией метода Лувена и использует для определения модулярности обобщенную формулу (2). На рис. 3 приведены результаты вычислений для графа  $G$ . Оптимальное значение модулярности  $Q = 0,6256$  достигается при  $r \in [0, 8; 1, 6]$ , число кластеров  $k = 8$ .

Результаты экспериментов показывают, что для исследуемого графа сообщества писателей можно использовать для разбиения на кластеры разрешение  $r = 1$  с максимальным значением модулярности. При этом число кластеров  $k = 8$  и состав самих кластеров совпадает для обоих методов. На рис. 4 показана визуализация разбиения графа  $G$  на сообщества  $c_i, i \in [1; 8]$ .

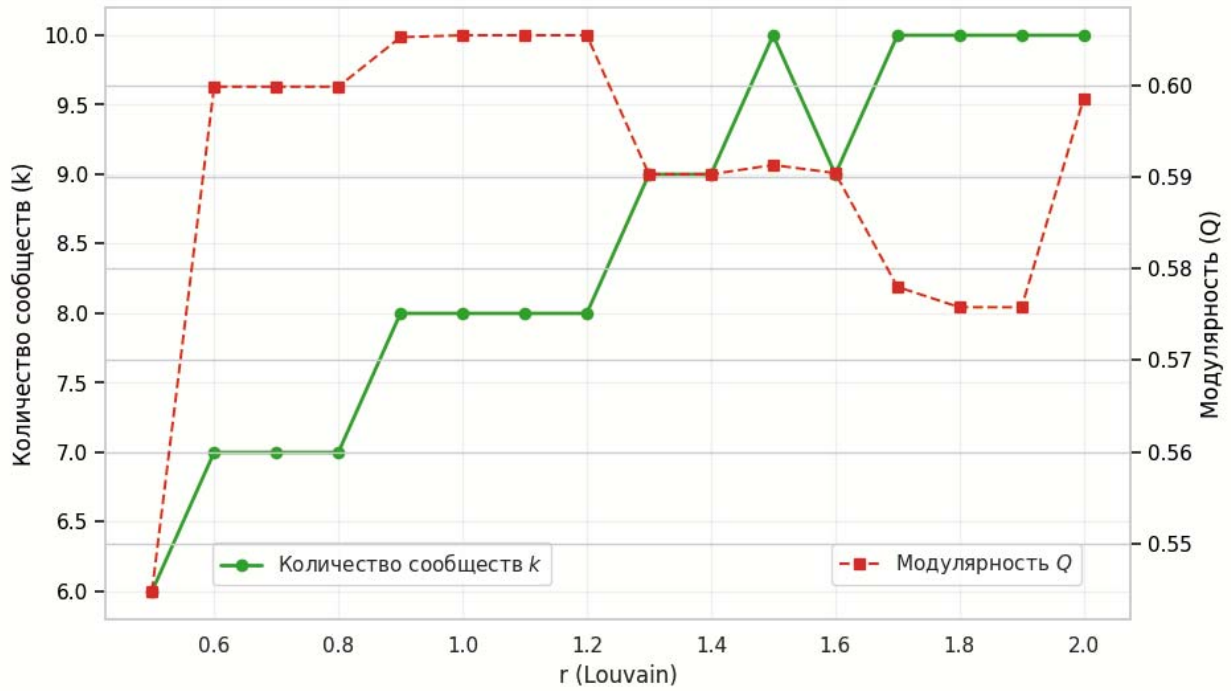


Рис. 2. Параметры разбиения графа  $G$  на кластеры методом Лувена  
 Fig. 2. Parameters for clustering the graph  $G$  with the use of Louvain's method

