

МРНТИ 20.53.19

Подходы к разработке информационной системы мониторинга мнений и оценки социального самочувствия

Мамыкова Ж.Д., Казахский национальный университет им. аль-Фараби,
г. Алматы, Республика Казахстан, E-mail: Zhanl.Mamykova@kaznu.kz
Мутанов Г.М., Казахский национальный университет им. аль-Фараби,
г. Алматы, Республика Казахстан, E-mail: Rector@kaznu.kz
Сундетова Ж.Т., Казахский национальный университет им. аль-Фараби,
г. Алматы, Республика Казахстан, E-mail: Zhanerke.Sundetova@kaznu.kz
Торекүл С.М., Казахский национальный университет им. аль-Фараби,
г. Алматы, Республика Казахстан, E-mail: Saule.Torekul@kaznu.kz

Стремительное распространение социальных онлайн-сервисов и развитие технологий Больших Данных инициировали интерес к использованию сведений из социальных сетей в различных отраслях. Сегодня приобретают популярность технологии "мониторинга социальных сетей" (social listening) и контент-анализа. Данные сервисы в основном представлены зарубежными разработками, лингвистические словари, которых построены на английском языке, и они плохо адаптированы под русский и казахский языки.

В представленной статье описан процесс проектирования и разработки информационной системы для "мониторинга социальных сетей адаптированной под специфику казахского и русского языка, а также форм сленгового общения, с целью определения эмоционального окраса пользовательского восприятия контента, и оценки социального самочувствия в стране/регионе. В работе представлен алгоритм информационной системы с описанием функциональных возможностей, а также методы и модели. Кроме этого, рассматривается производственная модель оценки социального самочувствия и модель семантического профиля восприятия обществом событий. Проведен сравнительный анализ разработок в вопросах мониторинга мнений для выявления сильных и слабых сторон. Информационная система даст возможность проводить мониторинг СМИ, социальных сетей, анализа управления репутацией, анализа пользовательского восприятия контента сети Интернет на региональном, республиканском и международном уровне.

Ключевые слова: мониторинг социальных сетей, анализ комментариев, социальное самочувствие, оценка пользовательского восприятия.

Пікірлерге мониторинг жүргізу және әлеуметтік көңіл кәйді бағалау ақпараттық жүйесін жобалау және құрастыру

Мамыкова Ж.Д., әл-Фараби атындағы Қазақ ұлттық университеті,
Алматы қ., Қазақстан Республикасы, E-mail: Zhanl.Mamykova@kaznu.kz
Мутанов Г.М., әл-Фараби атындағы Қазақ ұлттық университеті,
Алматы қ., Қазақстан Республикасы, E-mail: Rector@kaznu.kz
Сундетова Ж.Т., әл-Фараби атындағы Қазақ ұлттық университеті,
Алматы қ., Қазақстан Республикасы, E-mail: Zhanerke.Sundetova@kaznu.kz
Торекүл С.М., әл-Фараби атындағы Қазақ ұлттық университеті,
Алматы қ., Қазақстан Республикасы, E-mail: Saule.Torekul@kaznu.kz

Әлеуметтік онлайн-сервистердің кең таралуы және Үлкен Деректер технологияларының дамуы әлеуметтік желілердегі деректерді әртүрлі салаларда қолдануға деген қызығушылық тудырды. Қазіргі таңда контент анализ және "әлеуметтік желілерді мониторинг" технологиялары танымалдылыққа ие болуда. Мұндай сервистердің көпшілігі шетелдік әзірлемелер болып табылады және лингвистикалық сөздіктер ағылшын тілінде құрастырылған, сондықтан қазақ және орыс тілдеріне бейімделмеген.

Мақалада елдің/аумақтың әлеуметтік көңіл күйін бағалау және пайдаланушының контентті қабылдауының эмоционалды түсін анықтауға арналған, қазақ және орыс тілдерінің, сонымен қатар ауызекі сөйлеу формаларының ерекшеліктеріне бейімделген "әлеуметтік желілерді мониторинг" ақпараттық жүйесін жобалау және әзірлеу процесі көрсетілген. Жұмыста ақпараттық жүйе алгоритмінің функционалдық мүмкіндіктерінің сипаттамасы және моделдері мен әдістері көрсетілген. Сонымен қатар, әлеуметтік көңіл күйді бағалаудың өнімдік моделі және қоғамның оқиғаларды қабылдауының семантикалық моделі қарастырылған. Мықты және әлсіз тұстарын анықтау мақсатында пікірлерді мониторинг саласындағы әзірлемелерге салыстырмалы анализ жүргізілген. Ақпараттық жүйе аудандық, республикалық және халықаралық деңгейде БАҚ-ты және әлеуметтік-желілерді мониторингқа, беделдерді басқаруға анализ жасауға, интернет контентті пайдаланушының қабылдауына анализ жасауға мүмкіндік береді.

Түйін сөздер: әлеуметтік желілерді мониторинг, комментарийлерге анализ жасау, әлеуметтік көңіл-күй, пайдаланушының қабылдауын бағалау.

Approaches of the development of information monitoring system of social wellness opinions and evaluation

Mamykova Zh.D., al-Farabi Kazakh National university,
Almaty, Republic of Kazakhstan, E-mail: Zhanl.Mamykova@kaznu.kz
Mutanov G.M., al-Farabi Kazakh National university,
Almaty, Republic of Kazakhstan, E-mail: Rector@kaznu.kz
Sundetova Zh.T., al-Farabi Kazakh National university,
Almaty, Republic of Kazakhstan, E-mail: Zhanerke.sundetova@kaznu.kz
Torekul S.M., al-Farabi Kazakh National university,
Almaty, Republic of Kazakhstan, E-mail: Saule.Torekul@kaznu.kz

A dynamic spreading of social on-line services and development of the Big Data technologies have caused interest to use information from social media in different spheres. Nowadays the "social listening" and content analysis technologies gain popularity. The given services are basically represented by foreign development works, where linguistic dictionaries are made in English and they are badly adapted to Russian and Kazakh.

