



КРИБРУМ
Мы слушаем сеть

Описание системы

Система графического анализа распространения информационного сигнала в социальных сетях «Крибрум. Сигнал»

Москва, 2020

Аннотация

В данном документе приведено описание системы графического анализа распространения информационного сигнала в социальных сетях «Крибрум. Сигнал».

Система графического анализа распространения информационного сигнала в социальных сетях «Крибрум. Сигнал» предназначена для определения источников и путей распространения информации в социальных сетях:

- Facebook;
- ВКонтакте;
- Twitter;
- Instagram.

Система «Крибрум. Сигнал» реализует функции:

- поиска информации в хранилище данных Крибрум с платформ социальных сетей в сети Интернет;
- определения времени публикаций в социальных сетях и путей наследования данных аккаунтами социальных сетей;
- определение первоисточников информации и цепочек ее распространения;
- визуализации характера распространения информации в социальных сетях в графическом интерфейсе пользователя.

Данные функции системы «Крибрум. Сигнал» позволяют пользователю легко оценивать характер распространения информации в социальных медиа, определять естественный или искусственный характер распространения.

ОГЛАВЛЕНИЕ

1	ОБЩИЕ СВЕДЕНИЯ О СИСТЕМЕ	6
1.1	Обозначение и наименование системы	6
1.2	Разработчик системы.....	6
1.2.1	Программное обеспечение, необходимое для функционирования системы	6
1.3	Описание системы	6
1.3.1	Назначение системы.....	6
1.4	Задачи системы	6
1.5	Доступ к системе	6
1.5.1	Структура интерфейса системы.....	7
1.5.2	Область создания и настройки параметров графа	8
1.5.3	Область «Графы».....	9
1.5.4	Страница построения и просмотра сформированных графов	9
2	ФУНКЦИОНАЛЬНОЕ ОПИСАНИЕ РАЗДЕЛОВ СИСТЕМЫ.....	11
2.1	Главная страница	11
2.1.1	Область создания и настройки параметров графа	11
2.1.2	Область «Графы».....	12
2.2	Страница построения и просмотра сформированных графов	14
2.3	Страница «Подробности»	15
2.4	Страница «Статистика»	16
2.5	Примеры и описания графов	17

ОСНОВНЫЕ ТЕРМИНЫ И ПОНЯТИЯ

Автор – пользователь, от имени которого опубликовано сообщение на интернет ресурсе (площадке мониторинга).

Анализ графа распространения информации в социальных сетях – процесс анализа графа, построенного системой «Крибрум. Сигнал», с целью определения характера распространения информации в социальных сетях.

Граф - абстрактный математический объект, представляющий собой множество вершин графа (узлов) и набор рёбер (связей), то есть соединений между парами вершин. Графы являются основным объектом изучения теории графов.

Источник сообщения – площадка мониторинга, на которой обнаружено релевантное сообщение по отношению к поисковому запросу.

Оригинал - сообщение, опубликованное первым в ряду одинаковых или частично повторяющихся.

Перепечатки - ряд одинаковых или частично повторяющихся друг друга сообщений.

Платформы - социальные сети, по которым осуществляет поиск система «Крибрум. Сигнал» (Facebook, Вконтакте, Instagram, Twitter).

Поисковый запрос - последовательность символов, которую пользователь вводит в поисковую строку, чтобы найти интересующую его информацию. Чаще всего, поисковый запрос задаётся в виде набора слов или фразы, иногда используются расширенные возможности языка запросов поисковой системы.

Распространение информации - процесс, посредством которого некоторый информационный объект (информация, вирус, мнение) распространяется по коммуникационным каналам во времени и в пространстве среди узлов сети.

Система – в данном документе система поиска по социальным медиа «Крибрум. Сигнал».

Сообщение – это отдельная текстовая публикация в Интернете, содержащая осмысленное упоминание объекта мониторинга на одной из площадок мониторинга. Сообщение может иметь вид записи, комментария, поста, новости, заметки, статьи, рецензии в блоге, микроблоге, социальной сети, форуме, онлайн-СМИ, интернет-магазине или другом виде информационных Интернет-ресурсов и социальных медиа.

Социальный граф – система связей между людьми в социальных сетях, в сети Интернет или в реальной жизни.

Теория графов - раздел дискретной математики, изучающий свойства графов. В общем смысле граф представляется как множество вершин (узлов), соединённых рёбрами (связями).

Cassandra - система управления базами данных ¹, используемая в подсистеме сбора данных Крибрум.

RabbitMQ - платформа, реализующая систему обмена сообщениями².

PostgreSQL - свободная объектно-реляционная система управления базами данных (СУБД)³, используемая в системе «Крибрум.Сигнал».