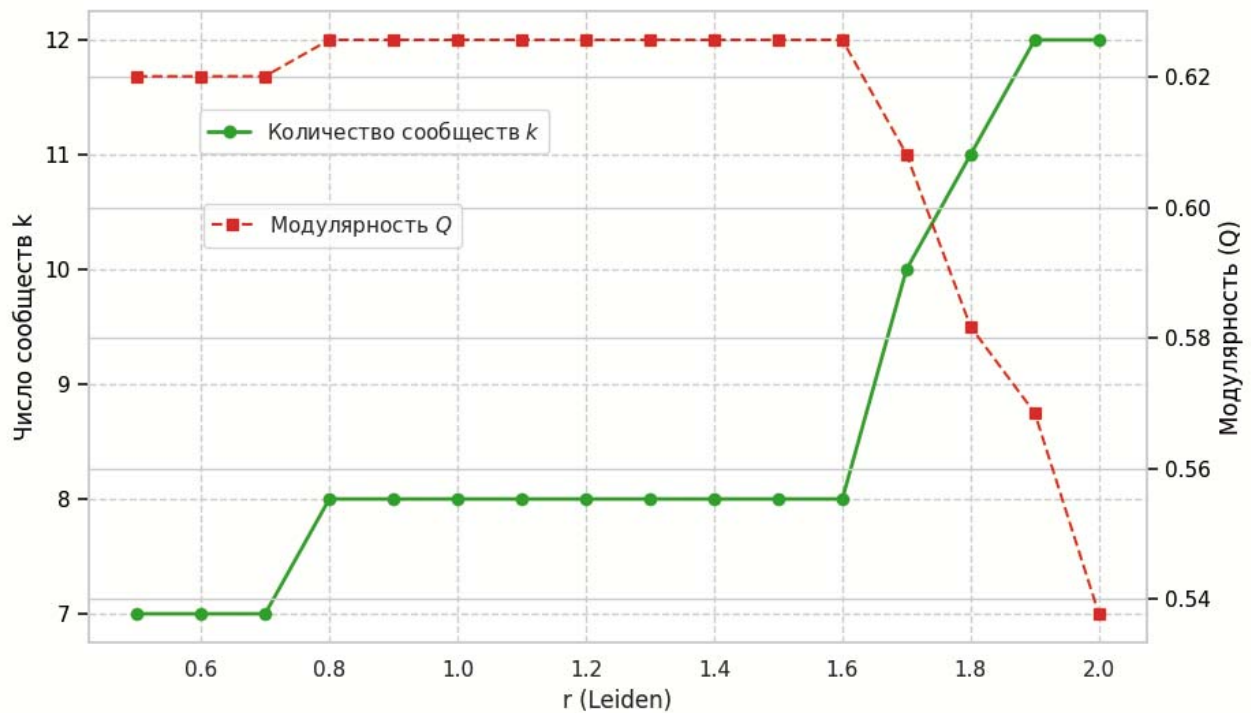


Рис. 3. Параметры разбиения графа  $G$  на кластеры методом Лейдена  
 Fig. 3. Parameters for clustering the graph  $G$  with the use of Leiden's method

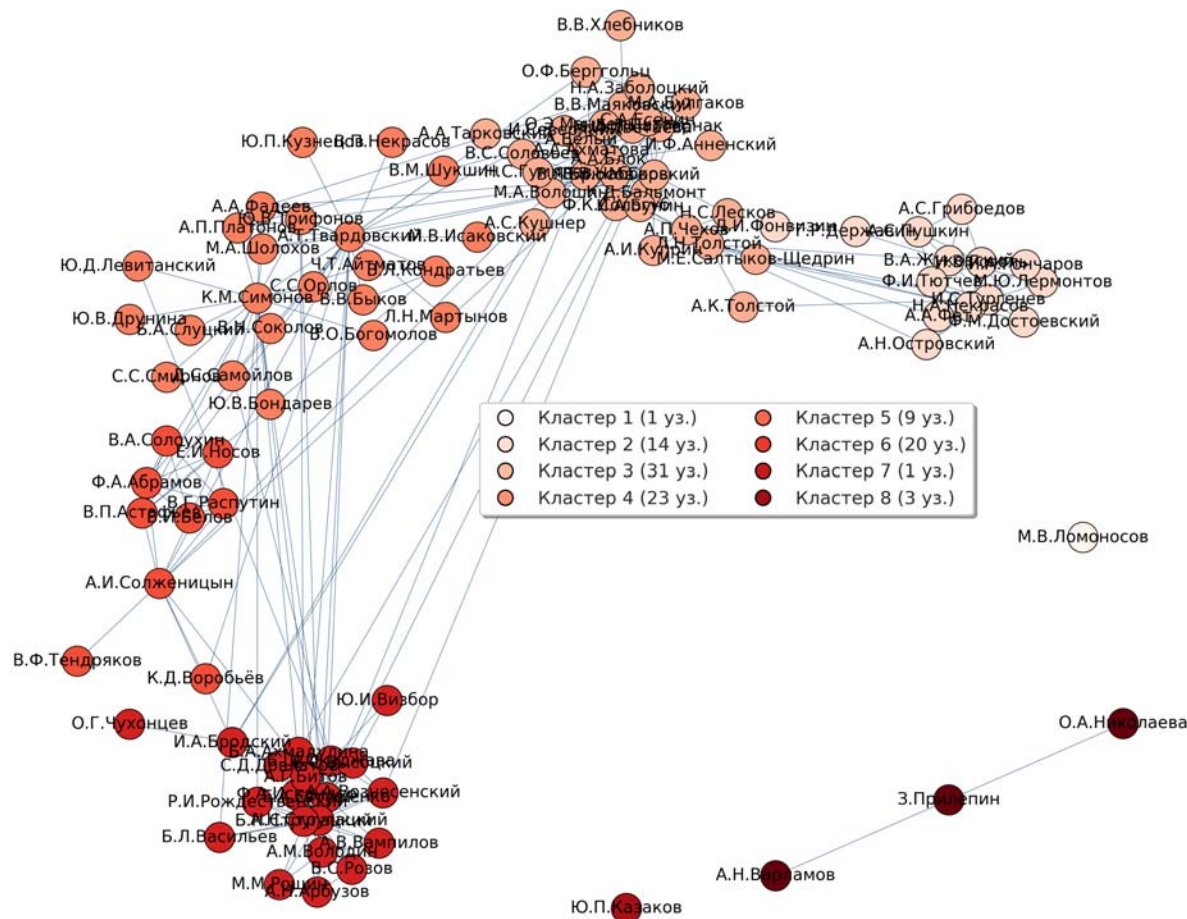


Рис. 4. Разбиение на кластеры  $c_i, i \in [1; 8]$  графа  $G$  по методу Лувена  
 Fig. 4. Clustering the graph  $G$  into  $c_i, i \in [1; 8]$  with the use of Louvain's method

### Ранжирование и выделение центральных узлов

Вопрос определения социального статуса писателей и репутации писательского труда является дискуссионным. В основном для определения рейтинга писателя учитывают мнение эксперта – литературного критика (см., например, в [10]). Анализ тесноты социального взаимодействия между писателями на основе личного общения может быть дополнительной альтернативной методикой для определения веса писателя, поскольку на профессиональную значимость личные связи оказывают значительное влияние. Ниже предлагается сравнить способы определения значимости узла на основе двух метрик: PageRank и комбинированной центральности по степеням.

