

Руководство пользователя системы
ИСАНД
isand.ipu.ru/engine

Содержание

1	Введение	3
2	Что такое ИСАНД?	3
3	Техническая информация	5
3.1	Схемы отображения	5
3.2	Общенаучные термины	5
3.3	Уровень	5
3.4	Отсечение по категориям	5
3.5	Отсечение по терминам	5
4	Руководство пользователя	6
4.1	Элементы раздела «Профиль»	9
4.1.1	Основные элементы	9
4.1.2	Выбор авторов и публикаций	9
4.1.3	Выбор схемы отображения	10
4.1.4	Учет общенаучных терминов	11
4.1.5	Уровень	11
4.1.6	Отсечение по категориям	11
4.1.7	Отсечение по терминам	12
4.1.8	Выбор пути отображения	12
4.1.9	Отсечение по годам	12
4.2	Элементы раздела «Ранжирование по категориям»	12
4.2.1	Основные элементы	12
4.2.2	Уровень	13
4.2.3	Область тезауруса	13
4.2.4	Выбор авторов	14
4.2.5	Все термины	14
4.2.6	Рейтинг статей	15
4.3	Элементы раздела «Графы»	15
4.3.1	Основные элементы	15
4.3.2	Выбор автора	16
4.3.3	Отсечение по частоте	16
4.3.4	Количество статей	17
4.3.5	Временная шкала	17
4.3.6	Общенаучные термины	18
4.3.7	Включение висячих узлов	18
4.3.8	Отображение названий терминов	19
4.3.9	Уровень терминов	19
4.4	Элементы раздела «Тезаурус»	20
4.4.1	Основные элементы	20
4.4.2	Выбор автора	20
4.4.3	Выбор раздела	21
4.4.4	Корень тезауруса	22
4.4.5	Режим расцветки	22
4.4.6	Подсветка терминов	23
4.4.7	Отображение названий терминов	24
4.5	Элементы раздела «Расстояния»	24
4.5.1	Основные элементы	25

4.5.2	Выбор авторов	25
4.5.3	Выделение участников конференций	26
4.5.4	Добавление проекций векторов конференций	26

1 Введение

В данной инструкции предоставлены указания по использованию веб-сайта isand.ipu.ru/engine. Пользователь может ознакомиться с основными функциями сайта, получить информацию по использованию веб-ресурса.

2 Что такое ИСАНД?

ИСАНД (Информационная Система Анализа Научной Деятельности) – система для определения области компетенции научных сотрудников по их работам: статьями, тезисам, монографиям, отдельным главам.

Основная гипотеза заключается в положительной корреляции между компетенцией автора в области X , и частотой использования автором в своих работах терминов, ассоциированных с областью X . Так, если автор использует термины из криптографии достаточно часто и последовательно: «целостность», «открытый ключ», «атака» и другие, то весьма вероятно что автор компетентен в этой области.

Предполагается, что использование терминов из области X является допустимым показателем компетенции авторов в этой области, поскольку именно в терминах заключён скелет семантики текста: невозможно написать работу по криптографии не воспользовавшись множеством её терминов. Обратное в общем случае не верно, и этот факт является основным источником ошибок в определении компетенции моделью ИСАНД.

С одной стороны, ассоциация терминов и областей знаний должна быть достаточно однозначной чтобы один термин принадлежал как можно более меньшему числу областей. С другой стороны сами области не должны быть слишком общими на практике.

С учётом этих требований, была разработана трёхуровневая иерархия связанных с теорией управления областей знаний, называемая тезаурусом. Тезаурус представляет собой граф в виде трёхуровневого дерева областей знаний (см. рис. 1). Область знаний из вершины потомка входит в область знаний вершины предка подобно тому, как «Группы и алгебры Ли» входит в «Алгебру и теорию чисел».

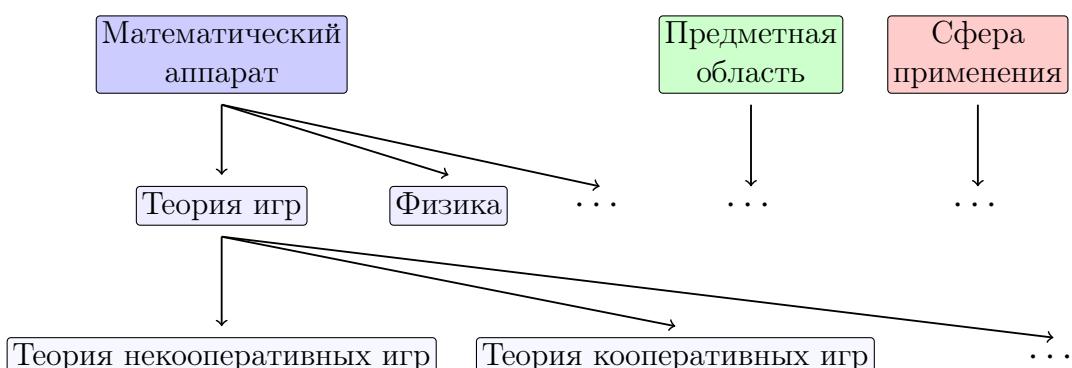


Рис. 1: Одно из поддеревьев дерева тезауруса

Профиль автора представляет собой назначение каждой вершине тезауруса положительного числа, называемого показателем компетенции, отражающего

компетентность автора в соответствующей области. Показатель компетенции вершины не может быть меньше чем этот показатель у её потомков.

Естественным образом концепция профиля может быть применена не только к отдельному автору, но и к целой группе учёных, организации, институту, журналу.

Такой профиль может быть использован следующим образом:

1. определение научных интересов автора, группы авторов, организации; в текущий или любой другой момент времени; прослеживание тенденций по их изменению;
2. поиск компетентных в данной научной сфере людей: рецензентов, докладчиков на конференцию, участников нового проекта, научных руководителей.

Существует несколько методов расчёта показателей компетенции, однако все они изначально опираются на число включений Δ_i^l термина i в рассматриваемом тексте l . Каждый метод по-своему решает задачу минимизацию ошибок определения профиля.

На данный момент для расчёта профилей модель ИСАНД использует базу данных работ, ядро которой составляет база данных ПРНД. Потому профиль может быть составлен только на основании анализа текстов работ только из этой базы данных.

3 Техническая информация

3.1 Схемы отображения

- «Абсолютный вектор» отражает суммарное количество вхождений терминов;
- «Стохастический вектор» — абсолютный вектор с нормализованными столбцами;
- «Булев вектор» — компоненты этого вектора могут принимать два значения: единица — если количество терминов больше значения «отсечение по терминам» и ноль — в остальных случаях;
- «По количеству использованных терминов» — вариант абсолютного вектора, когда «отсечение по терминам» убирает те столбцы, где «的独特性» терминов меньше, чем значение «отсечение по терминам». В случае абсолютного вектора «отсечение по терминам» работает с общим числом вхождений терминов, а в данном случае — с уникальным;
- «Термины» — просмотр терминов.

3.2 Общенаучные термины

Выставленный флагок «учитывать общенаучные термины» позволяет отображать на графике те термины, которые относятся к двум и более научным областям, и указывает их количество.

3.3 Уровень

Ползунок «уровень» регулирует глубину поддерева: все вершины с одним потомком, путь от которых до корня текущего поддерева больше значения ползунка «уровень», объединяются в одну фиктивную вершину, в которой лежат все термины этих вершин.

3.4 Отсечение по категориям

Для абсолютного вектора определяется составляющая с максимальным значением. Это значение умножается на значение ползунка, которое может меняться от нуля до единицы, и в результате получается число X. Затем при построении графика убираются несущественные составляющие, значение которых меньше X, чтобы сделать график более выразительным.