¹ **Apache Cassandra** — распределённая система управления базами данных, относящаяся к классу NoSQL-систем и рассчитанная на создание высокомасштабируемых и надёжных хранилищ огромных массивов данных, представленных в виде хэша.

² **RabbitMQ** — платформа, реализующая систему обмена сообщениями между компонентами программной системы (Message Oriented Middleware) на основе стандарта AMQP (Advanced Message Queuing Protocol). RabbitMQ выпускается под Mozilla Public License.

RabbitMQ создан на основе испытанной Open Telecom Platform, обеспечивающий высокую надёжность и производительность промышленного уровня и написан на языке Erlang.

³ **PostgreSQL** (произносится «Пост-Грэс-Кью-Эл») — свободная объектно-реляционная система управления базами данных (СУБД). Существует в реализациях для множества UNIX-подобных платформ, включая AIX, различные BSD-системы, HP-UX, IRIX, Linux, macOS, Solaris/OpenSolaris, Tru64, QNX, а также для Microsoft Windows.

Сильными сторонами PostgreSQL считаются:

- высокопроизводительные и надёжные механизмы транзакций и репликации;
- расширяемая система встроенных языков программирования: в стандартной поставке поддерживаются PL/pgSQL, PL/Perl, PL/Python и PL/Tcl; дополнительно можно использовать PL/Java, PL/PHP, PL/Py, PL/R, PL/Ruby, PL/Scheme, PL/sh и PL/V8, а также имеется поддержка загрузки C-совместимых модулей;
- наследование;
- легкая расширяемость.

1 ОБЩИЕ СВЕДЕНИЯ О СИСТЕМЕ

1.1 Обозначение и наименование системы

Наименование: «Крибрум.Сигнал».

Обозначение: Система графического анализа распространения информационного сигнала в социальных сетях «Крибрум. Сигнал».

1.2 Разработчик системы

Разработчиком системы является АО «Крибрум».

1.2.1 Программное обеспечение, необходимое для функционирования системы

Система мониторинга и анализа поведения аккаунтов в социальных медиа «Крибрум. Сигнал» является мульти-платформенной. Для использования системы достаточно наличие любого современного браузера.

1.3 Описание системы

1.3.1 Назначение системы

Система «Крибрум. Сигнал» предназначена для быстрого и легкого анализа источников и путей распространения информации в социальных сетях.

1.4 Задачи системы

Задачами системы являются:

- поиск информации в хранилище данных Крибрум с платформ социальных сетей в сети Интернет;
- определение времени публикаций в социальных сетях и путей наследования данных аккаунтами социальных сетей;
- определение первоисточников информации и цепочек ее распространения;
- визуализация характера распространения информации в социальных сетях в графическом интерфейсе пользователя.

1.5 Доступ к системе

Система представляет собой онлайн-сервис, доступный через Интернет-браузер, с авторизацией пользователей по логину и паролю или через сервисы популярных социальных сетей.

Установка дополнительного клиентского программного обеспечения для работы с Системой не требуется.

Вход в систему «Крибрум. Сигнал» осуществляется по ссылке <https://forest.kribrum.ru>

Для входа в систему требуется авторизация через популярные сервисы социальных сетей или электронную почту (рис.1):