**1. Метрика PageRank.** Алгоритм PageRank был разработан и успешно применяется с целью ранжирования веб-страниц в сетях большого объема с учетом важности (значимости), где важность страницы определяется важностью страниц, которые на нее ссылаются [14].

На текущий момент идея, лежащая в основе PageRank, применима во многих типах сообществ. Например, в академической среде импакт-фактор журнала зависит от того, сколько важных статей на него ссылаются. В социальных сетях авторитетный человек – это тот, на кого ссылаются (упоминают) другие авторитетные люди.

Данный подход может оказаться также полезным для определения социального статуса и репутации писателей, когда для рейтинга писателя анализируется не только содержание его произведений, а еще и структура социальных связей всей сети сообщества, куда входит данный писатель.

В работе был реализован метод вычисления метрики PageRank для построенного графа сообщества писателей  $G$  с учетом разбиения на 8 кластеров, в каждом кластере определены центральные узлы.

В табл. 2 приведены результаты вычисления метрики  $PR$  для центральных узлов кластеров. На рис. 5 показана структура сети с выделением узлов по размеру на основе метрики PageRank.



смешанных мер, вычисленных на локальной и глобальной структуре сети, описан и в других работах [18, 19]. Обобщенно метрику СОМС для узла  $i$  в неориентированном графе  $G$  можно представить в виде:

$$СОМС(i) = \alpha \cdot C_{local}(i) + (1 - \alpha) \cdot C_{global}(i),$$

где  $C_{local}(i)$  – коэффициент, отражающий локальную роль узла (например, его степень или локальную кластеризацию);  $C_{global}$  – коэффициент, отражающий глобальную роль узла (например, близость к другим узлам или посредничество);  $\alpha \in [0; 1]$ .

Метрика СОМС разработана для решения фундаментальной проблемы: традиционные меры центральности учитывают либо только локальные, либо только глобальные характеристики узлов, что приводит к снижению эффективности при работе с крупномасштабными сетями. Подход, основанный на комбинации СОМС, позволяет учесть узлы, находящиеся в промежуточных зонах сети, которые не являются ни явными центрами, ни изолированной периферией. Далее под термином СОМС будем понимать не одну фиксированную формулу, а семейство метрик, построенных по принципу выпуклой комбинации «локальности» и «глобальности». Такой подход к вычислению центральности узлов обеспечивает простой и мощный инструмент балансировки между микро- и макросвойствами сети, что особенно ценно при поиске структурно значимых объектов в социологических исследованиях, когда узел сети – это персоналия.

Учитывая специфику рассматриваемой сети сообщества писателей, комбинированный подход при выборе метрики, определяющей важность узла, достаточно логичен, поскольку для определения социального статуса писателя необходимо учитывать и его связи внутри социальной группы (литературные объединения, кружки, союзы писателей), и способность поддерживать социальные связи с ключевыми значимыми фигурами (редакторы, литературные агенты, издательства), способность влиять на общественное мнение (участие в фестивалях, лекциях, сотрудничество с медиа) и т. д.

В текущей работе для решения задачи ранжирования писателей и идентификации центральных узлов в кластерах предлагается использовать комбинированный показатель, основанный на взвешенной сумме метрик центральности (Degree, Betweenness, Closeness) узла  $i$ :

$$C_{DBC}(i) = \alpha_1 C_D(i) + \alpha_2 C_B(i) + \alpha_3 C_C(i), \quad (3)$$

где  $\sum_{j=1}^3 \alpha_j = 1$ . В локальную компоненту формулы (3) входит коэффициент  $C_D(i)$ , равный числу прямых связей (ребер) узла  $i$  с другими узлами сети (метрика Degree Centrality), который показывает центральность узла по степени и отражает его «локальную известность». Локальная компонента показывает, насколько вершина активна в своем ближайшем окружении. Глобальную компоненту формулы (3) формируют два коэффициента:  $C_B(i)$  вычисляется как число кратчайших путей через узел  $i$  (метрика Betweenness Centrality);  $C_C(i)$  вычисляется на основе среднего расстояния до всех узлов сети (метрика Closeness Centrality), которые в комбинации учитывают положение узла в общей топологии сети, его роль «моста» и скорость доступа к другим узлам. Глобальная компонента показывает, насколько вершина важна для всего графа в целом.