3.5 Отсечение по терминам

«Отсечение по терминам» убирает те столбцы, в которых число появившихся в тексте терминов меньше заданного значения. Это нужно, чтобы не учитывались те подфакторы, в которых использовано лишь немного терминов из возможных.

4 Руководство пользователя

После перехода по ссылке isand.ipu.ru/engine пользователь попадает на страницу (рис. 2):



Рис. 2: Главная страница

После нажатия на кнопку «Профиль» пользователь попадает на страницу (рис. 3):

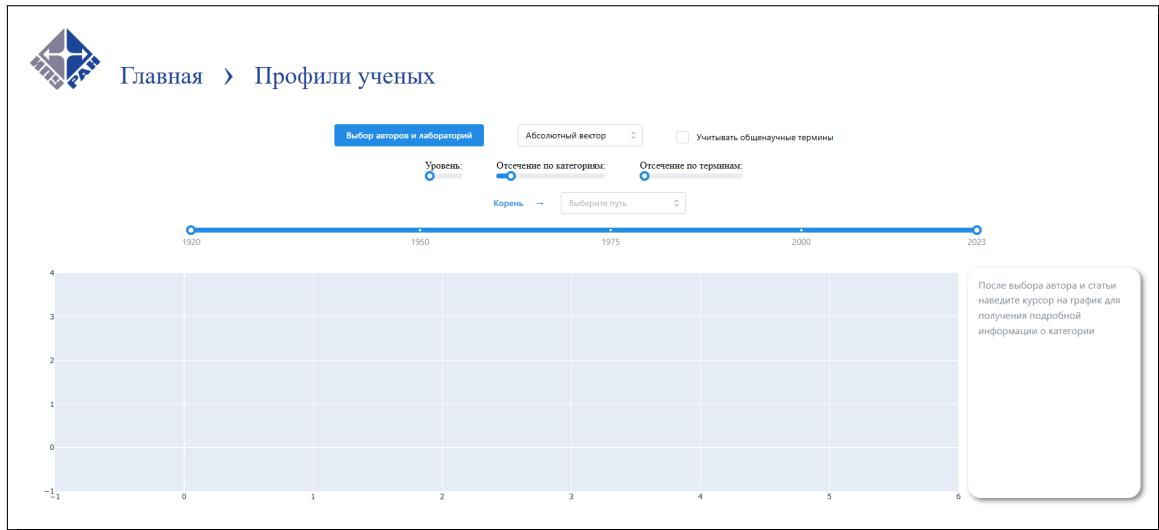


Рис. 3: Страница с визуализацией

После нажатия на кнопку «Ранжирование по категориям» пользователь попадает на страницу (рис. 4):



Рис. 4: Страница с ранжированием

После нажатия на кнопку «Графы» пользователь попадает на страницу (рис. 5):



Рис. 5: Страница с графиками связности терминов

После нажатия на кнопку «Тезаурус» пользователь попадает на страницу (рис. 6):

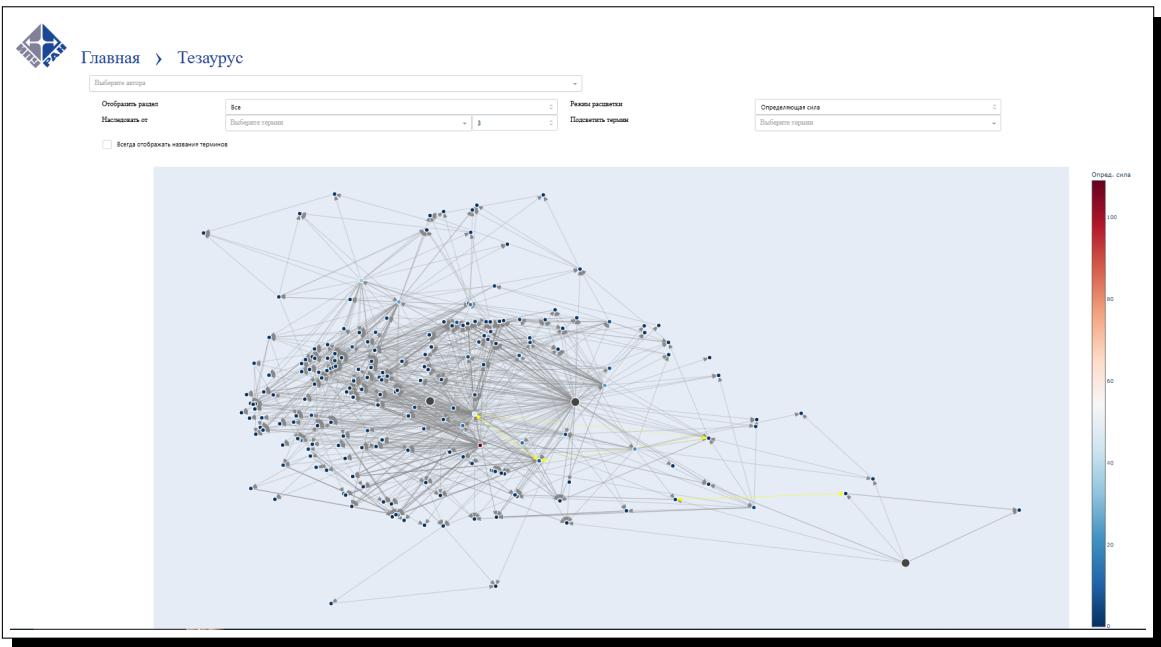


Рис. 6: Страница с графиком тезауруса

После нажатия на кнопку «Расстояния» пользователь попадает на страницу (рис. 7):



Рис. 7: Страница с проекциями дельт ученых и конференция на двумерное пространство

Рассмотрим подробнее элементы данных страниц.

4.1 Элементы раздела «Профиль»

4.1.1 Основные элементы

Страницу можно разделить на три основные области (рис. 8):

1. Меню
2. Область отображения графика
3. Информация о терминах

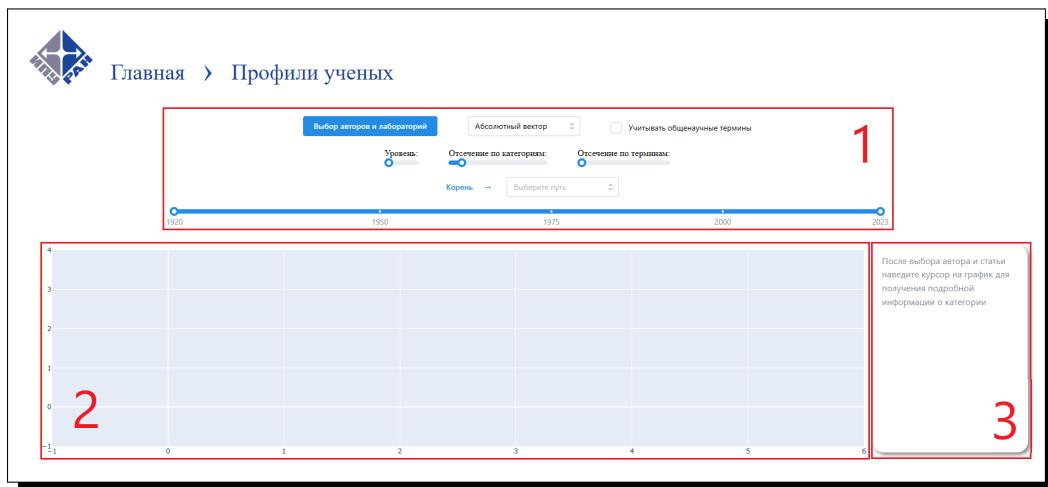


Рис. 8: Основные элементы страницы «Профиль»

Меню служит для выбора автора и публикаций, а также содержит фильтры для отображения информации.