The given article describes the process of designing and developing the information system to monitor "social media adapted to the specific character of the Kazakh and Russian languages and forms of slang communication as well to define the emotional coloring of a user's perception of the content and evaluation of the social well-being in the region and country. The work represents the algorithm of the information system describing functionalities and methods and models as well. Besides, a production model of social well-being evaluation and a semantic profile model for conceiving events by society has been considered. A comparative analysis of the development works has been made in the issues of opinions monitoring to reveal strengths and weaknesses. The information system gives an opportunity to monitor mass and social media, to analyze reputation management, to analyze a user's perception of the Internet content on the regional and a nationwide scale.

Key words: social media monitoring, comment analysis, social well-being, a user's perception evaluation.

1 Введение

В результате развития социальных сетей, СМИ, блогов, веб-ресурсов появилась острая необходимость, в инструменте определения оценки эмоционального окраса

тональности текстов, сообщений, возможность проведения репутационного аудита в режиме реального времени, выявления репутационных и информационных угроз, критик, негатива, дезинформации. Совместное использование неструктурированных и контентных данных потенциально позволяет применять социальные сети для решения широкого круга бизнес-задач: маркетинговая деятельность (мнение потребителей, анализ отзывов), управления брендом, мониторинг для бизнеса для отслеживания ключевых слов в режиме реального времени от основных новостных агентств и социальных медиа-платформ и др.

Проблемы систем мониторинга СМИ: слабое применение данных технологий для анализа интернет-пространства и оценки социального самочувствия в обществе; любая ситуация мониторится вручную; субъективизм в интерпретации данных событий; отсутствие практики применения методов Data Mining для обработки информационного массива потоков СМИ; отсутствие системы кодирования, позволяющей классифицировать текстовые сообщения в соответствии с категориями событий; сервисы мониторинга социальных сетей в основном представлены зарубежными разработками, лингвистические словари, которых построены на английском языке; отечественные системы мониторинга и анализа социальных сетей, плохо адаптированы под казахский язык.

В связи с чем, особую актуальность приобретает разработка Информационной системы мониторинга мнений и оценки социального самочувствия (далее - Система).

Основной целью разработки Системы является обеспечение сервиса мониторинга мнений интернет-пространства для проведения глубинного анализа неструктурированной информации, с целью извлечения знаний и получения важной информации из текстовых массивов по оценке пользовательского восприятия контента, и оценки социального самочувствия общества.

Для разработки системы анализа пользовательского восприятия контента социальных сетей, адаптированной под специфику казахского и русского языка, а также форм сленгового общения были поставлены следующие задачи: разработать модуль подключения к API-социальных сетей, разработать модуль управления тональным словарем, разработать модуль хранения данных для обработки и анализа результатов, разработать модуль моделирования социального самочувствия, программная реализация системы.

Реализация Системы позволит мониторить веб-ресурсы и социальные сети с подсистемой моделирования социального самочувствия общества, посредством оценки тональности сообщений/мнений.

2 Обзор литературы

Анализу проблем функционирования программных инструментов [1], предназначенных для мониторинга социальных сетей посвящены труды зарубежных ученых, которые позволили в полной степени рассмотреть характеристики программного обеспечения [2], типы социальных сетей, включая языковую поддержку [3], инструменты и сервисы, предназначенные для мониторинга социальной сети с использованием современных методов анализа синтаксиса и семантики на примере китайского языка [4]. В аспекте новых моделей мониторинга социальных сетей была рассмотрена модель для

мониторинга того, как социальные связи и социальная структура меняются внутри группы [5], абстрактная модель для сравнения метрик с одинаковым значением в разных социальных сетях [6] и концептуальные модели [7,8]. Проанализирован алгоритм мониторинга [9], который анализирует потоки из социальных сетей, сочетая в себе различные подходы машинного обучения с легкой степенью адаптированности и обогащен пользовательскими знаниями без необходимости сложной настройки, алгоритм сегментации по определению новых ключевых слов, где в качестве данных используются твиты, собранные из Twitter [10], алгоритм MMCRD (maximum monitoring contribution rate deployment) [11] и структура с открытым исходным кодом для мониторинга, анализа и поиска в нескольких социальных сетях [12].

Исследования в рамках разработки новых методов и методологии мониторинга социальных сетей представлены как методология информационного взаимодействия в социальных системах [13], методы измерения: пассивные, активные [14], с применением методов теории систем и кибернетики [15]. Методология, которая позволяет справиться с дефицитом данных при мониторинге реальных событий с использованием данных социальных сетей [16]. С точки зрения социального эффекта была изучена методология мониторинга социальных сетей, предназначенная для борьбы с кризисами (лесные пожары, вспышек заболеваний, преступления) [17].

Данные подходы позволили определить архитектуру разрабатываемой Системы с акцентом на построение лингвистического словаря для казахского языка и разработки алгоритмов чтения контента с "пониманием". Изучение зарубежного опыта нашло отражение в определении подходов в разработке алгоритмов машинного обучения для казахского языка и разработку интеллектуальной составляющей как платформы анализа данных.

3 Материал и методы

Происходящие глобальные изменения, связанные с интернет-технологиями и сетевыми инновациями, объясняют необходимость разработки инструментария и моделей для мониторинга общественного мнения и оценки социального самочувствия как одного из главных инструментов оценки по состоянию социально-экономической системы в контексте устойчивого развития.

На зарубежном рынке уже представлены инструменты по мониторингу социальных сетей и контент-анализу, как Hootsuite [18], Socialmention [19], BuzzLook [20], IQBuzz [21], Socialbakers [22], Keyhole [23], Brandwatch [24], YouScan [25] и другие аналоги, которые позволяют, применять социальные сети для решения широкого круга бизнес-задач: борьбы с мошенничеством, управления брендом, рекламы товаров и услуг, формирования новых каналов сбыта и др.

Рынок Казахстана не богато представлен отечественными разработками в вопросах мониторинга мнений (ИС "iMAS"[26], "Alem Media Monitoring"[27]) и разработка зарубежных компаний (ИС "Brand Analytics"[28], "Microsoft engagement"[29]).

