

Городские эмоции и цифровое участие: моделирование мотивации к использованию городских сервисов на основе анализа тональности постов и комментариев в социальных сетях

А.В. Чижик

Аннотация— В статье рассматриваются возможности анализа социального настроения горожан через тональность постов и комментариев в социальных сетях с целью дальнейшего анализа успешности функционирования городской среды, а также разработки модели мотивации к использованию цифровых городских сервисов. В статье описан разработанный метод оценки тональности текстов из социальных сетей с учётом их контекстуального содержания. Исследование демонстрирует, что тематическое моделирование в сочетании с анализом тональности хорошо подходит для данной задачи, создавая взаимосвязь «эмоции — социальная сфера — городской район». Однако значительной сложностью остаётся точное определение эмоций, поскольку тексты из социальных сетей содержат разговорные выражения, сарказм и культурные смещения. Это указывает на необходимость учёта контекста высказываний жителей для правильной оценки тональности их сообщений. Результатом исследования является метод и разработанная формула, которая позволяет выявлять эмоции на основе ряда факторов и учитывать контекст. Исследование помогло выявить взаимосвязь социального настроения и склонности горожан использовать возможности онлайн-коммуникации с городскими службами. В статье описывается созданная модель мотивации к использованию цифровых городских сервисов.

Ключевые слова— социальное настроение, анализ тональности, тематическое моделирование, языковые модели, модели мотивации.

I. ВВЕДЕНИЕ

В условиях современного мегаполиса цифровое участие становится критически важным компонентом взаимодействия жителей с социальной инфраструктурой города. В рамках данной статьи предлагается понимать под цифровым участием перевод части повседневной

активности горожан в онлайн-пространство. Это подразумевает, что поиск информации о городских услугах, взаимодействие с администрацией и интеракции с социальной инфраструктурой города могут и должны протекать в онлайн-режиме. Помимо того, что такой подход к взаимодействию между городом и горожанами удобен для жителей города, а также предполагает разгрузку городских сервисов в режиме оффлайн, этот процесс необходим для эффективного управления на основе данных, который является целью трендов на цифровую трансформацию [1]. В частности, это дает доступ к информации о нуждах жителей, что позволяет муниципальным властям принимать более обоснованные решения и оперативно реагировать на запросы общества.

Отметим, что идея цифровой трансформации городов на данный момент чаще всего опирается на логику создания экосистем цифровых городских сервисов, которые объединяют государственные и коммерческие структуры в единую платформу, доступную для граждан. Так, например, в Москве действует проект «Активный гражданин», который позволяет жителям участвовать в обсуждении городских инициатив, сообщать о проблемах и голосовать за варианты благоустройства районов [2]. Международный опыт также развивается в этом ключе. Например, в Шанхае (Китай) внедрена система цифровой интеграции через платформу Citizen Cloud, предоставляющая доступ к более чем 1 200 государственным сервисам, таким как медицинское обслуживание и социальное обеспечение [3]. Эта система является примером централизованного подхода к цифровому управлению, который помогает горожанам решать повседневные задачи и взаимодействовать с городской инфраструктурой онлайн. Статистика за 2023 год фиксирует, что ежедневно через эту платформу обрабатываются около 75 000 запросов, что снижает нагрузку на инфраструктуру региона.

Также стоит упомянуть опыт Вильнюса (Литва), где успешно осуществляется цифровая трансформация системы городских услуг (сервис Vilnius e-Services System [4]). Один из ключевых проектов касался модернизации системы электронных услуг муниципалитета, который помог улучшить навигацию и

□Статья получена 28 октября 2024 г.

Исследование выполнено за счет гранта Российского научного фонда и Санкт-Петербургского научного фонда № 23-28-10069 «Прогнозирование социального самочувствия с целью оптимизации функционирования экосистемы городских цифровых сервисов Санкт-Петербурга» (<https://rscf.ru/project/23-28-10069/>).

Статья подготовлена по итогам выступления на конференции «Интернет и современное общество», IMS-2024.

А.В. Чижик, Центр технологий электронного правительства Института дизайна и урбанистики Университета ИТМО (afrancuzova@mail.ru)

доступ к социальным услугам для граждан, особый фокус был сделан на пожилых людей. Команда разработчиков и социальных работников провела интервью с пользователями, чтобы выяснить их потребности и откорректировать цифровую платформу под потребности жителей города. На данный момент система объединяет функции по заказу и мониторингу различных услуг, касающихся социальной инфраструктуры города. Для горожан доступна оценка предоставляемых услуг, управление заявками и мониторинг их выполнения.