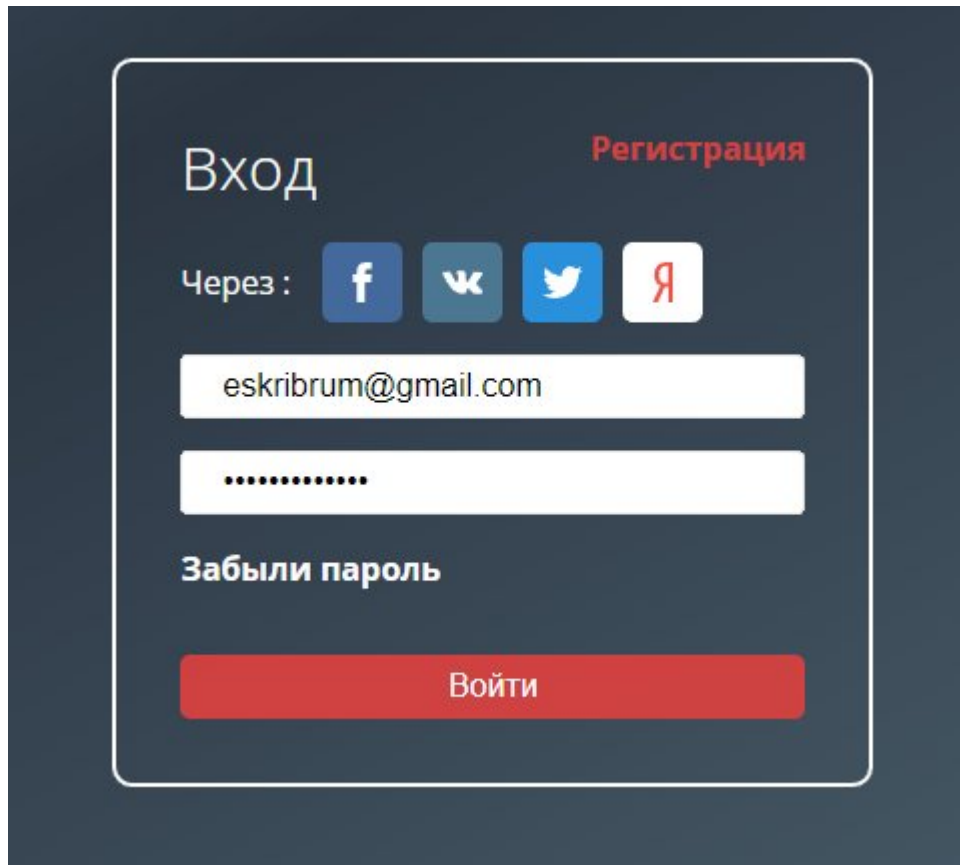


Рисунок 1. Страница авторизации в системе «Крибрум. Сигнал»

1.5.1 Структура интерфейса системы

Пользовательский графический интерфейс Системы «Крибрум. Сигнал» представляет собой интуитивно понятный инструмент анализа источников и путей распространения информации в социальных сетях.

После авторизации в системе показывается главная страница. Она содержит следующие элементы:

- Логотип;
- Область создания и настройки параметров графа;
- Область **Графы**;
- Футер (“подвал” сайта).

Название графа

Искать публикации

Социальная сеть
 Twitter ▼

Добавить в заставку ссылки

Игнорировать публикации ранее

Использовать не более найденных публикаций

Игнорировать деревья менее чем из вершин

Графы

Владелец	Статус	Название	Поисковый запрос	Подробности	Статистика
samosvat	посмотреть	Колумбайн в Керчи		i	i
samosvat	посмотреть	Valentino	Акула из ИКЕА	i	i
samosvat	посмотреть			i	i
samosvat	посмотреть	Сухих Г.Т.		i	i
samosvat	посмотреть	колумбайн-трек	"я устроил Колумбайн"	i	i
samosvat	посмотреть	колумбайн - сентябрь18	"Колумбайн"	i	i
samosvat	посмотреть	колумбайн_20.04.18	"20 апреля в школах России планируются нападения с убийствами и расправами"	i	i
samosvat	broken	Концертный зал в Зарядье	Концертный зал в Зарядье	i	i
samosvat	посмотреть	Секс в Зарядье	Секс в Зарядье	i	i
samosvat	посмотреть	А.Н.	Дожить до пенсии2	i	i
samosvat	broken	Навальный	Дожить до пенсии	i	i
samosvat	посмотреть	Пенсионный возраст		i	i
samosvat	посмотреть	колумбайн на марше		i	i
samosvat	посмотреть	Граф3	Митинги 26 марта	i	i
samosvat	посмотреть	Графа2		i	i
samosvat	посмотреть			i	i
samosvat	посмотреть			i	i
samosvat	посмотреть	Граф	Митинги 26 марта по всей России	i	i

Рисунок 2. Интерфейс системы «Крибрум. Сигнал»

1.5.2 Область создания и настройки параметров графа

Область создания и настройки параметров графа содержит поля, необходимые для создания графа, описания содержания распространяемой информации в виде поискового запроса, состоящего из ключевых слов и/или ссылок на публикации, настройки параметров создаваемого графа (рис.3).

Название графа

Искать публикации

Социальная сеть
 Twitter ▼

Добавить в заставку ссылки

Игнорировать публикации ранее

Использовать не более найденных публикаций

Игнорировать деревья менее чем из вершин

Рисунок 3. Структура области создания и настройки

1.5.3 Область «Графы»

В области библиотеки созданных графов отображается перечень сформированных графов в виде таблицы, содержащей следующие поля:

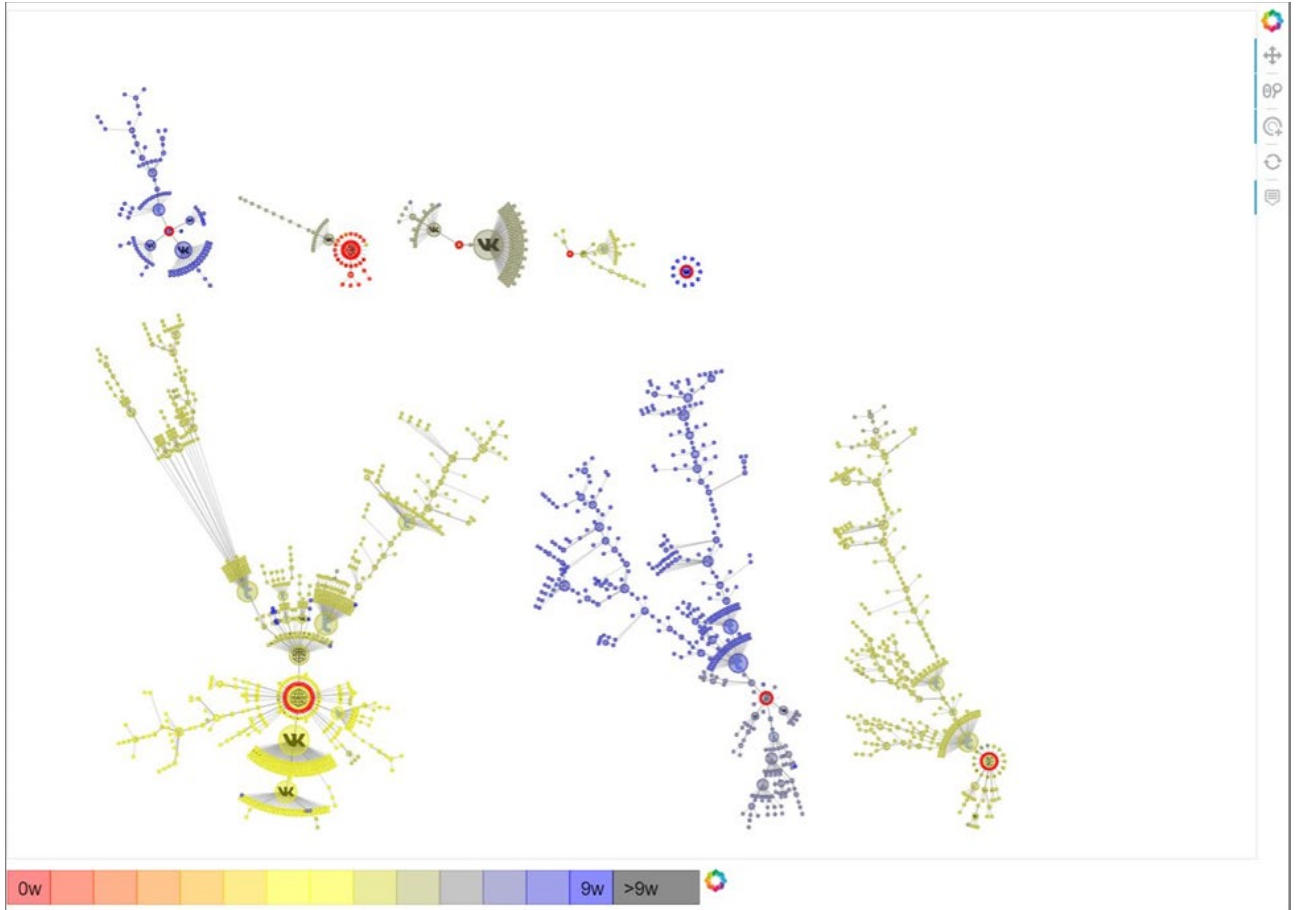
- **Владелец** – создатель графа;
- **Статус** – статус графа изменяется в процессе создания и формирования графа;
- **Название** – название графа;
- **Поисковый запрос**;
- **Подробности** – переход на страницу **Подробности**;
- **Статистика** – переход на страницу **Статистика**).

Графы					
Владелец	Статус	Название	Поисковый запрос	Подробности	Статистика
samosvat	посмотреть	Колумбайн в Керчи		↓	↓
samosvat	посмотреть	Valentino	Акула из ИКЕА	↓	↓
samosvat	посмотреть			↓	↓
samosvat	посмотреть	Сухих Г.Т.		↓	↓
samosvat	посмотреть	колумбайн-трек	"я устроил Колумбайн"	↓	↓
samosvat	посмотреть	колумбайн - сентябрь18	"Колумбайн"	↓	↓
samosvat	посмотреть	колумбайн_20.04.18	"20 апреля в школах России планируются нападения с убийствами и расправами"	↓	↓
samosvat	broken	Концертный зал в Зарядье	Концертный зал в Зарядье	↓	↓
samosvat	посмотреть	Секс в Зарядье	Секс в Зарядье	↓	↓
samosvat	посмотреть	А.Н.	Дожить до пенсии2	↓	↓
samosvat	broken	Навальный	Дожить до пенсии	↓	↓
samosvat	посмотреть	Пенсионный возраст		↓	↓
samosvat	посмотреть	колумбайн на марше		↓	↓
samosvat	посмотреть	Граф3	Митинги 26 марта	↓	↓
samosvat	посмотреть	Графа2		↓	↓
samosvat	посмотреть			↓	↓
samosvat	посмотреть	Граф	Митинги 26 марта по всей России	↓	↓