Выбор весов  $\alpha_1 = 0,5$ ,  $\alpha_2 = 0,3$ ,  $\alpha_3 = 0,2$  обусловлен «природой» метрик, а именно: локальная компонента (Degree) играет доминирующую роль при анализе внутрикластерной структуры, тогда как глобальные компоненты (Betweenness и Closeness) дополняют оценку, учитывая межузловые взаимодействия за пределами непосредственного окружения узла, при этом обе компоненты вносят равный вклад в общую метрику  $C_{DBC}$ . В зависимости от целей ранжирования коэффициенты можно варьировать. В табл. 3 показаны центральные узлы по комбинированной метрике.

Таблица 3. Центральные узлы кластеров  $c_i, i \in [1; 8]$  по комбинированной метрике

Table 3. Central nodes of clusters  $c_i, i \in [1; 8]$  according to the convex metric

№	Центральные узлы Central nodes	$C_{DBC}$
1	[‘М.В.Ломоносов’]	0,500
2	[‘В.А.Жуковский’]	0,593
3	[‘М.Горький’]	0,511
3	[‘А.А.Ахматова’]	0,484
3	[‘А.А.Блок’]	0,443
4	[‘К.М.Симонов’]	0,671
4	[‘А.Т.Твардовский’]	0,621
5	[‘А.И.Солженицын’]	0,674
6	[‘А.Н.Стругацкий’]	0,569
6	[‘Б.Н.Стругацкий’]	0,569
7	[‘Ю.П.Казаков’]	0,000
8	[‘З.Прилепин’]	1,000

Результаты вычислений показали, что комбинированная метрика  $C_{DBC}$  для графа  $G$  дает сходный с PageRank рейтинг и может так-

же успешно быть использована для определения рейтинга писателей, тогда как распределение значимости узла по отдельным метрикам  $C_D, C_B, C_C$  может значительно отличаться от PageRank. В качестве примера такой ситуации можно рассмотреть результаты ранжирования в кластере  $c_6$  (21 узел) из таблицы 4, где метрики  $PR$  и  $C_{COMC}$  дают схожий порядок между собой и отличный от метрики  $C_C$ .

Таблица 4. Ранжирование узлов кластера  $c_6$ , сравнение метрик  
Table 4. Ranking of cluster nodes  $c_6$ , comparison of metrics

№ ( $PR$ )	№ ( $C_{DBC}$ )	№ ( $C_C$ )
1 (0,093)	1 (0,569)	2 (0,741)
2 (0,093)	2 (0,569)	3 (0,741)
3 (0,087)	3 (0,556)	1 (0,769)
4 (0,069)	4 (0,423)	5 (0,645)
5 (0,063)	5 (0,400)	4 (0,667)
6 (0,063)	6 (0,389)	8 (0,625)
7 (0,057)	7 (0,364)	6 (0,645)
8 (0,057)	8 (0,364)	7 (0,645)
9 (0,055)	9 (0,349)	9 (0,606)
10 (0,053)	10 (0,313)	12 (0,526)
11 (0,047)	11 (0,291)	10 (0,540)
12 (0,035)	14 (0,223)	16 (0,487)
13 (0,034)	13 (0,231)	13 (0,526)
14 (0,033)	12 (0,233)	13 (0,526)
15 (0,029)	17 (0,195)	17 (0,476)
16 (0,028)	16 (0,203)	15 (0,512)
17 (0,028)	15 (0,205)	14 (0,526)
18 (0,018)	19 (0,131)	19 (0,408)
19 (0,017)	18 (0,138)	18 (0,444)
20 (0,014)	20 (0,095)	20 (0,351)
21 (0,014)	21 (0,095)	21 (0,350)

### ТВОРЧЕСКО-ЖАНРОВАЯ СПЕЦИАЛИЗАЦИЯ СООБЩЕСТВА

Теперь сравним соотношение авторов по жанровой специализации в графе  $G$  с количественным анализом более обширной группы писателей, опубликованным в экспертных работах С. Я. Суцего [12, 13], которые считаем эталонными для выполнения сравнения. Для этого в рассматриваемом графе  $G$  выделим временные периоды по году дебюта писателей аналогично тем, которые рассматривает эксперт, а именно: 1 период – 1781–1917 гг., 2 период – 1886–1985 гг. Ряд указанных им закономерностей оказалось возможным обнаружить и на структуре графа  $G$ .

Распределение авторов по специализациям за два периода показано на рис. 6 и 7. Мож-

но отметить, что в программу по литературе в средней и старшей школе включены преимущественно авторы, дебютировавшие в прозе и поэзии, при этом процент дебютантов по кодификатору близок к данным, указанным в статьях С. Я. Суцего, а также сходной является динамика в преобладании жанров.

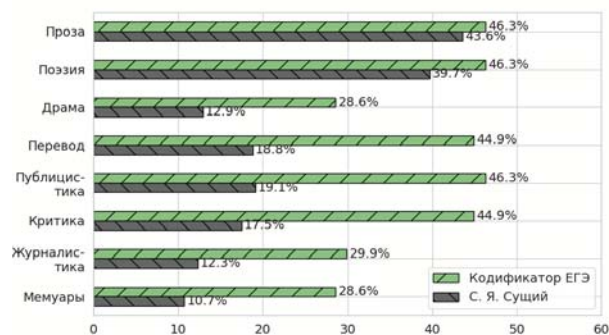


Рис. 6. Жанровая специализация литераторов XIX – начала XX в. (1781–1917 гг.)