4.1.2 Выбор авторов и публикаций

В левом верхнем углу расположена кнопка «Выбор публикаций» (рис. 9):

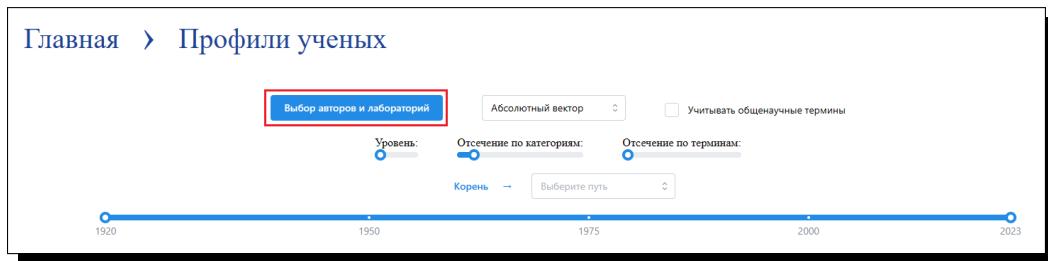


Рис. 9: Кнопка для доступа к выбору авторов и публикаций

После нажатия на кнопку пользователю открывается меню выбора авторов и их публикаций в левой части экрана (рис. 10):

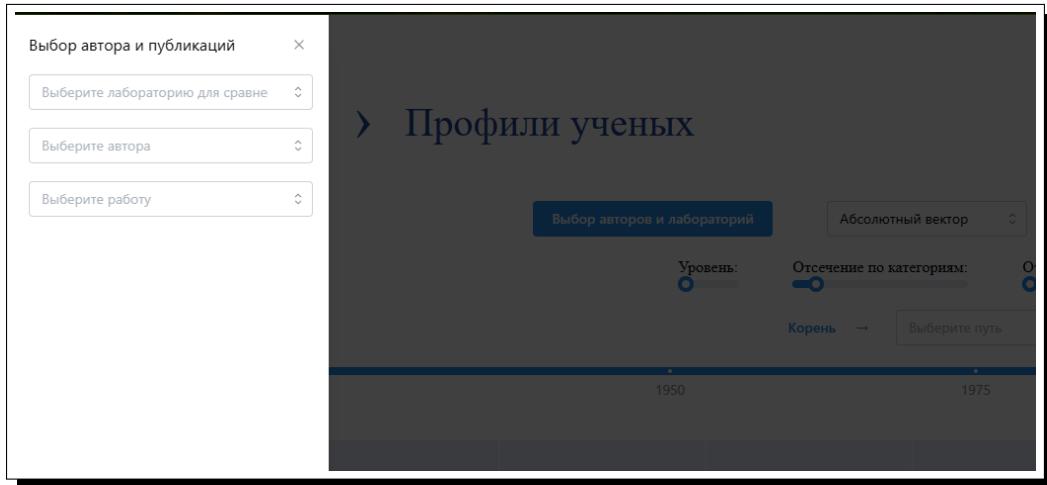


Рис. 10: Три выпадающих списка для выбора авторов и публикаций

Доступен выбор:

- одного автора и некоторых его публикаций;
- одного автора и всех его публикаций (пункт «Все работы» третьего выпадающего списка);
- нескольких авторов (автоматический выбор пункта «Все работы» третьего выпадающего списка);
- всех авторов и всех публикаций;
- всех авторов и/или публикаций, относящихся к выбранным организациям (лабораториям института), научным журналам, конференциям.

4.1.3 Выбор схемы отображения

Пользователю доступен выбор схем отображения данных (рис. 11):

- Абсолютный вектор
- Стохастический вектор
- Булев вектор
- По количеству используемых терминов
- Термины

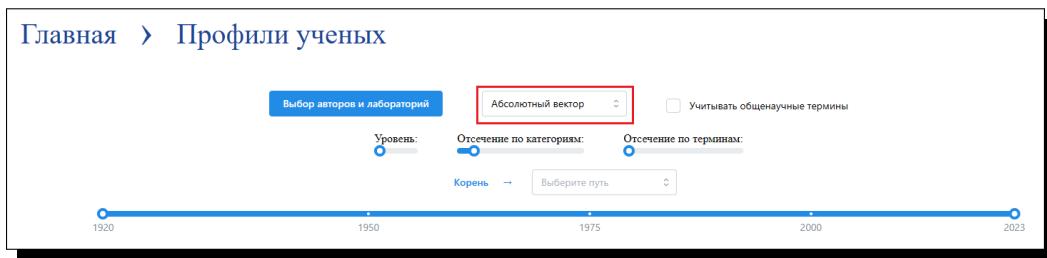


Рис. 11: Выбор схемы отображения

4.1.4 Учет общенациональных терминов

Если пользователь желает учитывать общенациональные термины при отображении данных, то необходимо поставить галочку в данном окошке (рис. 12):

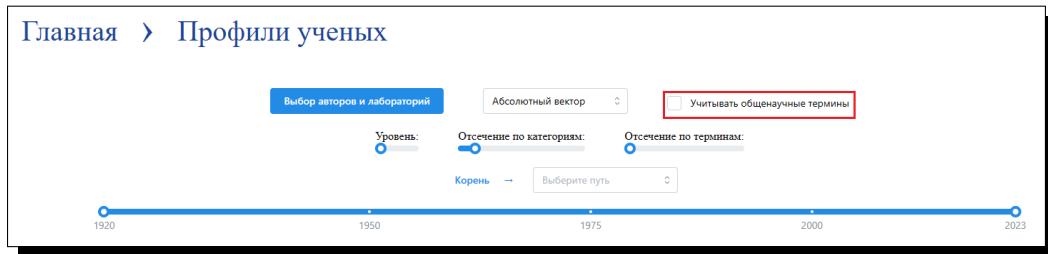


Рис. 12: Учет общенациональных терминов

4.1.5 Уровень

Ползунок «Уровень» влияет на отображение научных областей. Чем левее находится ползунок, тем более обширные научные области будут выводиться по вертикальной оси графика. Чем правее ползунок, тем более узконаправленными будут научные области.

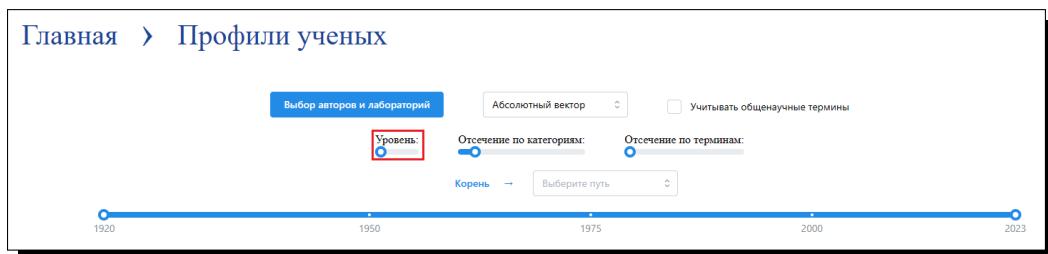


Рис. 13: Ползунок «Уровень»

4.1.6 Отсечение по категориям

Ползунок «Отсечение по категориям», согласно выбранному значению, будет отсекать те научные области, в которых количество входящих терминов меньше некоторого значения (рис. 14):

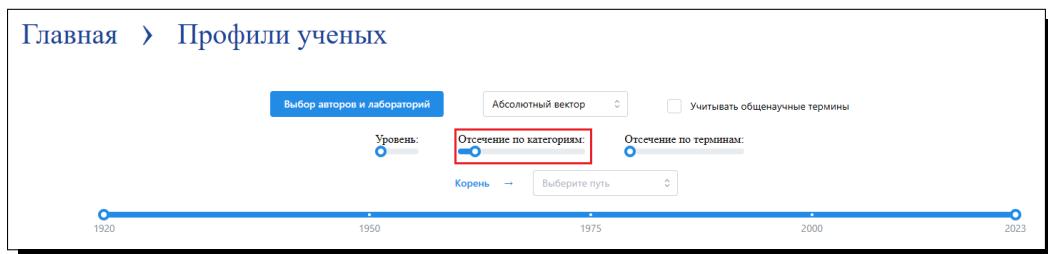


Рис. 14: Ползунок «Отсечение по категориям»