Отметим, что в системах городских цифровых услуг многих стран функции не ограничиваются только интерактивными возможностями, такими как подача заявок или оформление документов. Существенная часть сервисов носит информационный характер: они позволяют жителям получать проверенные данные от муниципалитетов, включая сведения о событиях, изменениях в расписании транспорта, доступных социальных услугах. Такие информационные сервисы ценны тем, что предоставляют жителям доступ к систематизированной, достоверной и актуальной информации, что повышает доверие и вовлечённость жителей в динамику городского пространства. К тому же интеграция такого рода сервисов делает подобные платформы не просто средством взаимодействия, но и мощным каналом информирования.

Несмотря на видимое удобство использования таких систем цифровых сервисов уровень вовлеченности горожан в онлайн-взаимодействие с городами пока остается под вопросом и требует изучения. В ряде исследований зафиксировано [5; 6], что проекты умных городов сталкиваются с низким уровнем мотивации горожан к участию в развитии подобных экосистем, что усложняет внедрение цифровых решений в качестве основного средства коммуникации между жителями и социальными инфраструктурами города.

Стоит также отметить, что в ряде научных статей подчеркивается, что концепции умного города часто не учитывают реальные нужды и интересы граждан, а сосредоточена на внедрении технологий без продуманного подхода к их адаптации для повседневных потребностей жителей [7].

Перечисленные проблемы демонстрируют необходимость глубокого исследования и анализа нескольких аспектов. Потребности жителей в городских сервисах (в том числе цифровых) напрямую связаны с социальными ролями, которые соотносятся с индивидом (ребенок, студент, родитель, работник и т.п.). Соответственно в идеальном случае сервисы полностью покрывают все социальные сферы, возникающие на фоне ролей, в которых находится горожанин, и чем проще индивиду удастся воспользоваться ими, тем более позитивное восприятие городской среды у него будет. Как следствие, социальное настроение горожан может являться критерием уровня удовлетворенности взаимодействием с городскими сервисами, так как индивидуальные настроения жителей города в конечном итоге складываются в общий эмоциональный фон. Стоит отметить, что социальное настроение имеет

смысл декомпозировать по двум критериям: 1. социальная сфера, с которой соотносится эмоциональный фон; 2. район (или любой другой вариант деления города на сектора), к которому относится эмоциональное проявление социальной группы. Целесообразность выделения конкретных сфер и локаций связано с тем, что, как правило, городская среда неоднородна, а, значит, прямая фиксация социального настроения всего города (как общей социальной группы, выделенной на фоне принадлежности к региону) не имеет смысла.

Итак, актуализируется задача детекции маркеров социального настроения, так как на фоне их анализа может быть выяснен вопрос успешности функционирования социальных инфраструктур города. При этом представляется, что эти маркеры стоит выявлять через мониторинг публичных дискуссий в социальных сетях, так как это дает возможность собрать объемное количество данных, а также структурировать их по двум важным параметрам – районная принадлежность и социальная сфера. К тому же такой подход подразумевает, что индивиды находятся в естественном контексте взаимодействия со своей социальной группой (обитатели одного района), таким образом становится возможным детектировать коллективное настроение, что больше соответствует цели, чем возможности классических методов сбора данных об эмоциональном фоне жителей (опросы, касающиеся социального самочувствия).

Успешность функционирования городской среды проявляется в наличии возможности у жителей вступить в коммуникацию с городской администрацией на разных уровнях и получить неформальную обратную связь. Таким образом выявление векторов социального настроения может быть полезно для поиска взаимосвязи между эмоциональным фоном горожан и доступностью (и оптимальностью) онлайн-коммуникации с городом при решении своих повседневных задач, связанных с социальными ролями индивида. Таким образом в конечном итоге актуализируется необходимость анализа зон проседающей коммуникации между жителями и городскими сервисами, так как проблема прерывающегося коммуникативного акта может быть решена посредством юстировки онлайн-пространства взаимодействия жителей с городом.