Рисунок 4. Структура области «Графы»

1.5.4 Страница построения и просмотра сформированных графов

Страница построения и просмотра сформированных графов содержит область отображения графов, сформированных по исследуемой теме, и элементы настройки области отображения (рис.5).



2 ФУНКЦИОНАЛЬНОЕ ОПИСАНИЕ РАЗДЕЛОВ СИСТЕМЫ


2.1 Главная страница

2.1.1 Область создания и настройки параметров графа

Для создания графа пользователь системы «Крибрум. Сигнал» заполняет поля области создания и настройки параметров графа.

В поле «Название графа» вносится произвольное название.

Предмет или тема исследования источников и путей распространения информации задаются с помощью поисковых параметров.

В поле **Искать публикации** вносятся ключевые слова, относящиеся к исследуемой теме. В данном разделе необходимо писать поисковый запрос на основе синтаксиса языка системы полнотекстового поиска Sphinx, подробнее с синтаксисом [можно ознакомиться](https://www.oreilly.com/library/view/introduction-to-search/9780596809546/ch04.html) в <https://www.oreilly.com/library/view/introduction-to-search/9780596809546/ch04.html>. Правильность составленного запроса можно проверить в сервисе <http://bs.kribrum.ru/> - в поле текст выбрав значок  и выбрав графу Sphinx.

В поле **Социальная сеть** можно выбрать одну из платформ социальных сетей: **Facebook**, **ВКонтакте**, **Twitter**, **Instagram**, или выбрать **Все**.

В поле **Добавить в заставку ссылки** необходимо внести все известные ссылки, имеющие отношение к исследуемой теме. Ссылки необходимо вставлять списком, без использования знаков препинания, используя только символ переноса строки.



Название графа

Искать публикации

Социальная сеть
Twitter

Добавить в заставку ссылки

Игнорировать публикации ранее 11/08/2018

Использовать не более 3000 найденных публикаций

Игнорировать деревья менее чем из 10 вершин

построить граф

Рисунок 6. Область создания и настройки параметров графа

Далее необходимо заполнить поля настройки.

В поле **Игнорировать публикации ранее** вносится дата, до которой данные не надо учитывать.

В поле **Использовать не более ... найденных публикаций** вносится число публикаций, на основе которых будет строиться граф. Максимальная глубина построения графа составляет 1 месяц.

В поле **Игнорировать деревья менее чем из ... вершин** вносится число вершин дерева графа, что позволяет исключить из рассмотрения деревья с малым, не значительным характером распространения информации.

После нажатия кнопки **Построить граф** система «Крибрум. Сигнал» формирует граф на основе заданных параметров.

Сформированные графы отображаются в области **Графы**.

2.1.2 Область «Графы»

В области библиотеки созданных графов отображается перечень сформированных графов в виде таблицы, содержащей следующие поля:

- **Владелец** – создатель графа;
- **Статус** – переход на страницу просмотра графа;
- **Название** – название графа;
- **Поисковый запрос** – поисковый запрос для построения графа;
- **Подробности** – переход на страницу **Подробности**;
- **Статистика** – переход на страницу **Статистика**.

Графы					
Владелец	Статус	Название	Поисковый запрос	Подробности	Статистика
samosvat	посмотреть	Колумбайн в Керчи		↓	↓
samosvat	посмотреть	Valentino	Акула из ИКЕА	↓	↓
samosvat	посмотреть			↓	↓
samosvat	посмотреть	Сухих Г.Т.		↓	↓
samosvat	посмотреть	колумбайн-трек	"я устроил Колумбайн"	↓	↓
samosvat	посмотреть	колумбайн - сентябрь18	"Колумбайн"	↓	↓
samosvat	посмотреть	колумбайн_20.04.18	"20 апреля в школах России планируются нападения с убийствами и расправами"	↓	↓
samosvat	broken	Концертный зал в Зарядье	Концертный зал в Зарядье	↓	↓
samosvat	посмотреть	Секс в Зарядье	Секс в Зарядье	↓	↓
samosvat	посмотреть	А.Н.	Дожить до пенсии2	↓	↓
samosvat	broken	Навальный	Дожить до пенсии	↓	↓
samosvat	посмотреть	Пенсионный возраст		↓	↓
samosvat	посмотреть	колумбайн на марше		↓	↓
samosvat	посмотреть	Граф3	Митинги 26 марта	↓	↓
samosvat	посмотреть	Графа2		↓	↓
samosvat	посмотреть			↓	↓
samosvat	посмотреть			↓	↓
samosvat	посмотреть	Граф	Митинги 26 марта по всей России	↓	↓

Рисунок 7. Область библиотеки созданных графов

Статус графа изменяется в процессе формирования графа.