4.1.7 Отсечение по терминам

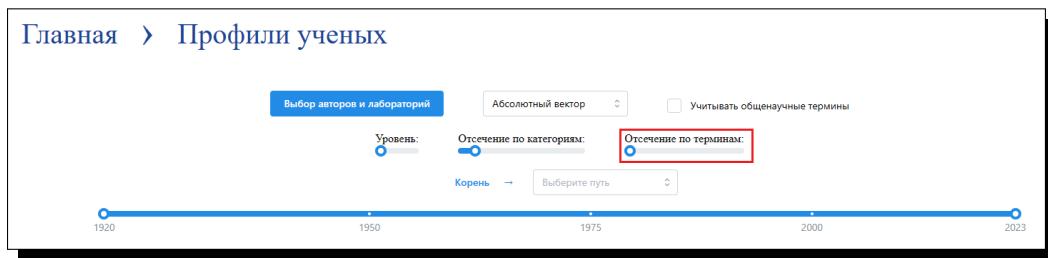


Рис. 15: Ползунок «Отсечение по терминам»

4.1.8 Выбор пути отображения

В выпадающем списке «Выбор пути» (рис. 16) отображаются показываемые в текущий момент на графике научные области. При выборе того или иного доступного варианта график изменится: на оси ОУ будут отображаться подкатегории выбранной научной области.

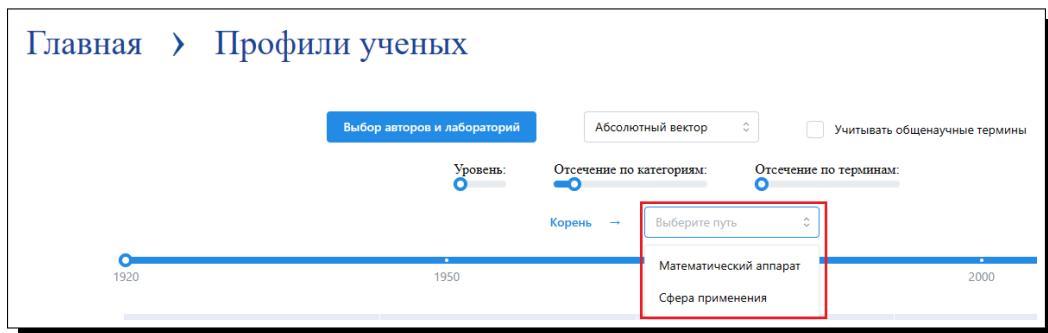


Рис. 16: Выбор пути отображения

4.1.9 Отсечение по годам

Ползунок отсечение по годам позволяет пользователю фильтровать публикации по времени их выхода.

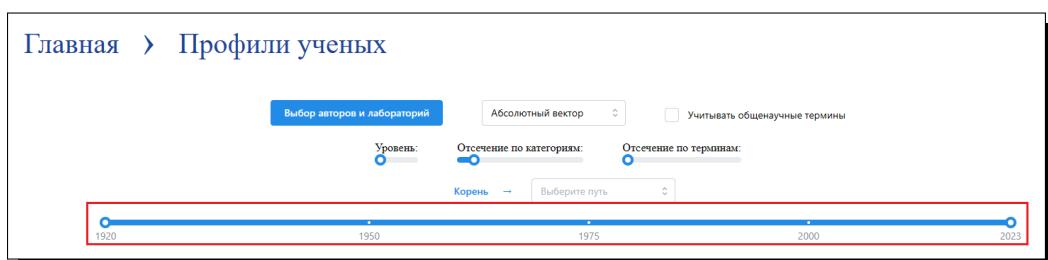


Рис. 17: Ползунок «Отсечение по годам»

4.2 Элементы раздела «Ранжирование по категориям»

4.2.1 Основные элементы

Страницу можно разделить на три основные области (рис. 18):

1. Меню

2. Область отображения графиков
3. Информация о терминах выбранного автора

Меню служит для выбора автора и публикаций, а также содержит фильтры для отображения информации. На графиках пользователь может увидеть список авторов, чьи работы подходят под заданные фильтры, а также, при наведении, количество этих работ. В области отображения информации о терминах пользователь может увидеть наиболее часто встречающиеся термины в работах выбранного автора.

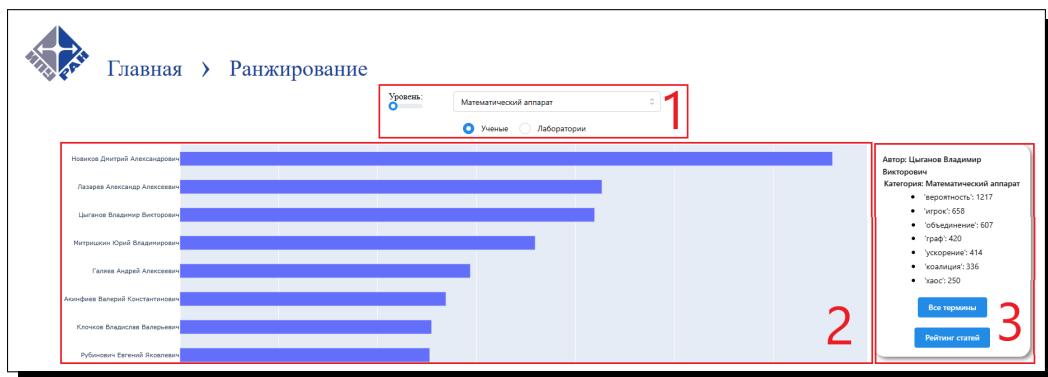


Рис. 18: Основные элементы страницы «Ранжирование»

4.2.2 Уровень

Ползунок «Уровень» (рис. 19) влияет на отображение научных областей. Чем левее находится ползунок, тем более обширные научные области будут учитываться при подборе авторов. Чем правее ползунок, тем более узконаправленными будут научные области.

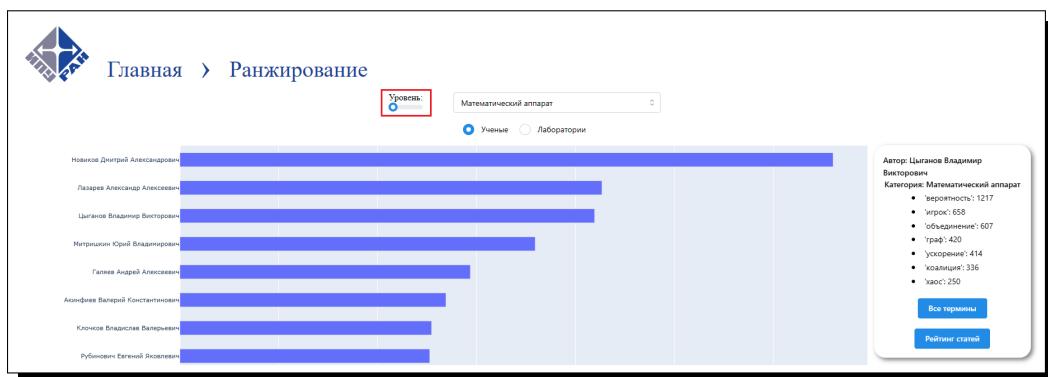


Рис. 19: Ползунок «Уровень»

4.2.3 Область тезауруса

В выпадающем списке «Выбор области тезауруса» (рис. 20) пользователь может выбрать и конкретизировать область тезауруса, по которой необходимо выполнить поиск.



