

Использование социальных сетей для оценки уровня обеспокоенности населения вопросами безопасности жизнедеятельности

ISSN 1996-8493

© Технологии гражданской безопасности, 2015

М.Э. Галиуллин

Аннотация

Рассмотрены вопросы использования общедоступной информации из социальных сетей, ее тематический и географический анализ в реальном времени, а также различные статистические оценки с целью определения динамики напряженности и заинтересованности населения отдельных территорий вопросами безопасности жизнедеятельности. Рассмотрены варианты технической реализации предложенного подхода, его актуальность, возможности применения, а также влияние использования подобных инструментов на уровень социальной напряженности общества и подготовку населения к действиям в условиях ЧС.

Ключевые слова: социальные сети; безопасность жизнедеятельности; социальная напряженность; потенциально опасные территории; информационные технологии; java enterprise edition; анализ текста; big data.

Using Social Networks to Assess the Level of Public Concern Issues of Life Safety

ISSN 1996-8493

© Civil Security Technology, 2015

M. Galiullin

Abstract

The article discusses the use of publicly available information from social networks, its thematic and geographical analysis in real time, as well as various statistical evaluation in order to determine the dynamics of tension and interest of the population of certain areas of life safety issues. The variants of the technical implementation of the proposed approach, its relevance, applicability, and the impact of the use of such tools on the level of social tension and prepare the population to act in emergency situations.

Key words: social networks; life safety; social tension; potentially dangerous territory; information technology; java enterprise edition; text analysis; big data.

Введение

За последние несколько лет область присутствия пользователей в сети Интернет сместилась в сторону социальных сетей и микроблогов. Новости в социальных сетях и микроблогах появляются раньше, чем на телевидении, радио и в газетах [1]. Здесь же размещаются комментарии, ведутся дискуссии. Таким образом, формируется общественное мнение по тому или иному вопросу [2]. Пользователи теперь имеют возможность самостоятельно оповестить других людей о каком-либо важном происходящем событии, дать оценку этому событию, и узнать мнение по этому событию от других пользователей сети. Причем общность людей, общающихся в сети интернет, все расширяется, общность охватывает все новые слои населения. В этой общности обсуждаются те вопросы, которые волнуют и интересуют людей.

Безусловно, вопросы безопасности жизнедеятельности всегда были и будут в центре внимания людей. Эти вопросы были и будут предметом обсуждения в блогосфере (социальных сетях, микроблогах, форумах) [4]. Но в современном мире вопросы безопасности — это сложные вопросы, которые не всегда имеют простое объяснение. Зачастую просто необходимо, чтобы люди слышали квалифицированное мнение специалистов по волнующему их вопросу.

Задача государственных структур — быть в курсе проблем и вопросов, волнующих население, и вовремя давать оценку тем или иным происходящим событиям или предметам дискуссий, возникающих в обществе.

Поэтому для таких структур, ответственных за связь с населением, крайне важно использование интернет технологий доступа к социальным сетям. Это позволит знать настроение и уровень обеспокоенности жителей существующими угрозами в регионе, что, в свою очередь, даст возможность избежать случайной паники, а в случае реальной угрозы оперативно привлечь к проблеме специалистов и донести их мнение до населения. Вовремя полученные знания позволят грамотно справиться с реально случившейся чрезвычайной ситуацией [3, 4].

Сориентироваться в море информации, обсуждающейся в современной блогосфере, достаточно сложно. Необходим специальный интернет инструментарий, с помощью которого работники государственных структур могли бы в реальном времени количественно отслеживать уровень обеспокоенности населения проблемами в области безопасности жизнедеятельности, чтобы во время привлекать специалистов к обсуждению возникающих проблем и доводить их мнение до пользователей социальных сетей. Результатом такой деятельности могло бы стать снижение уровня социального напряжения в обществе и изменение качества подготовки населения к действию в условиях ЧС [5].

Созданию такого инструмента, возникающим при создании подзадачам и путям их решения посвящена данная статья.

Постановка задачи

Формально постановка задачи звучит следующим образом. Необходимо создать информационную систему, выполняющую следующие задачи:

- сбор в реальном времени публичных сообщений пользователей социальных сетей, содержащих определенные ключевые слова;

- оценка релевантности собранной информации искомой тематике;

- определение дополнительных атрибутов выбранных сообщений, таких как, например, регион проживания автора сообщения;

- статистический анализ имеющейся информации, построение временных, территориальных и тематических срезов и т.п.;

- предоставление инструментов для специалистов, позволяющих в наглядном виде представлять текущие данные и всевозможные результаты анализа.

Технически система может состоять из трех частей.

Первая часть — серверное ядро системы — должно постоянно, в автоматическом режиме, собирать сообщения пользователей из наиболее популярных социальных сетей, анализировать полученные данные на предмет релевантности отслеживаемой тематике, сохранять отобранные данные в базе данных, выполнять фоновый статистический анализ имеющейся информации.