Таблица 1: Функциональные возможности программных продуктов.

	iMAS	Alem Media Monitoring	Предлагаемая авторами решение	Brand Analytics	Microsoft engagement
1	☑	☑	☑	☑	
2	☑	☑	☑	☑	☑
3	☑		☑	☑	☑
4	☑		☑	☑	☑
5	☑	☑	☑	☑	☑
6	☑	☑	☑	☑	☑
7			☑		
8	☑		☑	☑	☑
9			☑		☑
10			☑		☑
11			☑	☑	☑
12			☑	☑	☑
13			☑	☑	
14			☑		
15			☑		
16			☑		

Примечание (модули систем): 1. Поддержка казахского языка; 2. Поддержка русского языка; 3. Поддержка английского языка; 4. Географическо-интерактивная карта мнений; 5. Мониторинг социальных сетей; 6. Мониторинг интернет-ресурсов; 7. Моделирование социального самочувствия; 8. Набор слов-исключений; 9. Сервис настройки каталога ресурсов; 10. Сервис добавления новых ресурсов; 11. Сервис настройки профиля правил для темы поиска; 12. Детализация анализа; 13. Сервис настройки выгрузки данных; 14. Отчеты по КРІ системы; 15. Возможность приглашать "эксперта"; 16. Модуль управления тональным словарем.

Согласно данным Таблицы 1 все программные продукты поддерживают 3 языка (русский, английский и казахский) исключением является Microsoft engagement, у которого нет поддержки казахского языка. Для всех программных продуктов мониторинга мнений источниками являются социальные сети и интернет-ресурсы. Большинство программных решений позволяют увидеть географическо-интерактивную карту мнений, визуализацию результатов работы системы и дают возможность сузить круг поиска, с помощью добавления слов-исключений. Лишь отдельные программные продукты содержат в себе сервисы настройки каталога ресурсов, добавления новых ресурсов, настройки профиля правил для поиска, детализацию анализа и настройки выгрузки данных. В рассмотренных программных продуктах отсутствуют модули управления тональным словарем, моделирования социального самочувствия и возможность приглашения "эксперта".

Процесс проектирования подобных систем представляют собой платформенное решение. На основе использования функциональных возможностей аналогичных систем, нами была применена клиент-серверная архитектура программного

решения, использованы технологии проектирования информационных систем, технология программирования с применением языка Python, фреймворка Django, проектирование реляционной базы данных на основе PostgreSQL. Используются методы продукционной модели, позволяющие построить гибкие правила оценки социального самочувствия, базирующиеся на статистических показателях социально-экономического развития страны, модель семантического профиля на основе стандарта для оценки эффективности контентной политики и коммуникации на площадке социальных сетей, разработанная сервисом SocialBakers [30]. Данные модели и методы позволили нам построить платформу количественного анализа и систему организации хранения этих данных. Предложенная система мониторинга социальных сетей и анализа настроения, предполагает использование алгоритмов машинного обучения. Анализ литературных источников [3,4,16] показал, что для создания полноценной системы необходимо применить алгоритмы машинного обучения (классификация, кластеризация, TF-IDF, Word2Vec и т.д.) для казахского языка. Тем самым позволив построить лингвистический словарь, который сформирован для казахского языка и решить задачу автоматической оценки тональности текста, с последующим прогнозированием настроения пользователей. Для следующего этапа развития системы реализовать интеллектуальный модуль путём формирования корпуса текстовых массивов казахского языка и его разметки по тональности; интегрировать модуль машинного обучения с платформой мониторинга социальных сетей.

3.1 Аспекты проектирования системы

Архитектура системы базируется на основе трехуровневой модели клиент-серверного приложения (графический интерфейс пользователя - бизнес-логика - СУБД), которая представлена на рис.1.

Система состоит из:

1. Модуля "Коннектор позволяющий производить настройки подключения к интернет-ресурсам и API известных социальных сетей из каталога ресурсов;
2. Модуля "Лингвистический конструктор позволяющий формировать тональный словарь на казахском и русском языках, с возможностью динамического наполнения и определения оценки тональности;
3. Модуля "Платформа обработки данных и анализа позволяющий моделировать социальное самочувствие по принципу "лакмусовой бумажки" и строить количественный анализ результатов мониторинга социальных сетей, производить предобработку и обработку данных с целью определения оценки тональности, с применением алгоритмов машинного обучения;
4. Модуля "Визуализации формирующий отчеты в формате PDF и позволяющий строить графические представления результатов количественного анализа.

Основное назначение Системы является обеспечение сервиса мониторинга мнений интернет-пространства для проведения глубинного анализа неструктурированной информации, с целью извлечения знаний и получения важной информации из текстовых

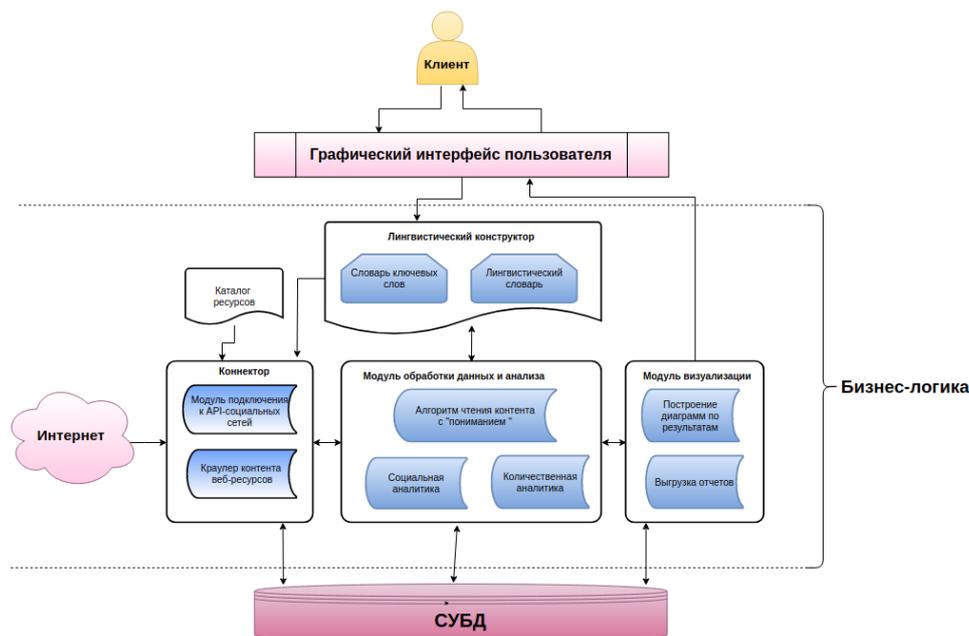


Рисунок 1: Архитектура Системы

массивов по оценке пользовательского восприятия контента, и оценки социального самочувствия общества.