Fig. 6. Genre specialization of writers of the 19th and early 20th centuries (1781–1917)

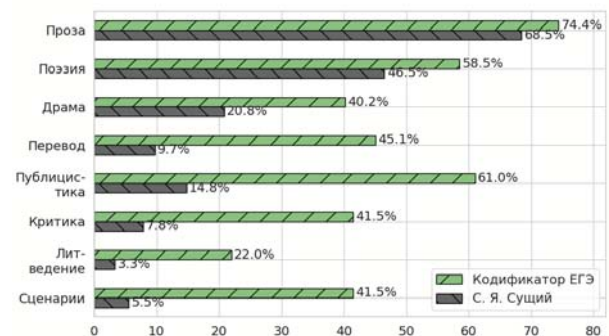


Рис. 7. Жанровая специализация литераторов советского периода (1886–1985 гг.)

Fig. 7. Genre specialization of writers of the Soviet period (1886–1985)

Так, согласно кодификатору, в русском писательском кругу XIX – начала XX в. попеременно доминировали поэзия и проза. В среднем за период и в прозе, и в поэзии дебютировали по 46,3% (43,6 и 39,7% – аналогичные показатели в [12] соответственно).

Остальные жанровые специализации также представлены в кодификаторе, но процент в целом существенно выше, чем описывает С. Я. Суцкий. Это можно объяснить спецификой поиска информации нейросетями DeepSeek, YandexGPT и разным принципом отнесения писателя к тому или иному жанру. В частности, в работах С. Я. Суцего в таблицах подсчитан именно процент дебютантов в том или ином жанре, тогда как нейросеть

относит писателя к определенному жанру тогда, когда у него имеются произведения в этом жанре, несмотря на то, что дебютировать он мог в другом жанре или совмещать специализации.

В частности, для публицистики по кодификатору оценка писателей, работавших в этом жанре, составила 46,3% за 1 период против 19,1% дебютантов в таблицах из работы [12]. Однако в тексте работы отмечается, что «... в 1850–1870-е гг. поэзия среди авторских дебютов оказалась даже не второй, но третьей специализацией, уступив вторую позицию публицистике. Почти каждый пятый литературный дебютант в середине века выбирал для профессиональной самореализации публицистику. Но и представители других творческих специализаций в 1850–1880-х гг. обращаются к ней куда чаще, чем их предшественники. Едва ли можно найти крупного русского писателя этого времени, который не отдал бы ей дань». Таким образом, полученный по кодификатору процент вполне обоснован.

В советский период, согласно кодификатору, в 1920-е гг. проза возвращает себе лидирующие позиции среди дебютантов – 66,2% дебютируют в прозе, при этом 33,7% совмещают прозу и поэзию (2/3 и 1/3 – аналогичные показатели в [13] соответственно). На всем

протяжении советского периода преобладала проза, процентные показатели по кодификатору близки к указанным диапазонам 50–90% в прозе и 35–55% в поэзии. Также данные, полученные по кодификатору, сопоставимы с данными по жанровой специализации всего профессионального сообщества русской литературы, которые представлены в «Справочнике Союза писателей СССР» [11], где основными жанровыми специализациями были проза и поэзия – более 70%.

Отдельный интерес представляет динамика лидерства жанров проза и поэзия для периода XIX – начала XX в., описанная С. Я. Суцким. Вычисления на множестве писателей из кодификатора также показывают перелом в соотношении ведущих жанров в конце 1820-х – 1830-е гг., когда ведущая роль переходит от поэзии к прозе, при этом в жанровом распределении дебютантов в 1830-е гг. они делятся между прозой и поэзией практически в равной степени. Второй переломный момент приходится на Серебряный век русской литературы (1890–1920-е гг.), когда к 1910–1917 гг. доля дебютантов в прозе сократилась. На рис. 8 можно увидеть смену предпочтений между прозой и поэзией в одни и те же периоды по кодификатору и согласно эталонным экспертным данным из статьи [12].