Рис. 20: Список «Выбор области тезауруса»

4.2.4 Выбор авторов

При помощи переключателя «Выбор авторов» (рис. 21) пользователь может выбирать - отбирать публикации по людям или по лабораториям.

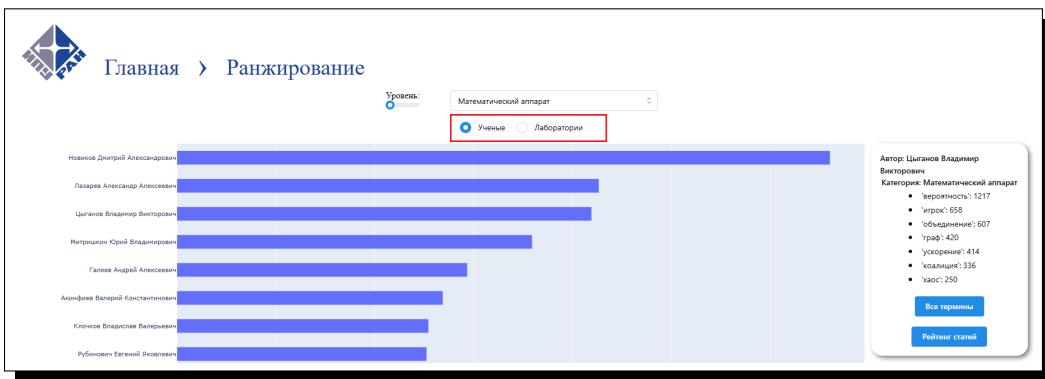


Рис. 21: Переключатель «Выбор авторов»

4.2.5 Все термины

Кнопка «Все термины» (рис. 22) позволяет пользователю вывести список всех терминов, использующихся в работах выбранного автора, а также количество вхождений термина в этих работах.

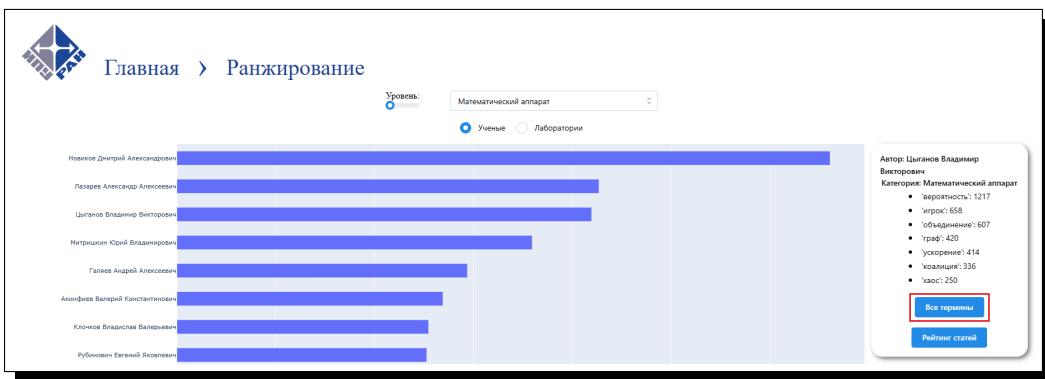


Рис. 22: Кнопка «Все термины»

4.2.6 Рейтинг статей

Кнопка «Рейтинг статей» (рис. 23) позволяет пользователю вывести список статей, наиболее сильно повлиявших на конкретную категорию у конкретного автора.

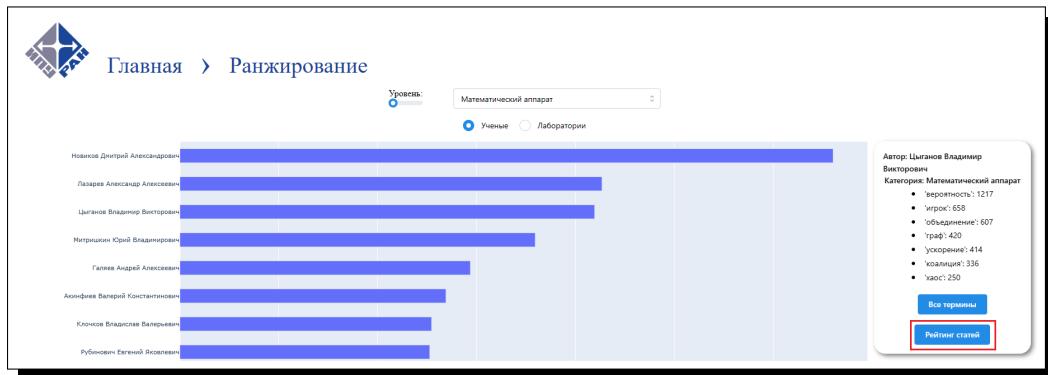


Рис. 23: Кнопка «Рейтинг статей»

4.3 Элементы раздела «Графы»

Страницу можно разделить на две основные области (рис. 24):

1. Меню
2. Область отображения графика

4.3.1 Основные элементы



Рис. 24: Основные элементы страницы «Графы»

4.3.2 Выбор автора



Рис. 25: Выбор автора

В выпадающем списке «Выбор автора» (рис. 25) пользователь может выбрать автора, на основе публикационной активности которого строится граф связности терминов. Если автор не выбран, то на экран выводится случайно сгенерированный граф.

4.3.3 Отсечение по частоте



Рис. 26: Ползунок «Отсечение по частоте»

Выбирая промежуток на ползунке «Отсечение по частоте» (рис. 26), пользователь устанавливает минимальное и максимальное значение встречаемости термина в статье. Термины, встречаемости которых попадают в промежуток, добавляются на график.

4.3.4 Количество статей



Рис. 27: Ползунок «Количество статей»

Выбирая значение на ползунке «Количество статей» (рис. 27), пользователь определяет предельное количество статей, термины из которых используются при построении графа. Данный лимит введен, чтобы не перегружать график, если ученого очень много публикаций.

4.3.5 Временная шкала

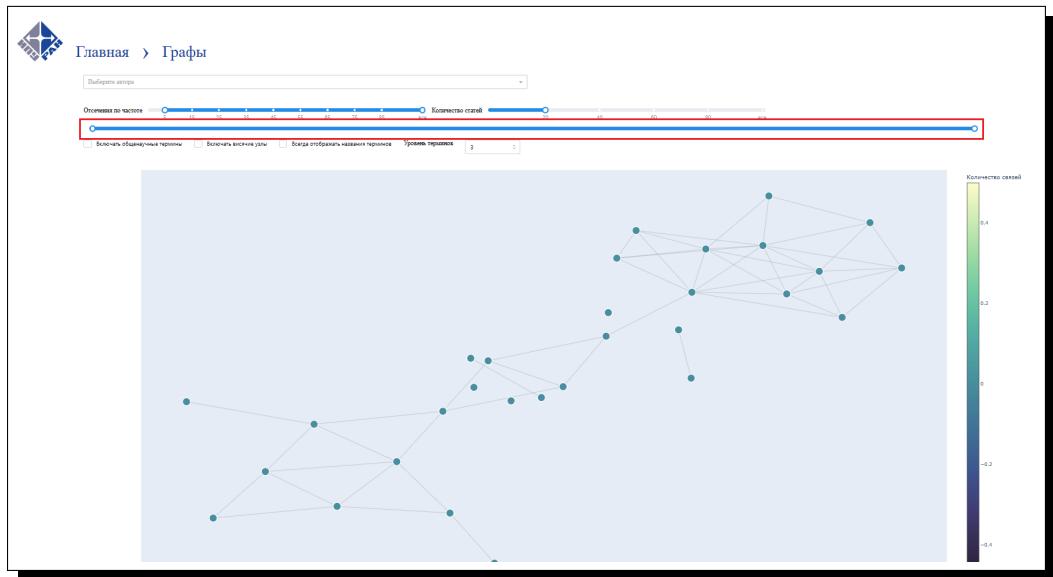


Рис. 28: Ползунок «Временная шкала»

Выбирая промежуток на ползунке «Временная шкала» (рис. 28), пользователь определяет диапазон лет. Только статьи, опубликованные в эти годы, учитываются при построении графа. По умолчанию промежуток включает весь период публикационной активности автора.