II. ОНЛАЙН-КОММУНИКАЦИИ КАК ОСНОВА ДЕТЕКЦИИ ЭМОЦИОНАЛЬНОГО ФОНА ГОРОДА

Динамика урбанизации приводит к усложнению социальной, экономической и экологической структуры городской среды, поэтому важной частью управления городом становится мониторинг эмоционального состояния горожан, который может обеспечить понимание восприятия городской среды жителями, поскольку социальное настроение является одним из основных факторов, воздействующих на самоощущение индивида и, как следствие, влияющий на его мнение о качестве жизни в городской среде (маркеры социального самочувствия). Для оценки социального самочувствия населения традиционно используются

классические социологические методы (например, опросы и фокус-группы) [8]. Они позволяют выявлять удовлетворенность граждан различными аспектами городской жизни: от качества медицинского обслуживания до доступности общественного транспорта и уровня безопасности [9]. Такие данные часто становятся основой для понимания того, какие социальные сферы испытывают дефицит ресурсов или поддержки со стороны городской администрации, и где проявляется наибольшее недовольство среди жителей [10;11]. Например, в статье [12] демонстрируется, что с помощью опросов удалось выявить недовольство жителей качеством социальных услуг муниципалитетов, так как зачастую горожане сталкиваются с отсутствием четкой и оперативной обратной связи. Эти проблемы обозначены и в более свежих результатах опросов жителей крупных городов [13;14].

Однако стоит подчеркнуть, что несмотря на привлекательность и репрезентативность такого рода методик проведения исследования степени удовлетворенности жителей городской средой, в данном случае речь всегда идет о зафиксированном случившемся событии, то есть о контексте, который имеет отношение к категории прошедшего времени. В то же время для эффективного управления городской средой необходимо быть в контексте настоящего времени, тогда взаимодействие с жителями может быть выстроено наиболее гармонично, в том числе и в зоне коммуникативного взаимодействия с ними (как было указано выше, это важный аспект, и часто жители нуждаются в первую очередь в обратной связи).

Поиск новых методов мониторинга эмоционального фона горожан приводит к методам обработки естественного языка, что позволяет выявлять паттерны общественных настроений и структуры социокультурных представлений, обратившись к текстовым данным, находящимся в публичных сегментах интернета. Текстовые данные, размещенные в интернете, представляют собой важный источник информации о формировании образа мира в массовом сознании. Источниками таких данных являются как традиционные средства массовой информации (СМИ), так и социальные сети. В СМИ тексты создаются журналистами, чьи материалы отражают как индивидуальные взгляды, так и определенный идеологический вектор, направленный на формирование общественного мнения. В социальных сетях пользователи выражают свои личные мнения, вступают в дискуссии, тем самым создавая децентрализованное и многоаспектное поле общественного взаимодействия. Для детекции эмоционального фона жителей города больше подходит поиск и сбор данных из социальной сети. Эти данные могут быть собраны, обработаны и проанализированы с использованием методов автоматического анализа текстов, что позволяет декомпозировать дискуссии по социальным сферам (тематическое моделирование) и провести анализ тональности.

Анализ текстовых данных (nlp) – непростая задача, требующая выбор актуальных средств автоматической

обработки и анализа естественного языка. Сравнительно долго основные методы анализа текстовых данных были связаны с выявлением именованных сущностей и детекцией устойчивых коллокаций (реже – биграмм). Надо отметить, что такие подходы в некоторых случаях оправданы до сих пор (и дают неплохие результаты), однако на данный момент анализ текста часто опирается на машинное обучение. При этом стоит отметить, что, например, поиск частотных слов и/или биграмм используются как вспомогательные методы на этапе предобработки текстовых данных. Если же смотреть онтологично на анализ текстовых данных с использованием алгоритмов машинного обучения, то можно выделить две стратегии: обучение с учителем (по сути, задача классификации данных), когда человек помогает обучать модель (ручная разметка данных), и обучение без учителя, когда компьютер находит закономерности в тексте с минимальным вмешательством человека (вмешательство сводится к указанию параметров, например, сколько кластеров модель должна попытаться найти).

Обучение без учителя практически для любых данных подразумевает классические алгоритмы кластеризации. Кластеризация — это процесс, используемый для разделения объектов на группы, где кластеры-группы, содержат внутри себя схожие между собой объекты. То есть все объекты, выделенные в один кластер, более похожи друг на друга, чем на объекты вне его. Суть алгоритмов кластеризации состоит в интерпретации параметров, составляющих каждый элемент набора данных, детектированные моделью признаки помогают сгруппировать элементы, чем более однозначные отличия между выделенными признаками, тем качественнее произойдет разделение (часто данные настолько похожи друг на друга, что разделение на кластеры может быть условным). Для текстовых данных кластеризацию можно использовать, однако часто результаты оказываются нерепрезентативными, так как естественный язык имеет множество одинаковых по словарному составу конструкций с различными смысловыми интерпретациями, это означает, что данные в векторном пространстве (любой анализ текста с использованием машинного обучения подразумевает перевод текста в векторное пространство) будут находиться близко, и стандартный алгоритм кластеризации воспримет их за элементы одного кластера. Поэтому для анализа текста чаще используют тематическое моделирование, которое осуществляется с помощью статистических алгоритмов. На сегодняшний день существует 4 наиболее распространенных и популярных метода: LSA, pLSA, LDA и более нового метода lda2vec, основанного на глубоком обучении. В целом все тематические модели основаны на следующих двух предположениях: 1. каждый документ состоит из смеси тем, и 2. каждая тема состоит из набора слов. Есть несколько успешных примеров использования тематического моделирования для анализа данных, полученных из социальных сетей. Например, Graph Neural Network (GCN) использовалась для исследования влияния виртуального социального