Вторая часть — набор инструментов для специалистов. Инструменты должны делать количественные и сравнительные оценки различного рода на основе отобранных ядром системы данных. Программный код инструментов выполняется также на сервере, но только по запросу пользователя.

Третья часть — пользовательский интерфейс доступа к инструментам — клиентская часть системы, доступная специалистам, использующим инструменты. Интерфейс должен быть удобным и интуитивно понятным для неподготовленных пользователей.

Подзадачи и требования к реализации системы

Количество ежеминутно оставляемых пользователями сообщений в социальных сетях огромно. Все популярные социальные сети имеют программный интерфейс доступа к своим данным (API), используя который можно получить список пользовательских постов (сообщений), содержащих определенные ключевые слова, за некоторый промежуток времени. С помощью предоставляемого API можно получить также некоторую второстепенную информацию (например, географические регионы проживания авторов сообщения), которая может помочь в группировке и анализе полученных данных. Количество полученных таким образом сообщений может оказаться по-прежнему довольно внушительным, и большинство из этих сообщений может не отвечать искомой тематике. Поэтому одной из важных подзадач является определение соответствия каждого полученного сообщения контексту поиска.

Зачастую, пользовательские сообщения, разбросанные по разным микроблогам, все же связаны между собой логически и представляют цепочки тематических обсуждений. Выявление таких цепочек, определение их смысловой темы и объекта обсуждения является также одной из серьезных подзадач.

Описанные выше задачи подразумевают использование алгоритмов определения частотности появления слов в тексте, морфологического, синтаксического и семантического анализа текста и т.п.

Подсистема фоновой анализа полученных данных и соответствующий набор инструментов должны реализовывать следующие возможности:

- выявлять новые темы и объекты обсуждений;

- выявлять всплески уровня активности пользователей социальных сетей в обсуждении различных тем, касающихся вопросов безопасности жизнедеятельности;

- находить тему, объект, географических регион и источник обсуждения, вызвавшего всплеск обеспокоенности;

- выполнять поиск в базе данных цепочек сообщений и находить обсуждение определенных интересующих специалистов тем по ключевым словам за определенный промежуток времени;

- оценивать общий уровень активности населения отдельных географических регионов за различные промежутки времени, и позволять наблюдать динамику уровня;

- позволять включаться специалистам в обсуждения, возникающие в социальных сетях;

Часть подсистемы анализа данных должна работать в момент сбора информации в реальном времени и производить первичный отбор данных. Другая часть, реализующая сложные и продолжительные по времени алгоритмы, должна обрабатывать полученные данные в фоновом режиме, но с определенной периодичностью. Наконец, программный код инструментов, связанных с пользовательским интерфейсом, должен включать только быстрые алгоритмы, срабатывающие в момент запроса пользователя.

Реализация пользовательского интерфейса может быть любой, и зависит от целевой аудитории специалистов-пользователей и их количества. Интерфейс может быть реализован в виде настольных, мобильных и веб-приложений. Наиболее универсальным решением может стать веб-интерфейс. Независимо от реализации, он должен быть понятным любому пользователю и включать интуитивно понятные блоки — графики, диаграммы, карты, сортированные списки, инструменты поиска и т.п.

Реализация системы и достигнутые результаты

Работы по созданию описанной системы были проведены в рамках совместного проекта с Общественным Советом корпорации «Росатом» при поддержке грантом Минобрнауки. В этом проекте по созданию «Информационной инфраструктуры для

работы с населением на территориях с радиационно опасными объектами» были реализованы многие из описанных выше алгоритмов и инструментов.

Для реализации системы в рамках проекта автором данной статьи было предложено использовать программную платформу Java Enterprise Edition (далее JEE), которая предназначена для создания систем уровня предприятия. JEE уже содержит в себе многие низкоуровневые сервисы и инфраструктуру, необходимую для создания любого крупного программного проекта — хранение объектов в базе данных, сетевое взаимодействие, обмен сообщениями, выполнение подзадач в отдельных потоках, веб-сервер и запуск сервлетов, и многое другое. Используя JEE-платформу (многие реализации которой абсолютно бесплатны, например WildFly), разработчики могут целиком сконцентрироваться на самой задаче.

В ходе работы над проектом были созданы модули, написанные на языке Java, обеспечивающие доступ к открытой информации через предоставляемый социальными сетями API, частотный и морфологический анализ текстов, и многие другие необходимые функции.

Созданная система работает с наиболее популярными социальными сетями — Вконтакте, Facebook, Twitter. Система также производит поиск информации в официальных новостных лентах. Искомая тематика отслеживаемых сообщений — радиационная безопасность на объектах атомной промышленности в России. Система фильтрует найденные сообщения пользователей в соответствии с алгоритмами релевантности заданным ключевым словам, группирует их по территориальному признаку и по признаку принадлежности обсуждения конкретному объекту атомной промышленности. Созданные инструменты для специалистов, доступные в виде веб-интерфейса, позволяют выявлять всплески и динамику активности пользователей как в целом по России, так и по отдельным регионам. Для лучшей наглядности многие инструменты сделаны в виде информеров (рис. 1—4).