4.3.6 Общенаучные термины



Рис. 29: Переключатель «Общенаучные термины»

Если пользователь желает учитывать общенаучные термины при отображении данных, то необходимо поставить галочку в окошке переключателя «Общенаучные термины» (рис. 29).

4.3.7 Включение висячих узлов



Рис. 30: Переключатель «Включать висячие узлы»

Если пользователь хочет, чтобы на графике присутствовали термины, не имеющие связей с другими терминами, то необходимо поставить галочку в окошке переключателя «Включать висячие узлы» (рис. 30).

4.3.8 Отображение названий терминов



Рис. 31: Переключатель «Всегда отображать названия терминов»

Если пользователь хочет, чтобы названия научных терминов отображались постоянно, а не только при наведении на вершину, то необходимо поставить галочку в окошке переключателя «Всегда отображать названия терминов» (рис. 31).

4.3.9 Уровень терминов



Рис. 32: Окно выбора «Уровень терминов»

Окно выбора «Уровень терминов» (рис. 32) влияет на отображение научных областей. Чем меньше уровней, тем из более обширных научных областей будут происходить термины. Чем больше уровней, тем более узконаправленными будут научные области.

4.4 Элементы раздела «Тезаурус»

Страницу можно разделить на две основные области (рис. 18):

1. Меню
 2. Область отображения графика

4.4.1 Основные элементы

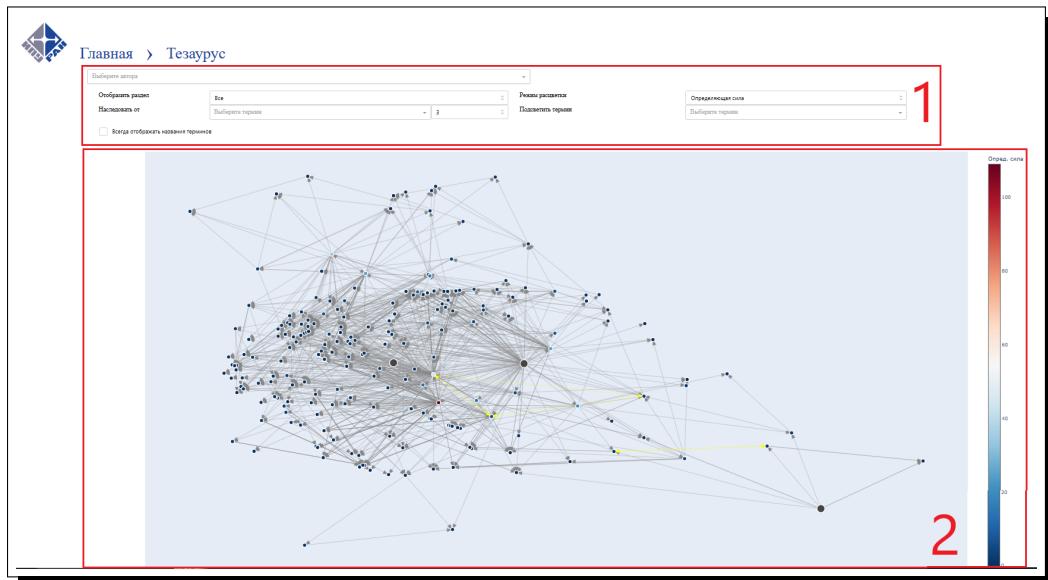


Рис. 33: Основные элементы страницы «Тезаурус»

4.4.2 Выбор автора

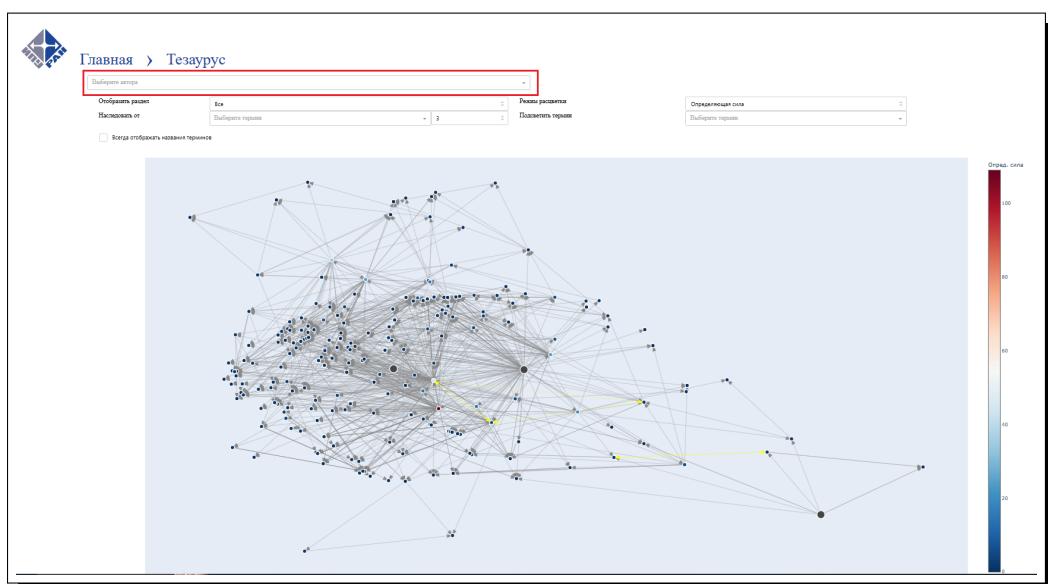


Рис. 34: Выбор автора

В выпадающем списке «Выбор автора» (рис. 34) пользователь может выбрать автора, публикационная активность которого накладывается на тезаурус. Термины, которые этот автор использовал, подсвечиваются на графике тезауруса.

4.4.3 Выбор раздела

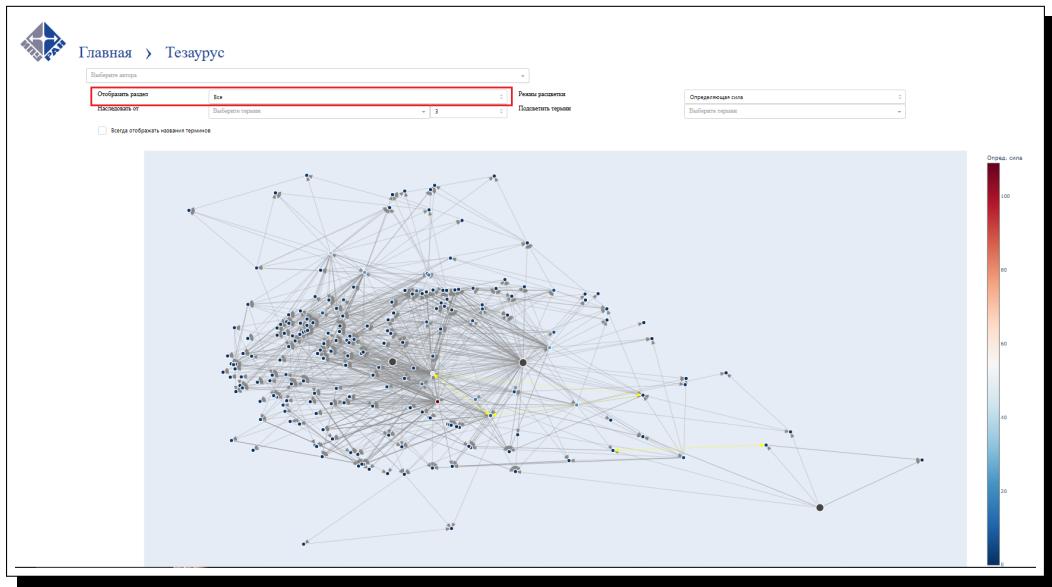


Рис. 35: Окно выбора «Отобразить раздел»