Разработанная система имеет встроенные модули подключения к API социальных сетей Вконтакте, Facebook, Twitter, на этапе разработки подключение к Instagram, функциональные возможности позволяют:

- Работать с текстами на 2-х языках (казахский, русский) с автоматическим определением языка текста (казахский, русский, смайлы, символы);
- Определять тональность темы/текста (отрицательная, положительная, нейтральная);
- Поддерживать тональный словарь на казахском и русском языках;
- Формировать базу данных слов с ошибками на казахском и русском языках, с целью увеличить результаты процедуры оценки тональности темы (это связано с тем, что комментарии написаны с ошибками);
- Фиксировать диапазон времени при процедуре мониторинга социальных сетей (от года, 6 мес., 3 мес., месяц, неделя, день);
- Строить интегральную оценку тональности темы;
- Выявлять профиль участника социальных сети, путем чтения данных по профилю участника; подсчета активности участника в теме (количество комментариев, лайков, репостов);

- Приглашать эксперта участвовать в дискуссии и наблюдать статистику мониторинга по изменению мнений под влиянием эксперта;
- Рассчитать индекс социального самочувствия общества;
- Работать с большим объемом данных (заложены принципы масштабирования и параллельной обработки);
- Создавать различные отчеты (по категории, ресурсам, оценки тональности, мультязычности и др) в виде различных графиков и диаграмм с выгрузкой в формате PDF.

Разработанная архитектура системы базируется на ролевой политике доступа к системе, которая обеспечивает организацию алгоритмического подхода работы пользователя. В основе любого алгоритма действий пользователя лежит тема, которая представляет собой набор заданных параметров суперпользователем, такие как: название, категория, подкатегория темы, набор ключевых слов, язык, ресурсы и временной интервал поиска.

С учетом выше названных пунктов определен алгоритм работы пользователя с системой.

Таблица 2: Алгоритм Системы

Шаг 1. Запускаем систему (социальные сети, интернет-ресурсы новостных порталов)
Шаг 2. Создаем новую тему для анализа (тема, категория, подкатегория, период, язык поиска, область поиска, ключевые слова)
Например: Проблемы безопасности граждан (Категория: общество, безопасность; Диапазон поиска: 10.10.18-10.07.18; рус./каз., соц.сети: Вконтакте, facebook, Twitter; ключевые слова: потерпевший, убийство, задержан, следствие, МВД, полицейские)
Шаг 3. Запускаем поисковый механизм (поиск в режиме реального времени, подключение тонального словаря на рус/каз языках с целью определения тональности темы и комментариев; расчет социального самочувствия).
Шаг 4. Изучаем результаты поиска и анализа (отчеты по результатам поиска, аналитика по: оценки тональности, теме, записям, ресурсам, авторам, языку, работе эксперта) Например: Анализ по теме "Проблемы безопасности граждан". Общее количество результатов: 3594; из них записей: 973, комментариев: 2621 Результатов из "Вконтакте": 3090, из "Twitter": 504 на русском языке: 2632; на казахском языке: 63, на других языках: 899 Оценка тональности темы: отрицательная; "отрицательные": 1743, "положительные": 468, "нейтральные": 1383 Активные участники: Ф. И. О., Ф. И. О., Ф. И. О. Популярные слова: "жертва", "потерпевший", "убийство". Местоположение: определено 49,14% (1766), из Казахстана: 52,5% (927), из России: 47,5% (839)
Шаг 5. Визуализация результатов анализа (диаграммы: оценка тональности, типы ресурсов, количество записей; оценка социального самочувствия)
Шаг 6. Сохранение результатов анализа (файл в формате pdf, разметка по датам поиска)

Длительность работы Системы по теме поиска занимает определенное время, которое зависит от периода поиска, количества ключевых слов и набора ресурсов для поиска, что в свою очередь влияет на количество возвращаемых Системой результатов мониторинга и на длительность обработки темы поиска. По завершению поиска по заданным параметрам можно получить "Сводный отчет по результатам темы". В отчете по результатам заданной темы отображаются все параметры, по которым производился поиск, количество найденных записей, количество найденных комментариев, общая тональность Темы.

Страница "Сводный отчет по результатам темы" представлена на рис.2.

Сводный отчет по результатам темы "Проблемы безопасности граждан"

Тема	Проблемы безопасности граждан
Категория	Общество
Подкатегория	Происшествия
Ключевые слова	потерпевший, жертва, убийство, задержан, следствие, МВД, полицейские;
Ресурсы	Vk, Twitter,
Язык	KZ, RU
Количество результатов	3594
Количество записей	973 Подробнее >>
Количество комментариев	2621 Подробнее >>
Диапазон времени	С 10 октября 2018 г. 3:08 по 17 июля 2018 г. 20:48
Оценка тональности темы	Отрицательный
Количество запросов модератору	
Участие эксперта	

Получить отчет в PDF

Посмотреть диаграмму по теме "Проблемы безопасности граждан"

Рисунок 2: Страница "Сводный отчет по результатам темы"

Кроме того, в Системе предусмотрен функционал визуализации различных отчетов по всем результатам (за весь период мониторинга, за квартал, за 1 месяц, за 1 неделю, за 1 день), диаграммы представлены на рис.3 а,б,в,г.

