

## **Обзор сервисов сбора данных о пользователях социальных сетях**

*Ковалева Ирина Валерьевна*

*Приамурский университет имени Шолом-Алейхема*

*студент*

*Размахнина Анна Николаевна*

*Приамурский университет имени Шолом-Алейхема*

*студент*

*Лучанинов Дмитрий Васильевич*

*Приамурский университет имени Шолом-Алейхема*

*Старший преподаватель кафедры информационных систем, математики и методик обучения*

### **Аннотация**

В статье описываются подходы, инструментов и методы обработки больших данных (Big Data) в социальных сетях. Рассмотрены алгоритмы сбора данных в социальной сети facebook, instagram и VK.

**Ключевые слова:** алгоритм сбора данных, большие данные, big data, социальные сети.

## **Overview of services for collecting data on users of social networks**

*Kovaleva Irina*

*Sholom-Aleichem Priamursky State University*

*student*

*Razmahnina Anna Nikolaevna*

*Sholom-Aleichem Priamursky State University*

*student*

*Luchaninov Dmitry Vasilievich*

*Sholom-Aleichem Priamursky State University*

*Senior lecturer in information systems, mathematics and teaching methods*

### **Abstract**

The article describes the approaches, tools and methods of processing big data (Big Data) on social networks. The algorithms of data collection in social network facebook, instagram и VK.

**Key words:** algorithm of data acquisition, big data, big data, social networks.

Социальные сети изначально создавались с целью возможности общения друзей, коллег, родственников между собой, но теперь все зависит от статуса человека в обществе в частности его участия в сфере бизнеса. Нынешнее молодое поколение и те, кто не имеет никакого отношения к сфере бизнеса, вероятно, не предполагают степень отслеживания и контроля их действия в социальных сетях для коммерческой выгоды.

Основная фундаментальная концепция, которую необходимо понимать в отношении социальных сетей, это то, что они являются одним из наиболее значимых источников больших данных (Big Data).

Основная информация в социальных сетях, на которую обращают внимание коммерческие организации для анализа клиентской аудитории – это лайки, комментарии, загрузки, ретвиты и акции. Все эта информация может использоваться для толкования того, чего хотят клиенты, и того, как организации должны показывать себя потребителям, если они хотят оставаться конкурентоспособными.

Прежде всего, стоит отметить, что под терминами «Big Data» подразумевается огромный набор информации. При этом ее объем настолько велик, что анализ таких больших объемов данных обычными программными средствами является очень сложной. Big Data – это метод хранения и обработки огромных объемов данных. Сам термин «Big Data» состоит из определения трех «V», что означает Volume – объем данных, Velocity – необходимость перерабатывать данные на большой скорости и Variety – многообразие и недостаточную структурированность данных.

Другая сторона вопроса – это многообразие и неорганизованность информации. Все чаще и чаще приходится взаимодействовать с медиа контентом, записями в блогах, неструктурированными документами и т.д.

Таким образом, когда мы говорим о big data, мы понимаем, что это связано с тремя аспектами: большим объемом информации, ее разнообразием или необходимостью обрабатывать данные очень быстро. С другой стороны, под этим термином часто понимают совершенно конкретный набор подходов и технологий, призванных решить данные задачи. В основе одного из таких подходов лежит система распределенных вычислений, где обработка больших объемов данных требует для себя не одну высокопроизводительную машину, а целую группу таких машин, объединенных в кластер.

Многие авторы занимались исследованием алгоритмов и обработкой больших объемов данных. Например, И.А. Сизов раскрывает важность изучения, и рассматриваются различные возможности использования технологий Big Data [1]. В.А. Бабурин и М.Е. Яненко показали возможность применения технологий Big Data для повышения конкурентоспособности предприятий [2]. В.Л. Шпер в своей работе описывает методологию Big Data — большие данные [3]. Н.В.Ситникова и др. проводили исследования алгоритмов классификации с применением методов обработки Big Data [4]. А.А. Кабдуллин исследовал преобразование больших данных в аналитической системе Oracle Big Data Appliance [5].

Поэтому, когда мы говорим о Big Data по отношению к социальным сетям, мы должны сначала понять, что они не делимы (или, по крайней мере, внутренне взаимозависимы друг от друга).