Статусы графа:

- **формируется** (в процессе создания);
- **посмотреть** (создан и готов к просмотру);
- **broken** (поврежден).

При нажатии кнопки **Посмотреть** происходит переход на страницу построения и просмотра сформированных графов.

2.2 Страница построения и просмотра сформированных графов

Страница построения и просмотра сформированных графов содержит область отображения графов, сформированных по исследуемой теме, и элементы настройки области отображения (рис. 8).

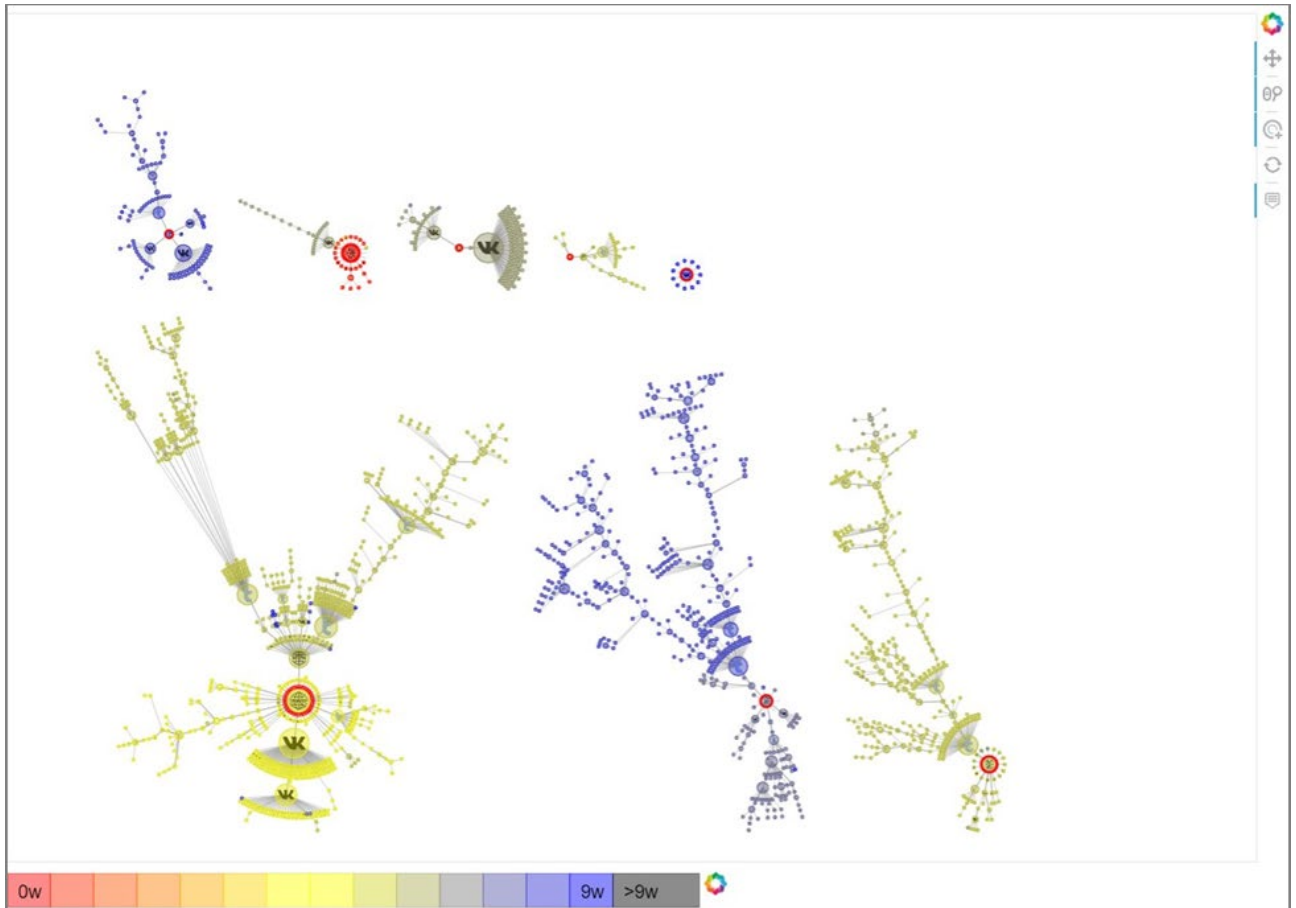


Рисунок 8. Структура страницы построения и просмотра сформированных графов

Под областью отображения графов находится «температурная шкала», характеризующая скорость распространения информации – чем краснее цвет узлов и связей графа, тем быстрее по этим узлам распространялась информация. Чем темнее цвет узлов и связей графа, тем медленнее информация распространялась по таким узлам.