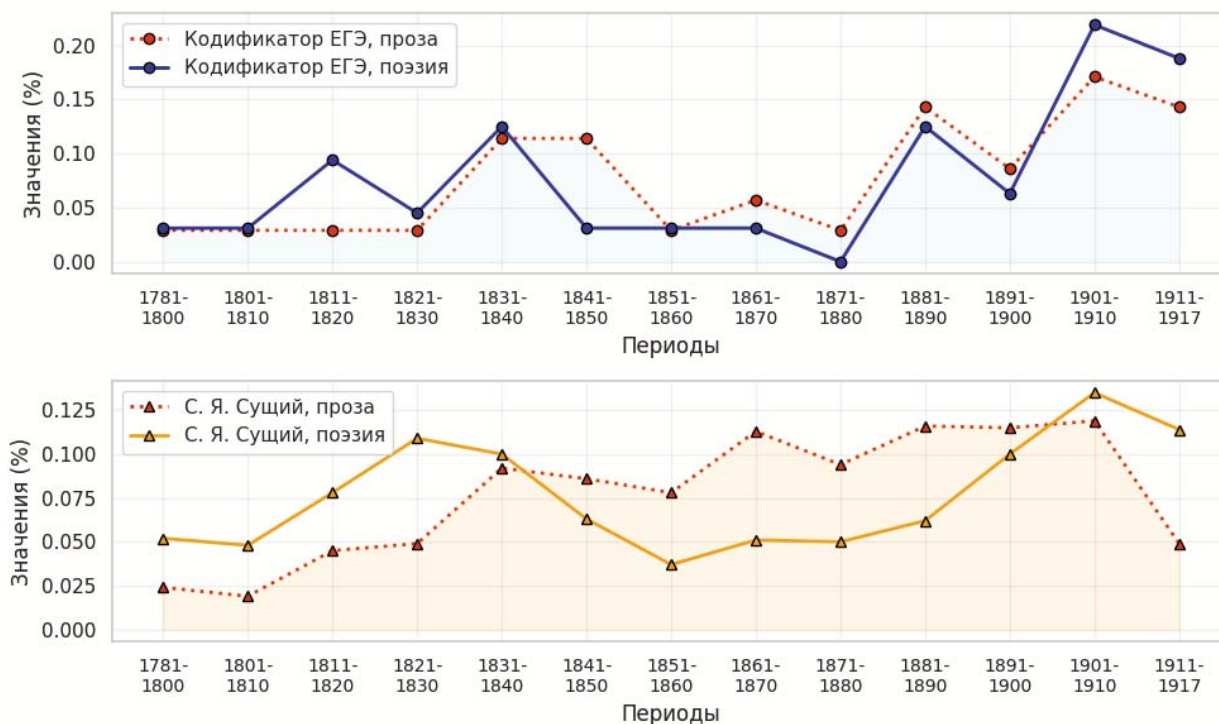


Рис. 8. Динамика доминирования жанров проза и поэзия за период XIX – начала XX в. (1781–1917 гг.)  
Fig. 8. Leadership dynamics between prose and poetry from the 19th to the early 20th centuries (1781–1917)

## ЗАКЛЮЧЕНИЕ

В работе предложен подход к построению модели профессионального сообщества литераторов на основе учета социальных связей между ними и метрики для ранжирования писателей на основе их социального веса.

Для построения математической модели в виде неориентированного графа в качестве множества вершин выбран набор писателей, чьи произведения учтены в официальном кодификаторе для проведения единого государственного экзамена по литературе. Наличие ребра между двумя вершинами графа определялось алгоритмом двухуровневого поиска с помощью нейросетей DeepSeek, YandexGPT с последующей верификацией описания связей экспертом.

Анализ структуры социального графа литераторов показал сходство свойств построенной сети с другими социальными сетями и веб-графами, а именно: разреженность, степенное распределение степеней вершин, свойство «малого мира», высокую кластеризацию и наличие плотных подграфов (сообществ). Граф состоит из четырех компонент связности, наибольшая из которых содержит 97 вершин из 102.

Для выделенного сообщества были исследованы показатели творческой специализации авторов в двух временных отрезках: XIX – начало XX в. и советский период в русской литературе.

Результаты сравнения с данными, приведенными в статьях С. Я. Сущего, показали идентичный характер соотношений между прозой и поэзией для выделенных временных периодов, а также сходную динамику смены лидирующих позиций между двумя стилями.

Таким образом, предложенный в работе способ построения набора данных для описания социальной сети литераторов показал хорошую согласованность с имеющимися в открытом доступе экспертными данными.

Кроме того, обнаруженные свойства позволяют рассматривать гипотезу, что построенный подграф по свойствам схож с графом, который мог быть получен в результате случайного блуждания.

Поскольку исходный граф всей сети литераторов неизвестен и не может быть построен, то достоверно утверждать, что подграф получен случайным блужданием, невозможно, однако можно отметить совместимость с этой гипотезой, т. к. наблюдаемая структура типична для случайного блуждания. Этот вопрос требует дополнительного изучения.

## ЛИТЕРАТУРА

1. *Казак В.* Лексикон русской литературы XX века. М.: Культура, 1996. 512 с.

2. *Карточка* проекта фундаментальных и поисковых научных исследований, поддержанного Российским научным фондом // РНФ [Электронный ресурс]. URL: <https://rscf.ru/project/24-18-00761/> (дата обращения: 10.01.2026).

3. *Кодификатор* проверяемых требований к результатам освоения основной образовательной программы среднего общего образования и элементов содержания для проведения Единого государственного экзамена по литературе [Электронный ресурс]. URL: [https://doc.fipi.ru/ege/demoversii-specifikacii-kodifikatory/2026/li\\_11\\_2026.zip](https://doc.fipi.ru/ege/demoversii-specifikacii-kodifikatory/2026/li_11_2026.zip) (дата обращения: 15.02.2026).

4. *Мир* русской культуры. Энциклопедический справочник / Науч. ред. А. Н. Мячин. М.: Вече, 1997. 624 с.

5. *Райгородский А. М.* Математические модели интернета // Квант. 2012. №. 4. С. 12–16.

6. *Русские писатели. 1800–1917: Биографический словарь: В 5 т./ Гл. ред. П. А. Николаев.* М.: Сов. энцикл.; Большая рос. энцикл., 1989–2007.