Модуль "Мониторинга социальных сетей" отвечает за выявление тональности пользовательских мнений, путем мониторинга аккаунтов новостных порталов в известных социальных сетях. Для этого используется тональный словарь на казахском и русском языках. При проектировании тонального словаря были предусмотрены возможности создавать слова обычной формы, сленговые варианты слов, слова с ошибками, и предусмотрен функционал для добавления новых слов с целью динамического наполнения словаря.

Результативность работы разработанной системы можно определить путем оценки семантического профиля по результатам поиска. Для этого были рассмотрены модели стандартов оценки вовлеченности, которые основаны на способе измерения показателей

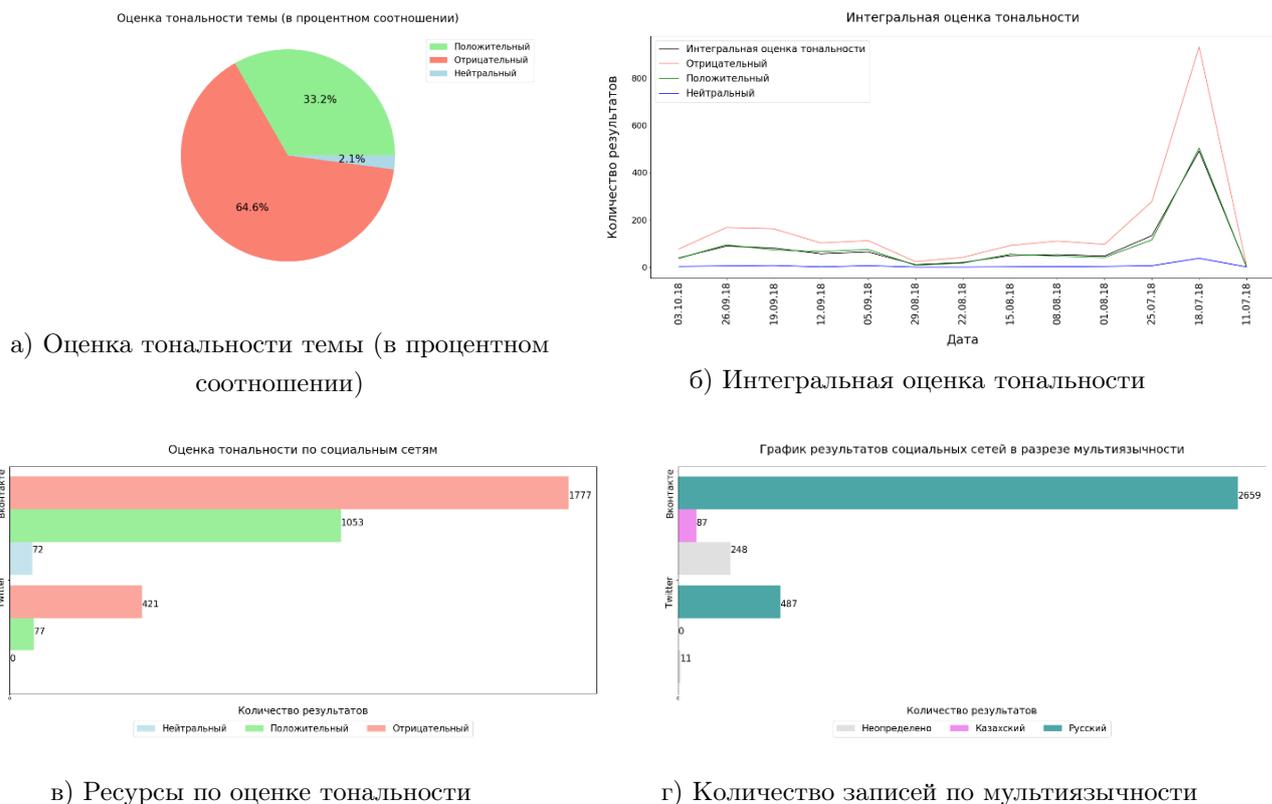


Рисунок 3: Диаграммы модуля "Визуализации мониторинга"

социальных сетей для SMM менеджмента, по формуле SocialBakers [31] на примере Facebook [32]. Исходя из которых, были определены следующие показатели для оценки семантического профиля восприятия обществом событий: уровень заинтересованности темой в обществе (R_{CT}), уровень активности обсуждения темы (R_{CE}), уровень настроения общества (R_{TS}).

Формула SocialBakers для расчета вовлечения аудитории имеет следующий вид:

$$\text{Daily ER} = \frac{\text{Total Actions Taken}}{\text{Daily coverage}} \times 100\% \quad (1)$$

где, Total Actions Taken – общее число действий, а именно – Лайки, Комментарии, Шэры, Вступления; а Daily Coverage – общий дневной охват контента, включая виральный и оплаченный [33]. На основании формулы 1 предлагаем расчет уровня активности обсуждения темы (R_{CE}), как:

$$R_{CE} = \frac{C}{CP} \times 100\% = R_{CE} = \frac{L+R+C}{CS} \times 100\%.$$

где, S – сумма количества подписчиков источников, CP – количество найденных постов по теме, CE – количество взаимодействий участниками по теме (комментарии (C), лайки (L), репосты (R)).

Уровень заинтересованности темой в обществе (RCT) определяется, как:

$$R = \frac{\times 100\%}{\max_{CT}}$$

где, – количество найденных записей по данной теме, \max – максимальный объем записей для темы (должно устанавливаться экспертом за определенный период времени).

Уровень настроения общества (RTS) рассчитывается, как суммарная оценка результатов тем мониторинга социальных сетей (положительная, нейтральная, отрицательная).

Помимо семантического профиля в основе программного продукта лежит оценка социального самочувствия, которая позволит определить, что послужило формированию такого настроения пользователей в социальных сетях.