2.3 Страница «Подробности»

Страница **Подробности** содержит детальную информацию о созданном графе, информацию о параметрах настройки графа и полный перечень ссылок, использованных для построения графа.

Страница **Подробности** содержит следующие поля:

- **Владелец;**
- **Название;**
- **Поисковый запрос;**
- **Граф: посмотреть** (ссылка на граф);
- **Прочие параметры** (настраиваемые при создании графа).

При нажатии кнопки **Граф: посмотреть** происходит переход на страницу построения и просмотра сформированных графов (рис 9).

О графе

Владелец: samosvat
Название: Граф3
Поисковый запрос: Митинги 26 марта
Граф: [посмотреть](#)
Прочие параметры: {'platform': '', 'ignore_before': '2017-03-01T00:00:00', 'ignore_small_nets': 10, 'query': 'Митинги 26 марта', 'query_limit': 3000}

Затравка

<https://twitter.com/Aleks13041970/status/1042454631161491456>
https://vk.com/wall42003331_13897
https://twitter.com/Arctic_bear6/status/1050977129461092352
https://vk.com/wall182136258_4628
https://vk.com/wall2914148_1779
https://vk.com/wall420877639_37348
https://vk.com/wall238389760_4126
https://vk.com/wall11175648_38317
https://vk.com/wall211024766_38651
https://vk.com/wall161551944_74387
https://vk.com/wall-47385736_77877
<https://twitter.com/KIB2569/status/1044591870993735685>
<https://fedorka56.livejournal.com/678586.html>
<https://twitter.com/AlenaSvetoch/status/1043747586132377600>
<https://twitter.com/antrijowa/status/1044550737676709889>
https://vk.com/wall144980582_63265
<https://twitter.com/ladybenia/status/1043692428765605888>
https://vk.com/wall-165774844_1419
https://vk.com/wall-59990444_119315
https://vk.com/wall271900382_629
https://vk.com/wall-2007972_49362
https://www.youtube.com/watch?v=_xgK4Mc5FSw
<https://ovdinfo.org/express-news/2018/09/20/sud-v-igove-postanovil-osvobodit-dmitriya-krepkina-osuzhdenного-po-delu-26>
https://vk.com/wall165472878_9288
<https://twitter.com/soskotter/status/1043749853719748608>
https://vk.com/wall32673660_10065
<https://twitter.com/Vrklvch/status/1041022261854130177>
<https://twitter.com/leontevaleksan3/status/1043713408389062656>
<https://twitter.com/nalog399/status/1041151661593571329>
<https://twitter.com/Rari17906174/status/1050729934329667585>
<https://viking-nord.livejournal.com/23289873.html>
<https://solvaigsamara.livejournal.com/5112409.html>
<https://vk.com/@mbkhmedia-spyaschii-advokat-i-shizo-iz-za-podderzhki-sobchak-figurant>
<http://nymphetovhnsn3gpf.onion/viewtopic.php?f=17&t=501&start=2720#p172692>
<https://www.facebook.com/876760289127004/posts/1339116982891330>
https://vk.com/wall163512067_19344
<https://twitter.com/Vichka77ss/status/1043710434099048448>

Рисунок 9. Структура страницы «Подробности»

2.4 Страница «Статистика»

Страница **Статистика** содержит статистическую информацию созданного графа, включающую в себя:

- ссылку на источник исходной публикации;
- время исходной публикации;
- количество публикаций по заданной теме (сумма постов, комментариев, репостов и внешних ссылок);
- количество постов по теме;
- количество комментариев;
- количество репостов публикаций;
- количество внешних ссылок.