В окне «Отобразить раздел» (рис. 35) пользователь может выбрать, какой раздел тезауруса по теории управления отобразить. По умолчанию отображается весь тезаурус. Доступные для отображения разделы:

- основные термины
- объекты управления
- воздействия и сигналы
- виды управления
- управляющие объекты
- системы управления
- законы управления
- элементы систем управления
- структуры систем управления
- состояния систем управления
- свойства систем управления
- общие термины

4.4.4 Корень тезауруса

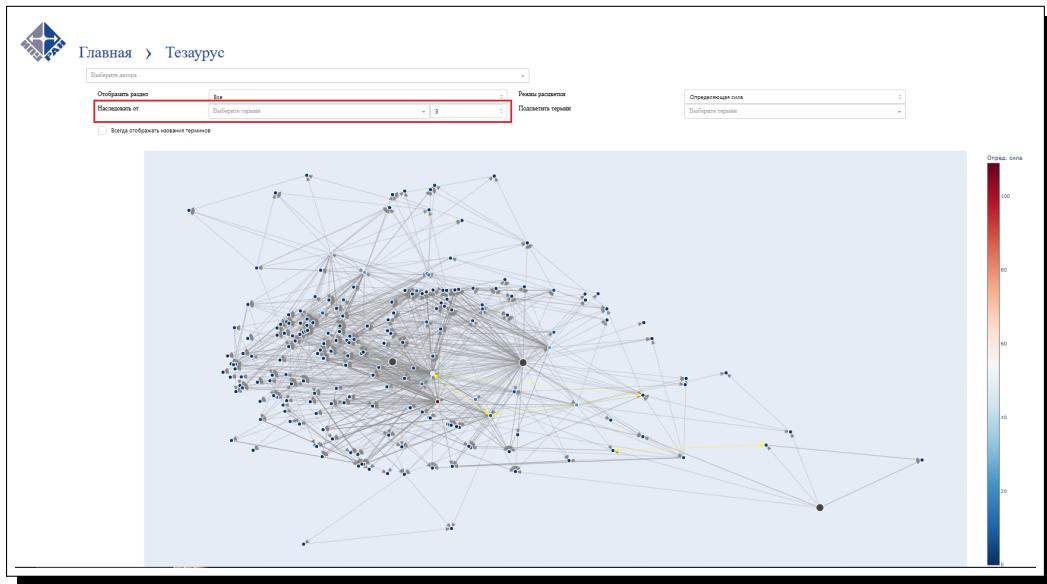


Рис. 36: Окна выбора «Наследовать от» и «Глубина»

В данном окне (рис. 36) пользователь может выбрать, какой термин использовать в качестве корня тезауруса. При построении тезауруса от заданного корня, в граф включаются только те термины, которые определяются через заданный корень. В соседнем окне «Глубина» (рис. 36) можно выбрать, насколько глубоко по цепочке определений строится граф.

4.4.5 Режим расцветки

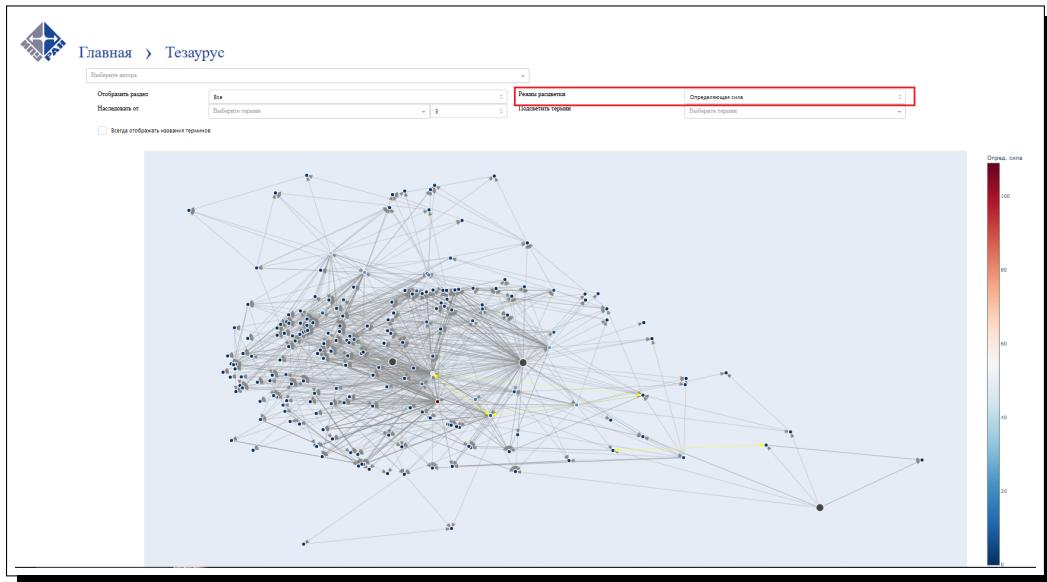


Рис. 37: Окно выбора «Режим расцветки»

В данном окне (рис. 37) пользователь может выбрать правило, в соответствии с которым вершинам графа назначается цвет. Доступные режимы расцветки:

- Определяющая сила

Цвет вершины, соответствующей некоторому термину задается отношением количества терминов, определенных через данный термин к количеству терминов, через которые определяется он.

- Раздел

Цвет вершины, соответствующей некоторому термину, задается разделом, к которому данный термин относится.

4.4.6 Подсветка терминов

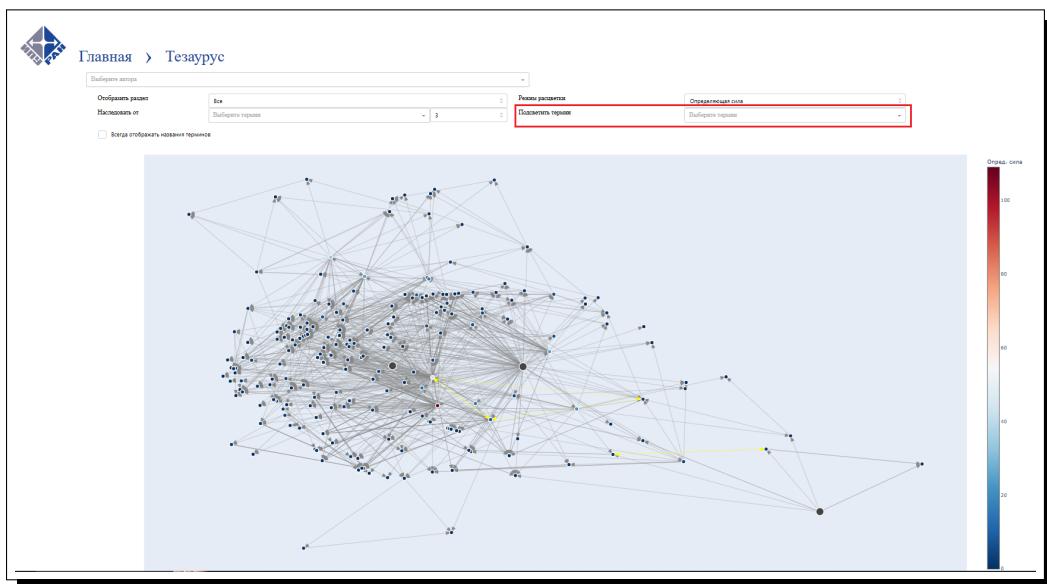


Рис. 38: Окно выбора «Подсветить термин»

В данном окне (рис. 38) пользователь может выбрать, какой термин выделить на графике. Все термины, помимо выделенного, будут отображаться полупрозрачными, ребра, исходящие из него, будут подсвечены красным, а входящие – синим.