Проведенный анализ на тему "Проблемы безопасности граждан" показал, анализ настроения общества - отрицательный. Согласно оценке индекса социального самочувствия, проблема настроения в обществе лежит глубже. Например, "число зарегистрированных преступлений" по г.Алматы на 2016 году составило 68704 преступлений, а в 2017 году - 63797 [34], являются самыми высокими по всей республике, чем и объясняется причина отрицательного настроения общества. Предлагаемый в данной Системе подход к анализу социального самочувствия, путем оценки продукционной модели ключевых показателей статистики, позволяет определить эмоциональное состояние общества по принципу "лакмусовой бумажки".

Для моделирования социального самочувствия использованы следующие статистические данные [34]: Занятое население; Безработное население; Уровень безработицы; Среднемесячная номинальная заработная плата; Среднемесячный прожиточный минимум населения; ВВП на душу населения; Доля населения с доходами ниже величины прожиточного минимума в общей численности; Число зарегистрированных преступлений. В основу расчета продукционной модели взят алгоритмический подход, где сравниваются значения показателей за текущий и предыдущие годы, либо берется отношение значения показателя к среднему или минимуму. В результате были получены четыре выходных результата определяющих индекс социального самочувствия: социальное напряжение - критическая зона, не ярко выраженное социальное напряжение, удовлетворительное социальное самочувствие, стабильное социальное самочувствие. По индексу социального самочувствия формируется диаграмма оценки социального самочувствия общества (рис. 4).

4 Результаты и обсуждение

Сравнительный анализ аналогичных зарубежных программных продуктов позволил сделать вывод о том, что данные решения при введении входных данных на казахском и русском языках не возвращают результаты мониторинга, такие как, тексты и количественные данные записей и комментариев по теме поиска. Это объясняется тем, что информационное поле поиска изученных систем ориентировано на интернет-пространство только зарубежных стран, и не охватывает отечественное интернет-пространство, а алгоритмы обработки данных и алгоритм определения

Главная → Моделирование социального самочувствия → Расчет социального самочувствия → Диаграмма по оценке социального самочувствия общества

Диаграмма по оценке социального самочувствия общества в разрезе всей страны

* Диаграмма построена на основе данных статистики 2013 и 2017 гг.
* Шаг шкалы Период: 1 год.

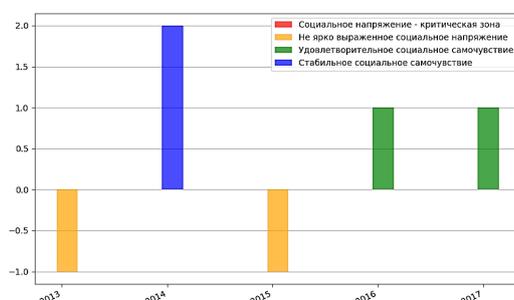


Рисунок 4: Диаграмма по оценке социального самочувствия общества с 2013 по 2017 гг.

оценки тональности текста зарубежных программных продуктов не адаптированы для казахского языка. Однако, по результатам мониторинга разработанная Система выдает всю детальную информацию по теме поиска. Таким образом, сравнительно-сопоставительный анализ функциональных возможностей зарубежных систем с разработанной Системой является не рациональным.

Практическая значимость исследования состоит в том, что предлагаемая информационная система позволит не только мониторить социальные сети и интернет-ресурсы, но и выявлять причины пользовательского мнения, путем расчета индекса социального самочувствия и с помощью управления тональным словарем можно обучать систему для более точных результатов.

В результате проведенных работ информационная система мониторинга мнений является первой отечественной разработкой с поддержкой тонального словаря на казахском языке, построенного с применением алгоритмов машинного обучения.

5 Заключение

Потребность в Казахстане в подобной системе есть. Так как, важно выявлять мнение потребителей, например, образовательных, медицинских и государственных услуг, с целью выявления "узких мест" в работе соответствующих ведомств, оптимизации работы и сервисов, принятия оперативных решений по проблемным вопросам. Ранний доступ к анализу и информации обеспечивает лучшее принятие решений и более построения качественных информированных ответов, для немедленного реагирования.

Работа по мониторингу социальной сети и контент-анализу позволит: осуществить детектирование информационного воздействия; определять круг интересов и распространения информации; осуществлять сбор расширенных социальных профилей для маркетинговой деятельности; отслеживать динамику вовлеченности пользователей в ту или иную тему; выявлять информационные поводы и отношения к ним; мониторить настроения интернет-пользователей на предмет восприятия конкретных тем обсуждения.

Данная система позволит развить новое направление рынка программных решений – разработка информационной системы мониторинга мнений и социальных сетей для обработки и анализа текстовых данных, с целью получения информации по оценке пользовательского восприятия контента и оценки социального самочувствия общества.