7. *Русские писатели, XIX век: Биобиблиографический словарь: в двух частях / Под ред. П. А. Николаева.* М.: Просвещение, 1996.

8. *Русские писатели XX века: Биографический словарь / Гл. ред. и сост. П. А. Николаев.* М.: Большая рос. энцикл.; Рандеву-АМ, 2000. 806 с.

9. *Рязанцев А. П.* Российское литературное поле в XXI веке: оценка состояния в интервью с критиками // Социологические исследования. 2025. Т. 51, № 8. С. 132–141. doi: 10.7868/S3034601025080112

10. *Рязанцев А. П.* Социальный статус и репутация писателей в современной России: тематический анализ // Интеракция. Интервью. Интерпретация. 2023. Т. 15, № 1. С. 27–44. doi: 10.19181/inter.2023.15.1.2

11. *Справочник* Союза писателей СССР. М.: Советский писатель, 1981. 821 с.

12. *Сущий С. Я.* Русская литература XIX – начала XX в.: социодинамический анализ писательского сообщества // Социологические исследования. 2020. № 2. С. 128–143. doi: 10.31857/S013216250008500-3

13. *Сущий С. Я.* Русская литература советского периода – социодинамика писательского сообщества // Социологические исследования. 2021. № 2. С. 112–129. doi: 10.31857/S013216250009902-5

14. *Avrachenkov K., Dreveton M.* Statistical analysis of networks. Boston-Delft: Now Publ., 2020. 248 p. doi: 10.1561/9781638280514

15. *Barabási A. L., Albert R.* Emergence of scaling in random networks // *Science*. 1999. Vol. 286, no. 5439. P. 509–512. doi: 10.1126/science.286.5439.509

16. *Blondel V. D., Guillaume J.-L., Lambiotte R., Lefebvre E.* Fast unfolding of communities in large networks // *J. Stat. Mech.: Theory Exp.* 2008. October. P10008. doi: 10.1088/1742-5468/2008/10/P10008

17. *Edunov S., Diuk C., Filiz I. O., Bhagat S., Burke M.* Three and a half degrees of separation // *Research at Facebook blog*. 4 February 2016. URL: <https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/> (дата обращения: 10.01.2026).

18. *Hajarathaiyah K., Enduri M. K., Anamalamudi S., Sangi A. R.* Algorithms for finding influential people with mixed centrality in social networks // *Arab. J. Sci. Eng.* 2023. Vol. 48, no. 8. P. 10417–10428. doi: 10.1007/s13369-023-07619-w

19. *Mohammad B. B., Dhuli S., Enduri M. K.* Isolating centrality-based generalization of traditional centralities to discover vital nodes in complex networks // *Arab. J. Sci. Eng.* 2025. Vol. 50, no. 15. P. 12003–12025. doi: 10.1007/s13369-024-09628-9

20. *Mohammad B. B., Dhuli V. S., Enduri M. K., Cenkeramaddi L. R.* A novel convex combination-based mixed centrality measure for identification of influential nodes in complex networks // *IEEE Access*. 2024. Vol. 12. P. 123897–123920. doi: 10.1109/ACCESS.2024.3450296

21. *Newman M. E. J., Girvan M.* Finding and evaluating community structure in networks // *Phys. Rev. E*. 2004. Vol. 69, no. 2. Art. 026113. doi: 10.1103/PhysRevE.69.026113

22. *Reichardt J., Bornholdt S.* When are networks truly modular? // *Phys. D: Nonlinear Phenom.* 2006. Vol. 224, no. 1-2. P. 20–26. doi: 10.1016/j.physd.2006.09.009

23. *Schnegg M.* Reciprocity and the emergence of power laws in social networks // *Internat. J. Modern Phys. C*. 2006. Vol. 17, no. 07. P. 1067–1076. doi: 10.1142/S0129183106009473

24. *Traag V. A., Waltman L., Van Eck N. J.* From Louvain to Leiden: guaranteeing well-connected communities // *Sci. Rep.* 2019. Vol. 9, no. 1. Art. 5233. doi: 10.1038/s41598-019-41695-z

25. *Travers J., Milgram S.* An experimental study of the small world problem // *Sociometry*. 1969. Vol. 32, no. 4. P. 425–443.

## REFERENCES

1. *Kazak V.* Lexicon of Russian literature of the 20th century. Moscow: Kul'tura; 1996. 512 p. (In Russ.)

2. Project card for fundamental and exploratory scientific research supported by the Russian Science Foundation. (In Russ.). *RNF = Russian Science*

*Foundation*. URL: <https://rscf.ru/project/24-18-00761/> (accessed: 10.01.2026).

3. Codifier of verifiable requirements for the results of mastering the basic educational program of secondary general education and content elements for the Unified State Examination in literature. (In Russ.). URL: [https://doc.fipi.ru/ege/demoversii-specifikacii-kodifikatory/2026/li\\_11\\_2026.zip](https://doc.fipi.ru/ege/demoversii-specifikacii-kodifikatory/2026/li_11_2026.zip) (accessed: 15.02.2026).