Алгоритм работы с данными социальных сетей определяется следующим алгоритмом:

- Сбор данных;
- Обработка данных;
- Анализ данных.

На сегодняшний день выделяют ряд инструментальных средств для сбора, анализа и обработки данных в социальных сетях.

Изучение социальных сетей превращается в задачу обработки Больших Данных, когда бизнес-руководителям или специалистам по информационным системам нужно прогнозировать поведение участников сообщества, чтобы добиться повышения эффективности маркетинга или продаж. У многих социальных сайтов имеется от 10 до 200 млн пользователей, поэтому стержнем большинства исследований является работа с выборками данных. Оптимальным, хотя и сильно затратным по времени, было бы извлечение знания из всего среза данных.

Для обработки больших объемов данных и получения информации о пользователях используются онлайн сервисы подробного анализа личных страниц и сообществ в социальных сетях.

**Fanpage Karma** – это аналитический сервис, предоставляющий подробный анализ Facebook-страницы, а также сравнение ее со страницами конкурентов.

Для анализа используются следующие параметры:

- Количество подписчиков и их рост в сравнении с предыдущим месяцем;
- Люди, которые обсуждают контент данной страницы;
- Стоимость постов – рассчитывается, сколько бы стоил охват подписчиков за месяц, если бы давалась реклама (берется средний CPM по отрасли);
- Среднее количество постов в день;
- Показатель ответов на публикации подписчиков;
- Время ответа – как быстро бренд отвечает на публикации подписчиков;
- Показатель взаимодействий – среднее число комментариев, лайков и просмотров по отношению числу подписчиков;
- Карма – взвешенный показатель вовлеченности, где репосты играют большую роль, чем лайки;
- Продуктивность страницы – высчитывается из 100%, за основу берутся показатели роста подписчиков и вовлеченности;

**Iconosquare** – бесплатная платформа, позволяющая просматривать статистику по собственному профилю в Instagram.

Основными функциями являются:

1. Viewer - просмотр ленты, подписчиков, лайков и комментариев;
2. Statistics (статистика) – раздел, где производится расчет подробной статистики фотографий с наибольшим количеством лайков, новых подписчиков и возможность продвигать свою страницу. Данный раздел будет полезен для владельцев корпоративных аккаунтов;
3. Manage (управление) – данный раздел, в первой вкладке Comment tracker собраны 5 последних комментариев к вашим фото. Здесь вы также сможете ответить на эти комментарии. Во второй вкладке Private message можно вести личную переписку с пользователями.
4. Promote (продвижение) - Здесь найдутся очень полезные вещи для вашей рекламы и продвижения: можно скопировать ссылку, чтобы дать ее тем, кто не использует Инстаграм, для просмотра вашего профиля с компьютера. Можно создать виджеты и кнопки для сторонних сайтов, а также коллаж из фотографий профиля.

**Allsocial** – это сервис для просмотра статистики сообществ социальной сети VK, который удобен и прост в использовании, к тому же не требует регистрации. Данный сервис имеет встроенную аналитику и собственное ранжирование и отлично подходит для проведения рекламной компании.

Основными особенностями сервиса являются:

- Сортировка исследуемых сообществ и групп по различным тематикам;
- Статистика сообществ по параметрам: охват, посетители, прирост, подписчики;
- Фиксирование изменений в названиях групп и сообществ;
- Поиск популярных страниц Вконтакте.

Данный сервис был выбран для исследования статистики сообществ. AllSocial – это бесплатный сервис, но он работает только с социальной сетью VK. С помощью данного сервиса можно получить абсолютно любые данные по различным сообществам, включая их админов. Регистрация и оплата для работы не нужна, достаточно просто открыть страничку сервиса и можно начинать анализ.

На рисунке 1 показан пример работы сервиса, который по заданным параметрам выводит результат. Например было задано название сообщества «Киномания» и определены критерии сортировки: аудитория, период, тип, категории и др. Так сортировка по подписчикам дает возможность увидеть популярные сообщества, оценить прирост за день, за неделю или за месяц.