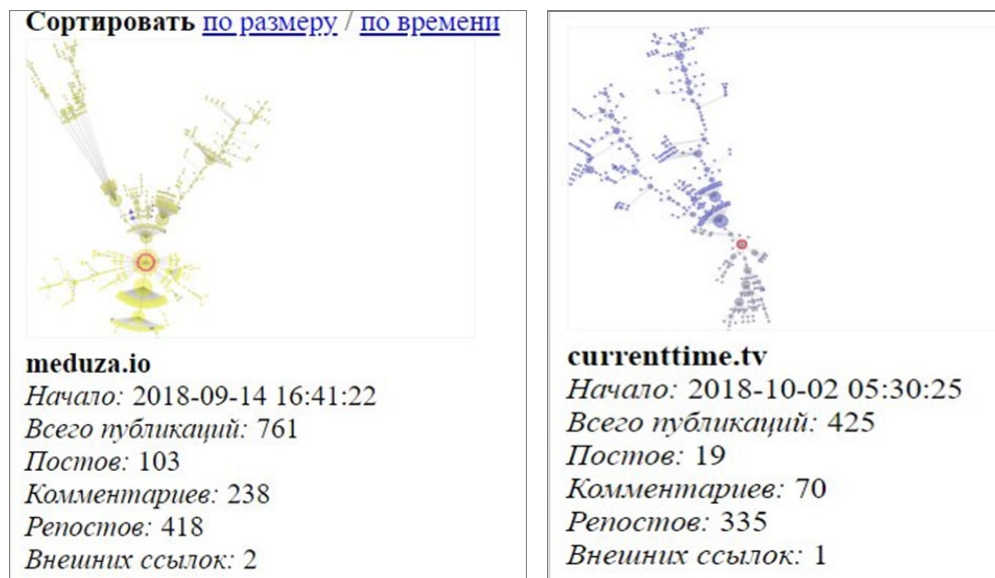


Рисунок 10. Структура страницы «Статистика»

Страница **Статистика** содержит фильтр сортировки сформированных графов по времени исходных публикаций или по размерам графа.

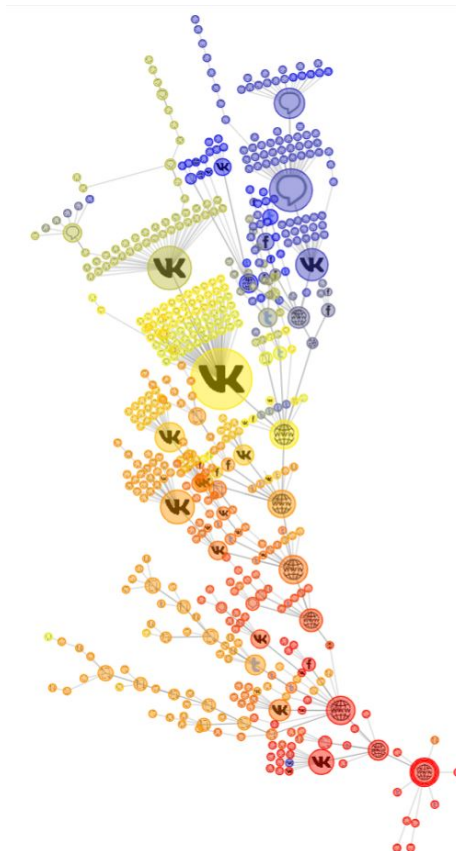
2.5 Примеры и описания графов

Граф №1



Распространение информации в социальных медиа носит **естественный характер**: цепочки распространения (направления обсуждения) визуально схожи и не имеют значимых отклонений, размеры и положение ключевых узлов (медийный вес публикаций) закономерны, а «температура» узлов (скорость распространения) равномерна.

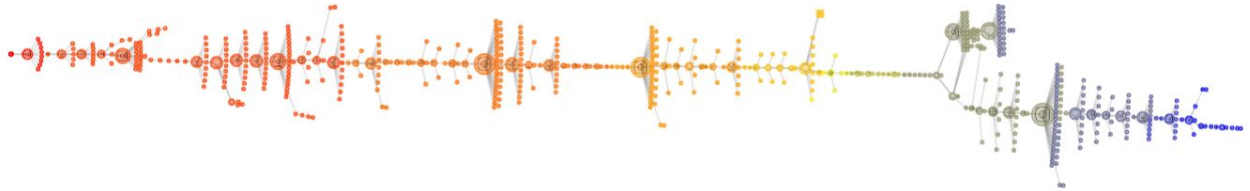
Граф №2



Длинные последовательные цепочки распространения информации являются одним из главных признаков того, что организовано искусственное распространение. Разная «температура» узлов и положение ключевых узлов

говорят о попытках актуализации темы, спустя значительное время после публикации информации первоисточником.

Граф №3



Признаком искусственного характера распространения информации также является аномальный интерес автора (аккаунта) к исследуемой теме. В данном случае автор в течение месяца регулярно публикует сообщения по теме, несмотря на отсутствие значимого интереса к ней у его аудитории, с целью поддержания ее присутствия в медиапространстве.