4. *Myachin A. N. (ed.)*. The world of Russian culture. Encyclopedic reference book. Moscow: Veche; 1997. 624 p. (In Russ.)

5. *Rajgorodskij A. M.* Mathematical models of the Internet. *Kvant = Quantum*. 2012;4:12–16. (In Russ.)

6. *Nikolaev P. A. (ed.)*. Russian writers. 1800–1917. Biographical dictionary: In 5 vol. Moscow: Sov. entsikl.; Bol'shaya ros. entsikl.; 1989–2007. (In Russ.)

7. *Nikolaev P. A. (ed.)*. Russian writers of the 19th century: Biographical dictionary. Vol. 1–2. Moscow: Prosveshchenie; 1996. (In Russ.)

8. *Nikolaev P. A. (ed.)*. Russian writers of the 20th century: Biographical dictionary. Moscow: Bol'shaya ros. entsikl.; Rendezvous-AM; 2000. 806 p. (In Russ.)

9. *Ryazantsev A. P.* The Russian literary field in the 19th century: an assessment of the state in interviews with critics. *Sotsiologicheskie issledovaniya = Sociological Research*. 2025;51(8):132–141. (In Russ.). doi: 10.7868/S3034601025080112

10. *Ryazancev A. P.* Social status and reputation of writers in modern Russia: thematic analysis. *Interaktsiya. Interv'yū. Interpretatsiya = Interaction. Interview. Interpretation*. 2023;15(1):27–44. (In Russ.). doi: 10.19181/inter.2023.15.1.2

11. Directory of the Union of writers of the USSR. Moscow: Sovetskii pisatel'; 1981. 821 p. (In Russ.)

12. *Sushchij S. Ja.* Russian literature of the 19th and early 20th centuries: asociodynamic analysis of the writers' community *Sotsiologicheskie issledovaniya = Sociological Research*. 2020;2:128–143. (In Russ.). doi: 10.31857/S013216250008500-3

13. *Sushchij S. Ya.* Russian literature of the Soviet period: the sociodynamics of the writers' community. *Sotsiologicheskie issledovaniya = Sociological Research*. 2021;2:112–129. (In Russ.). doi: 10.31857/S013216250009902-5

14. *Avrachenkov K., Dreveton M.* Statistical analysis of networks. Boston-Delft: Now Publ.; 2020. 248 p. doi: 10.1561/9781638280514

15. *Barabási A. L., Albert R.* Emergence of scaling in random networks. *Science*. 1999;286(5439):509–512. doi: 10.1126/science.286.5439.509

16. *Blondel V. D., Guillaume J.-L., Lambiotte R., Lefebvre E.* Fast unfolding of communities in large networks. *J. Stat. Mech.: Theory Exp.* 2008;10:P10008. doi: 10.1088/1742-5468/2008/10/P10008

17. *Edunov S., Diuk C., Filiz I. O., Bhagat S., Burke M.* Three and a half degrees of separation. *Research at Facebook blog*. 4 February 2016. URL: <https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/> (accessed: 10.01.2026).
18. *Hajarathaiyah K., Enduri M. K., Anamalamudi S., Sangi A. R.* Algorithms for finding influential people with mixed centrality in social networks. *Arab. J. Sci. Eng.* 2023;48(8):10417–10428. doi: 10.1007/s13369-023-07619-w
19. *Mohammad B. B., Dhuli S., Enduri M. K.* Isolating centrality-based generalization of traditional centralities to discover vital nodes in complex networks. *Arab. J. Sci. Eng.* 2025;50(15):12003–12025. doi: 10.1007/s13369-024-09628-9
20. *Mohammad B. B., Dhuli V. S., Enduri M. K., Cenkeramaddi L. R.* A novel convex combination-based mixed centrality measure for identification of influential nodes in complex networks. *IEEE Access*. 2024;12:123897–123920.

doi: 10.1109/ACCESS.2024.3450296

21. *Newman M. E. J., Girvan M.* Finding and evaluating community structure in networks. *Phys. Rev. E*. 2004;69(2):026113. doi: 10.1103/PhysRevE.69.026113
22. *Reichardt J., Bornholdt S.* When are networks truly modular? *Phys. D: Nonlinear Phenom.* 2006;224(1-2):20–26. doi: 10.1016/j.physd.2006.09.009
23. *Schnegg M.* Reciprocity and the emergence of power laws in social networks. *Internat. J. Modern Phys. C*. 2006;17(07):1067–1076. doi: 10.1142/S0129183106009473
24. *Traag V. A., Waltman L., Van Eck N. J.* From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* 2019;9(1):5233. doi: 10.1038/s41598-019-41695-z
25. *Travers J., Milgram S.* An experimental study of the small world problem. *Sociometry*. 1969;32(4):425–443.

Поступила в редакцию / received: 10.04.2026; принята к публикации / accepted: 22.05.2026.  
 Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

#### СВЕДЕНИЯ ОБ АВТОРЕ:

**Бородина Александра Валентиновна**  
 канд. физ.-мат. наук, старший научный сотрудник  
*e-mail: borodina@krc.karelia.ru*

#### CONTRIBUTOR:

**Borodina, Alexandra**  
 Cand. Sci. (Phys.-Math.), Senior Researcher