4.4.7 Отображение названий терминов

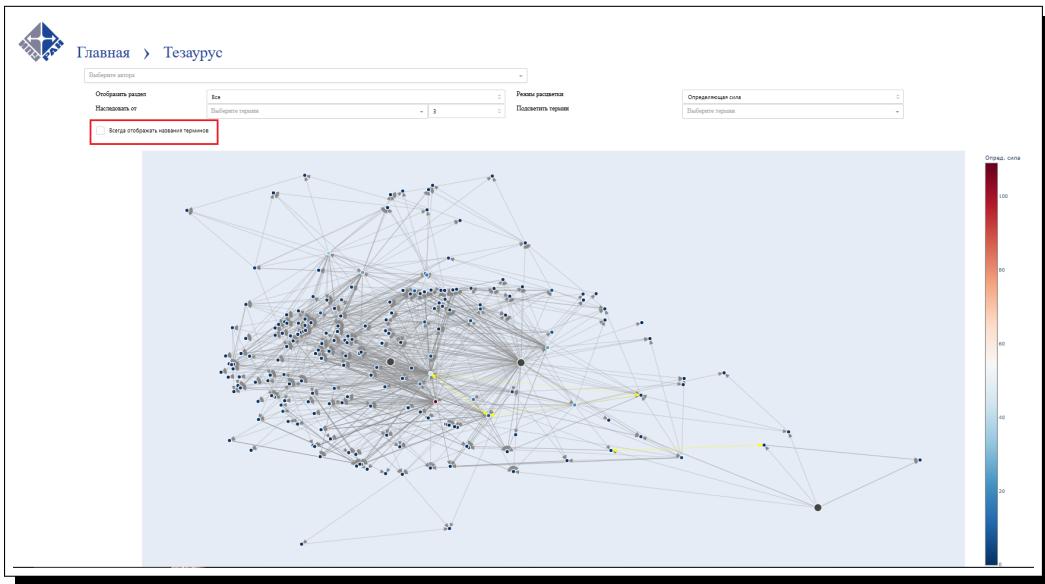


Рис. 39: Переключатель «Всегда отображать названия терминов»

Если пользователь хочет, чтобы названия научных терминов отображались постоянно, а не только при наведении на вершину, то необходимо поставить галочку в окошке переключателя «Всегда отображать названия терминов» (рис. 39).

4.5 Элементы раздела «Расстояния»

Страницу можно разделить на две основные области (рис. 18):

1. Меню
2. Область отображения графика

4.5.1 Основные элементы



Рис. 40: Основные элементы страницы «Расстояния»

4.5.2 Выбор авторов



Рис. 41: Окно выбора авторов

В выпадающем списке «Выбор авторов» (рис. 41) пользователь может выбрать авторов, чьи проекции векторов необходимо выделить на отображении.

4.5.3 Выделение участников конференций

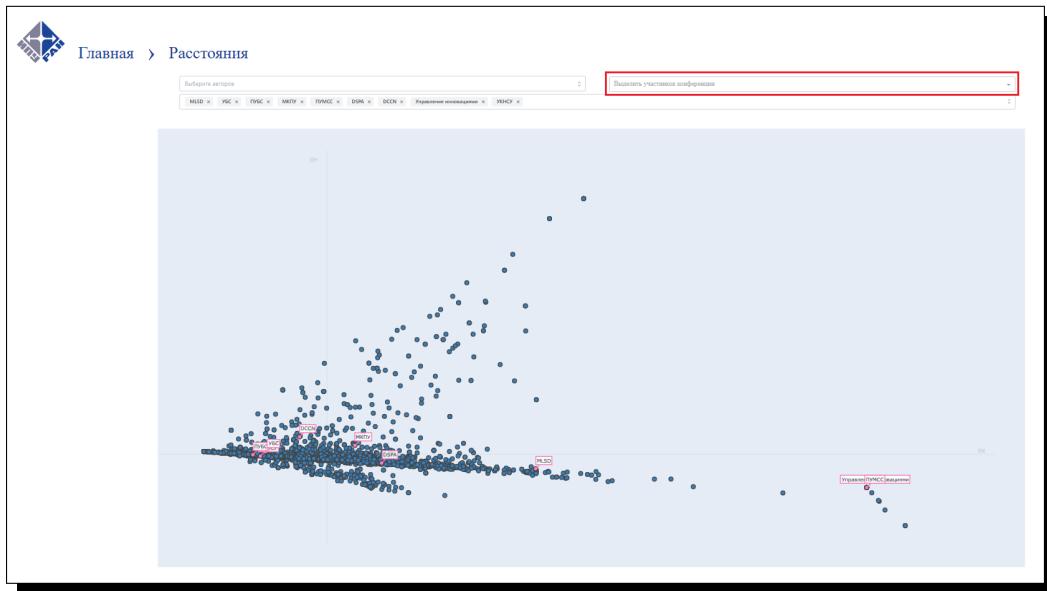


Рис. 42: Окно выбора «Выделить участников конференции»

В данном окне (рис. 42) пользователь может выбрать конференцию, проекции векторов участников которой необходимо выделить на отображении.

4.5.4 Добавление проекций векторов конференций

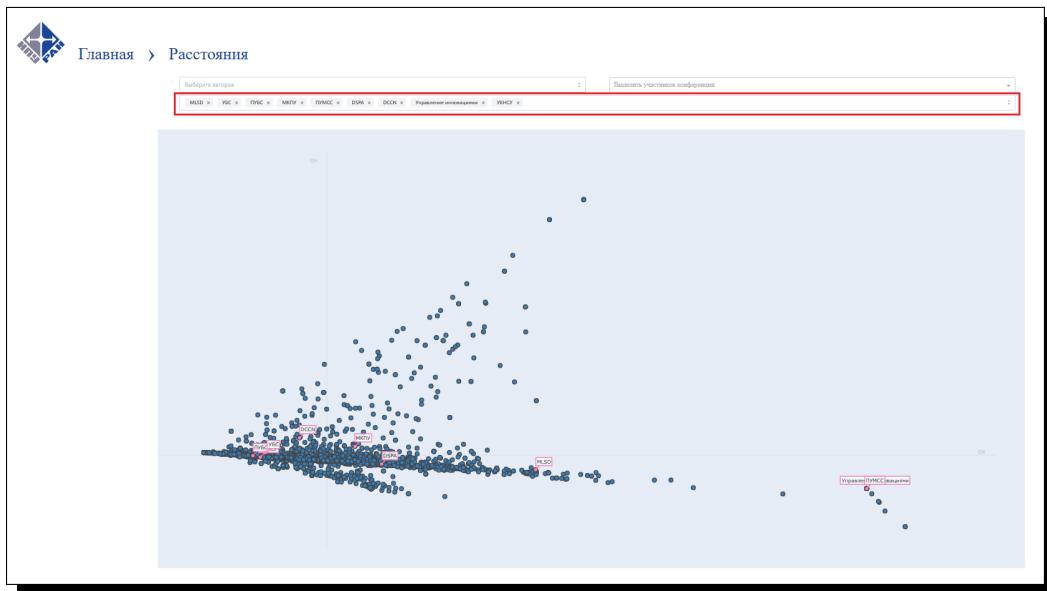


Рис. 43: Окно выбора «Выделить конференции»

В данном окне (рис. 43) пользователь может выбрать конференции, чьи проекции векторов необходимо выделить на отображении. Вектор конференции строится по работам, которые были представлены на ней за последние пять лет. По умолчанию выделены все доступные конференции.