Список литературы

- [1] *Benedetto F., Tedeschi A.* Big Data Sentiment Analysis for Brand Monitoring in Social Media Streams by Cloud Computing // Sentiment Analysis and Ontology Engineering. – 2016. – № 639. – С. 341-377.
- [2] *Gu J., Xiao B., Gu Z.* Software Design for Network Monitoring System // International Conference on Information Computing and Applications. – Berlin: Springer, 2012. – С. 41-46.
- [3] *Laine M.O.J., Frühwirth C.* Monitoring Social Media: Tools, Characteristics and Implications // International Conference of Software Business. – Berlin: Springer, 2010. – С. 193-198.
- [4] *Liu Y.* Social Media Monitoring // Social Media in China. – 2016. – С. 185-193
- [5] *Gesell, S.B., Barkin, S.L. & Valente, T.W.* Social network diagnostics: a tool for monitoring group interventions // Implementation Science. – 2013. – № 8(116). – С. 1-12.
- [6] *Tesconi M., Gazzè D., Duca A.L.* SocialTrends: A Web Application for Monitoring and Visualizing Users in Social Media // International Conference on Social Informatics. – Berlin: Springer. – 2012. – С. 535-538.
- [7] *Averchenkov V., Budylskii D., Podvesovskii A., Averchenkov A., Rytov M., Yakimov A.* Hierarchical Deep Learning: A Promising Technique for Opinion Monitoring and Sentiment Analysis in Russian-Language Social Networks" // Creativity in Intelligent Technologies and Data Science. – 2015. – № 535. – С. 583-592.
- [8] *Psallidas F., Ntoulas A., Delis A.* Soc Web: Efficient Monitoring of Social Network Activities // International Conference on Web Information Systems Engineering - Heidelberg: Springer. – 2013. – С. 118-136.
- [9] *Kirsch B., Giesselbach S., Knodt D., Rüping S.* Robust End-User-Driven Social Media Monitoring for Law Enforcement and Emergency Monitoring" // Community-Oriented Policing and Technological Innovations. – 2018. – С. 29-36.
- [10] *Xiaoyan W., Kai X., Ying S., Jian-long T., Li G.* Research of New Words Identification in Social Network for Monitoring Public Opinion // International Conference on Trustworthy Computing and Services. – Heidelberg: Springer. – 2013. – С. 598-603.
- [11] *Guo Z., Wang Z., Zhang R.* MMCRD: An Effective Algorithm for Deploying Monitoring Point on Social Network // China Conference on Knowledge Graph and Semantic Computing. – Singapore: Springer. – 2018. – С. 40-51.
- [12] *Schinas M., Papadopoulos S., Apostolidis L., Kompatsiaris Y., Mitkas P.A.* Open-Source Monitoring, Search and Analytics Over Social Media // International Conference on Internet Science. – Greece: Springer. – 2017. – С. 361-369.
- [13] *Grüblbauer J., Haric P.* Social Media Monitoring Tools as Instruments of Strategic Issues Management // Handbook of Social Media Management. – 2013. – С. 671-687.
- [14] *Gyarmati L., Trinh T.A.* Measurement Methods of User Behavior in Online Social Networks // Computational Social Networks. – 2012. – С. 157-175.
- [15] *Stiehm S., Welter F., Richert A., Jeschke S.* Shaping the Future Through Cybernetic Approaches of Social Media Monitoring // Automation, Communication and Cybernetics in Science and Engineering. – 2016. – С. 179-191.
- [16] *Souza R.C.S.N.P., de Brito D.E.F., Cardoso R.L., de Oliveira D.M., Meira W., Pappa G.L.* An Evolutionary Methodology for Handling Data Scarcity and Noise in Monitoring Real Events from Social Media Data // Ibero-American Conference on Artificial Intelligence. – Chile: Springer. – 2014. – С. 295-306.
- [17] *Sparks R.* Social Network Monitoring: Aiming to Identify Periods of Unusually Increased Communications Between Parties of Interest // Frontiers in Statistical Quality Control 11. – 2015. – С. 3-13.
- [18] *Hootsuite* [Электрон. ресурс]. – URL: <https://hootsuite.com/> (дата обращения: 21.11.2018).
- [19] *Socialmention* [Электрон. ресурс]. – URL: <http://socialmention.com/about/> (дата обращения: 21.11.2018).

- [20] *BuzzLook* [Электрон. ресурс]. – URL: <https://br-analytics.ru/about/> (дата обращения: 21.11.2018).
- [21] *IQBuzz* [Электрон. ресурс]. – URL: <http://iqbuzz.pro/about.php> (дата обращения: 21.11.2018).
- [22] *Socialbakers* [Электрон. ресурс]. – URL: <https://www.socialbakers.com/companу/> (дата обращения: 21.11.2018).
- [23] *Keyhole* [Электрон. ресурс]. – URL: <https://keyhole.co/about-us> (дата обращения: 23.11.2018).
- [24] *Brandwatch* [Электрон. ресурс]. – URL: <https://www.brandwatch.com/about/> (дата обращения: 23.11.2018).
- [25] *YouScan* [Электрон. ресурс]. – URL: <https://youscan.io/product/> (дата обращения: 23.11.2018).
- [26] *iMAS* [Электрон. ресурс]. – URL: <https://imas.kz> (дата обращения: 24.11.2018).
- [27] *Alem Media Monitoring* [Электрон. ресурс]. – URL: <https://alem.kz> (дата обращения: 24.11.2018).
- [28] *Brand Analytics* [Электрон. ресурс]. – URL: <https://br-analytics.ru> (дата обращения: 24.11.2018).
- [29] *Microsoft engagement* [Электрон. ресурс]. – URL: <https://dynamics.microsoft.com/ru-ru/microsoft-social-engagement> (дата обращения: 24.11.2018).
- [30] *Официальный интернет ресурс SocialBakers* [Электрон. ресурс]. – URL: <https://socialbakers.com/> (дата обращения: 24.11.2018).
- [31] *Тихонова Е.* Что такое Engagement Rate и для чего это нужно? [Электрон. ресурс]. – 2018. – URL: <https://blog.sociate.ru/chto-takoe-engagement-rate-i-dlya-chego-eto-nuzhno> (дата обращения: 29.11.2018)
- [32] *Engagement Rate: A Metric You Can Count On* [Электрон. ресурс]. – 2018. – URL: <https://www.socialbakers.com/blog/1427-engagement-rate-a-metric-you-can-count-on> (дата обращения: 29.11.2018)
- [33] *Rumyantseva D.* ER - методы расчета и значение показателей [Электрон. ресурс]. – 2015. – URL: <https://www.cossa.ru/155/101492/> (дата обращения: 29.11.2018)
- [34] *Официальный интернет ресурс Комитета по статистике Министерства национальной экономики РК* [Электрон. ресурс]. – URL: <https://.stat.gov.kz> (дата обращения: 29.11.2018).