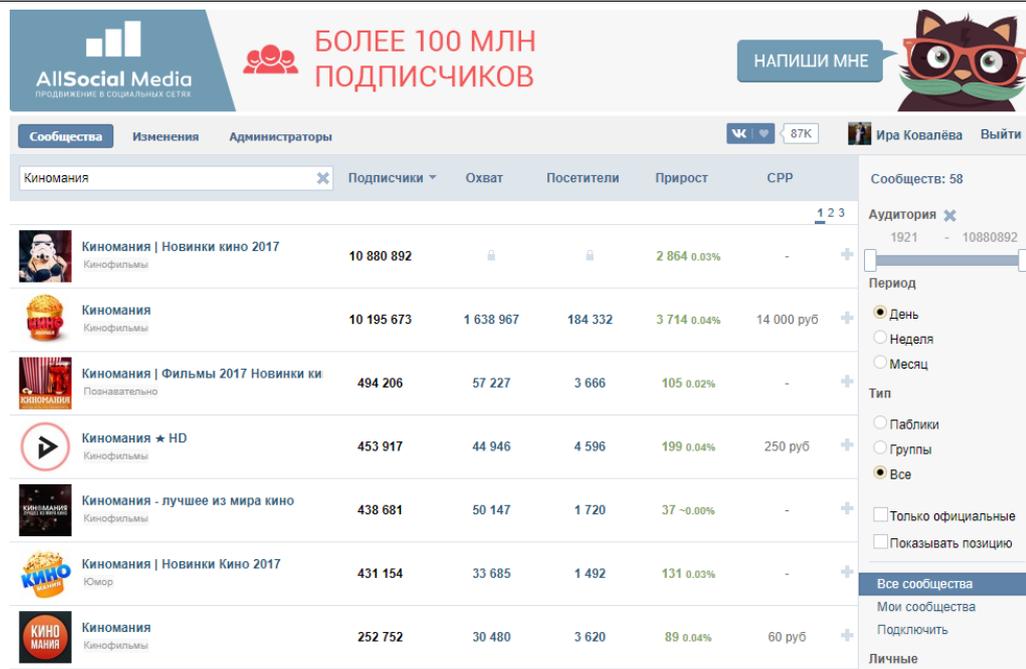


Рисунок 1. Пример пользования Allsocial

Вы не потратите время и деньги впустую, размещая рекламные объявления на неэффективных площадках. Правильный выбор сообщества – залог успеха вашей рекламной кампании.

Качественный и быстрый анализ сообществ ВКонтакте с помощью сервиса Allsocial.ru поможет каждому пользователю всегда знать о самых лучших и популярных пабликах. Если нужно разметить рекламу – тогда это возможность найти подходящую платформу и сделать рекламную кампанию более эффективной. Если же вы являетесь администратором сообщества, тогда можете увидеть свой рейтинг среди других пабликов, а также место ваших конкурентов.

Извлечение «секретной» информации из социальных сетей, посредством специальных сервисов, получает все большую популярность. Особенно это актуально, если речь идет о коммерческих страницах и высоко конкурентных областях, где популярность и доход влияют на продвижение ресурса. Так как зная статистику конкурентов, можно ставить и реализовывать конкретные цели с большим успехом.

В данном исследовании были описаны сервисы для сбора, анализа и обработки информации в социальных сетях, таких как Facebook, Instagram и VK. На примере рассмотрено применение сервиса Allsocial для социальной сети VK.

### Библиографический список

1. Сизов И.А. Big Data - большие данные в бизнесе // Экономика. Бизнес. Информатика. 2016. №3. С. 8-23.
2. Бабурин В.А., Яненко М.Е. Технологии Big Data в сервисе: новые рынки, возможности и проблемы // Техничко-технологические проблемы сервиса.

- 
2014. №1(27). С. 100-105.
3. Шпер В.Л. Big Data — большие данные // Методы менеджмента качества. 2013. №6. С. 56-57.
  4. Ситникова Н.В., Парингер Р.А., Куприянов А.В. Исследование алгоритмов классификации с применением методов обработки Big Data // XIII Королёвские чтения. Международная молодёжная научная конференция, сборник трудов. 2015. С. 137-138.
  5. Кабдуллин А.А. Система Oracle Big Data для анализа больших данных // WORLD SCIENCE: PROBLEMS AND INNOVATIONS. Сборник статей победителей IX Международной научно-практической конференции: в 2 частях. 2017. С. 171-174.