Работы в рамках проекта были нацелены на развертывание созданной системы в Общественных Приемных корпорации «Росатом», существующих в радиационно опасных регионах. Имея под рукой инструменты, предлагаемые данной системой, специалисты этих структур могут получить необходимую обратную связь с населением своего региона, и наладить с ним непосредственный диалог. Это, в свою очередь, может способствовать уменьшению уровня обеспокоенности жителей регионов вопросами радиационной безопасности, и качественно изменить подготовку населения к действиям в условиях ЧС [3, 4].

Заключение

На сегодняшний день готовых решений, реализующих подобный функционал, не существует. Это отчасти связано с новизной задачи, использующей тенденции и информационные технологии, получившее



Рис. 1. Информеры-графики. В реальном времени позволяют с одного взгляда обнаружить всплеск активности в социальных сетях

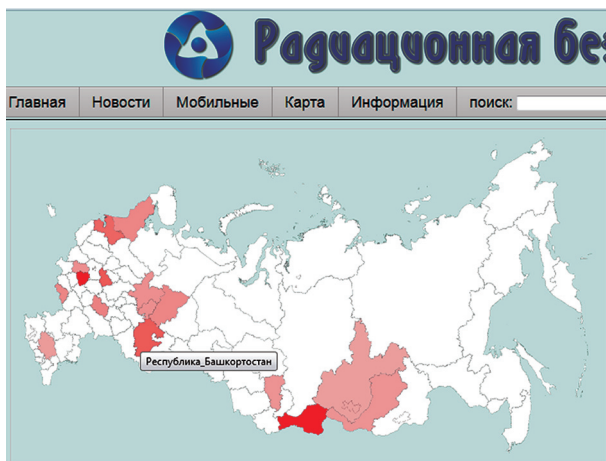


Рис. 2. Информер-карта. Наглядно показывает регионы России, повысившие свою активность в социальных сетях за последнее время — день, неделю, месяц



Рис. 3. Раздел новостей из социальных сетей. Позволяет найти и прочитать сообщения за определенный период времени, из конкретного региона, по конкретному объекту

Информеры - рейтинги

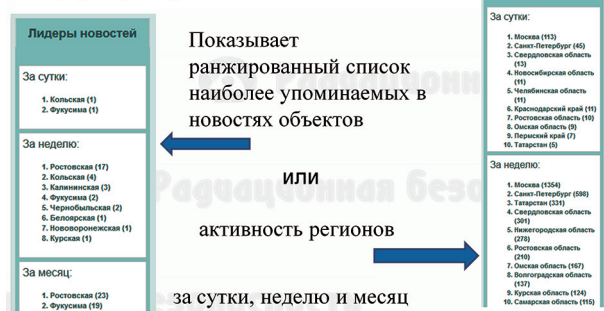


Рис. 4. Информеры-рейтинги активности за период времени

развитие лишь в последние годы. Но не в последнюю очередь отсутствие реализаций объясняется также сложностью возникающих в ходе решения подзадач, задействующих различные области смежных разнородных знаний.

Описанные в данной статье идеи и реализации могут стать основой для создания будущих систем работы с населением на территориях с расположенными потенциально опасными объектами.

Работа поддержана грантом Минобрнауки № RFMEFI57414X0038 в рамках реализации ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014—2020 годы».

Литература

- Браславец Л.А. Социальные сети как средство массовой информации: к постановке проблемы // Вестник воронежского государственного университета. Сер.: Филология. Журналистика. 2009. № 1. С. 125.
- Губанов Д. А., Новиков Д. А., Чхартишвили А.Г. Социальные сети: модели информационного влияния, управления и протестов. М.: Изд-во физико-математ. лит.-ы. 2010. Т. 228.
- Марченко Т.А., Мельникова Т.Б. Оценка информационной составляющей социально-психологической обстановки на радиоактивно загрязненных территориях уральского региона // Технологии гражданской безопасности. 2014. № 1 (39). С. 24.
- Морозова Е.В., Мирошников И.В. Сетевые сообщества в условиях чрезвычайных ситуаций: типы поведения акторов // Человек. Сообщество. Управление. 2010. № 4. С. 16.
- Фещенко А.В. Социальные сети в образовании: анализ опыта и перспективы развития // Открытое дистанционное образование. 2011. № 3. С. 43.

Сведения об авторах

Галиуллин Марат Энзимович: ФГБОУ ВПО «Удмуртский государственный университет», ассистент каф. 426000, г. Ижевск, ул. Университетская, 1, корп. 6, каб. 309. Тел.: (3412) 68-38-31. E-mail: m.a.galiullin@gmail.com

Information about authors

Galiullin Marat E.: Udmurt State University, Assistant of the Department. 426000, Izhevsk, University str., 1, bldg. 6, room. 309. Tel.: (3412) 68-38-31. E-mail: m.a.galiullin@gmail.com