References

- [1] Benedetto F., Tedeschi A., "Big Data Sentiment Analysis for Brand Monitoring in Social Media Streams by Cloud Computing", (Sentiment Analysis and Ontology Engineering, 639 (2016), accessed on November 14, 2018, https://doi.org/10.1007/978-3-319-30319-2_14
- [2] Gu J., Xiao B., Gu Z., "Software Design for Network Monitoring System", (International Conference on Information Computing and Applications, 2012) : 41-46.
- [3] Laine M.O.J., Frühwirth C., "Monitoring Social Media: Tools, Characteristics and Implications", (International Conference of Software Business, 2010) : 193-198.
- [4] Liu Y., "Social Media Monitoring", (Social Media in China, 2016), accessed on November 14, 2018, https://doi.org/10.1007/978-3-658-11231-8_10
- [5] Gesell, S.B., Barkin, S.L. & Valente, T.W., "Social network diagnostics: a tool for monitoring group interventions", (Implementation Science, 2013), accessed on November 14, 2018, <https://doi.org/10.1186/1748-5908-8-116>
- [6] Tesconi M., Gazzè D., Duca A.L., "SocialTrends: A Web Application for Monitoring and Visualizing Users in Social Media", (International Conference on Social Informatics, 2012) : 535-538.
- [7] Averchenkov V., Budylyskii D., Podvesovskii A., Averchenkov A., Rytov M., Yakimov A., "Hierarchical Deep Learning: A Promising Technique for Opinion Monitoring and Sentiment Analysis in Russian-Language Social Networks" in *Creativity in Intelligent Technologies and Data Science*, ed Kravets A., Shcherbakov M., Kultsova M., Shabalina O. (Springer, Cham, 2015) : 583
- [8] Psallidas F., Ntoulas A., Delis A., "Soc Web: Efficient Monitoring of Social Network Activities", (International Conference on Web Information Systems Engineering, 2013) : 118-136.

- [9] Kirsch B., Giesselbach S., Knodt D., Rping S., "Robust End-User-Driven Social Media Monitoring for Law Enforcement and Emergency Monitoring" in *Community-Oriented Policing and Technological Innovations*, ed Leventakis G., Haberfeld M. (SpringerBriefs in Criminology. Springer, Cham, 2018) : 29.
- [10] Xiaoyan W., kai X., Ying S., Jian-long T., Li G., "Research of New Words Identification in Social Network for Monitoring Public Opinion", (International Conference on Trustworthy Computing and Services, 2013) : 598-603
- [11] Guo Z., Wang Z., Zhang R., "MMCRD: An Effective Algorithm for Deploying Monitoring Point on Social Network", (China Conference on Knowledge Graph and Semantic Computing, 2018) : 40-51.
- [12] Schinas M., Papadopoulos S., Apostolidis L., Kompatsiaris Y., Mitkas P.A., "Open-Source Monitoring, Search and Analytics Over Social Media", (International Conference on Internet Science, 2017) : 361-369
- [13] Grüblbauer J., Haric P., "Social Media Monitoring Tools as Instruments of Strategic Issues Management" in *Handbook of Social Media Management. Media Business and Innovation*, ed Friedrichsen M., Mühl-Benninghaus W. (Springer, Berlin, Heidelberg, 2013) : 671
- [14] Gyarmati L., Trinh T.A., "Measurement Methods of User Behavior in Online Social Networks", (Computational Social Networks, 2012), accessed on November 14, 2018, https://doi.org/10.1007/978-1-4471-4051-1_7
- [15] Stiehm S., Welter F., Richert A., Jeschke S., "Shaping the Future Through Cybernetic Approaches of Social Media Monitoring", in *Automation, Communication and Cybernetics in Science and Engineering 2016*, ed Jeschke S., Isenhardt I., Hees F., Henning K. (Springer, Cham) : 179
- [16] Souza R.C.S.N.P., de Brito D.E.F., Cardoso R.L., de Oliveira D.M., Meira W., Pappa G.L., "An Evolutionary Methodology for Handling Data Scarcity and Noise in Monitoring Real Events from Social Media Data", (Ibero-American Conference on Artificial Intelligence, 2014) : 295-306
- [17] Sparks R., "Social Network Monitoring: Aiming to Identify Periods of Unusually Increased Communications Between Parties of Interest", in *Frontiers in Statistical Quality Control 11*, ed Knoth S., Schmid W. (Frontiers in Statistical Quality Control. Springer, Cham, 2015) : 3
- [18] Official internet resource Hootsuite, accessed November 21, 2018, <http://socialmention.com/about/>
- [19] Official internet resource Socialmention, accessed November 21, 2018, <http://socialmention.com/about/>
- [20] Official internet resource BuzzLook, accessed November 21, 2018, <https://br-analytics.ru/about/>
- [21] Official internet resource IQBuzz, accessed November 21, 2018, <http://iqbuzz.pro/about.php>
- [22] Official internet resource Socialbakers, accessed November 21, 2018, <https://www.socialbakers.com/company/>
- [23] Official internet resource Keyhole, accessed November 23, 2018, <https://keyhole.co/about-us>
- [24] Official internet resource Brandwatch, accessed November 23, 2018, <https://www.brandwatch.com/about/>
- [25] Official internet resource YouScan, accessed November 23, 2018, <https://youscan.io/product/>
- [26] Official internet resource iMAS, accessed November 24, 2018, <https://imas.kz>
- [27] Official internet resource Alem Media Monitoring, accessed November 24, 2018, <https://alem.kz>
- [28] Official internet resource Brand Analytics, accessed November 24, 2018, <https://br-analytics.ru>
- [29] Official internet resource Microsoft engagement, accessed November 24, 2018, <https://dynamics.microsoft.com/ru-ru/microsoft-social-engagement>
- [30] Official internet resource SocialBakers, accessed November 24, 2018, <https://socialbakers.com/>
- [31] Tihonova E., "Chto takoe Engagement Rate i dlya chego eto nuzhno?", [What is the Engagement Rate and what is it for?], accessed November 29, 2018, <https://blog.sociate.ru/chto-takoe-engagement-rate-i-dlya-chego-eto-nuzhno>
- [32] "Engagement Rate: A Metric You Can Count On", accessed November 29, 2018, <https://www.socialbakers.com/blog/1427-engagement-rate-a-metric-you-can-count-on>
- [33] Rumyantseva D., "ER - metodyi rascheta i znachenie pokazateley", [ER - calculation methods and indicators], accessed November 29, 2018, <https://www.cossa.ru/155/101492/>
- [34] Official internet resource Committee on Statistics Ministry of National Economy of the Republic of Kazakhstan, accessed November 29, 2018, <https://.stat.gov.kz>