общения на предпочтения пользователей [15; 16]. Также стоит упомянуть исследование, проведенное в Сербии с целью выявить причины отказа от вакцинации от COVID-19 [17]. Тематическое моделирование было применено к постам на сербском языке, использовались модели LDA и NMF для классификации текстов с негативной коннотацией, включая темы о побочных эффектах, недоверии к властям и теориях заговора. Для повышения точности применялась модель BERT для классификации и фильтрации данных, что позволило глубже понять доминирующие тревоги населения и разработать точные рекомендации для муниципального управления в Сербии.

Анализ тональности текстов также является популярной задачей применения nlp к текстам, собранным из социальных медиа. Так, например, анализ настроений с использованием нейронных сетей стал предметом нескольких крупных исследований, в том числе [18]. За основу были взяты шесть базовых эмоций [19] (гнев, радость, удивление, страх, отвращение и печаль), которые с использованием моделей оценки настроений детектировались в текстах пользователей социальных сетей. Выявленный эмоциональный фон использовался для предсказания поведения социальной группы.

В исследовании [20], посвящённом анализу постов на тему экологии и изменения климата в социальных сетях (Reddit и др.) применялись модели тематического моделирования и сентимент-анализ для понимания общественных настроений. Подход включал использование Pointwise Mutual Information (PMI) для оценки эмоционального окраса текста. Методология включала сбор данных, предобработку, обучение модели и тестирование на новых данных, что позволило эффективно анализировать настроения, выраженные в тысячах сообщений пользователей.

В последние годы появился еще один подход к анализу эмоционального состояния индивидов с целью последующего прогнозирования и/или интерпретации социального настроения. Он заключается в том, что базовые эмоции необходимо связывать с набором ключевых слов, отражающих контекст темы. Таким образом, используется объединение тематического моделирования и анализа тональности текстов [21].

Перечисленные исследования и подходы показывают важность как классических социологических методов, так и применения nlp и мониторинга социальных сетей. Логичным представляется проводить анализ социальных сетей, структурируя собранные массивы в логику «район-социальная сфера-эмоция». Таким образом появляется возможность декомпозировать город по локациям и анализировать эмоциональный фон горожан, соотносящийся с ними, а также социальные сферы, к которым относятся эмоции. Роль же классических опросов видится в уточнении причин затруднений, с которыми сталкиваются индивиды при взаимодействии с социальными инфраструктурами города в тех сферах, на которые указал анализ текстовых данных.

III. АНАЛИЗ ПОСТОВ И КОММЕНТАРИЕВ ЖИТЕЛЕЙ В СОЦИАЛЬНЫХ СЕТЯХ С УЧЕТОМ КОНТЕКСТА ВЫСКАЗЫВАНИЙ

Анализируя эмоциональные полярности дискуссий в социальных сетях, можно выявлять общие конструкции культурных смещений и векторы эмоциональных восприятий, которые в последствии становятся базисом мировоззрения общества. Таким образом, публичные тексты из социальных сетей могут быть полезны в том числе для оценки общественно-политической обстановки, уровня функционирования социальной сферы, социального самочувствия, а также для анализа удовлетворенности жителей городской средой. Фокусируя внимание на последнем, важно отметить, что понимание эмоциональных реакций жителей на различные аспекты городской жизни может помочь муниципалитетам и городским планировщикам принимать более обоснованные решения. Например, если анализ показывает повышенные уровни недовольства или стресса в определённых районах, это может служить сигналом к необходимости улучшения городской инфраструктуры, общественного транспорта, или жилищных условий в этих районах.

Значит, возникает две актуальные сущности «район-эмоция» и ее производная – «район-эмоция-социальная сфера». Важно отметить, что в случае изучения текстов из социальных сетей речь идет о латентном взаимодействии с индивидом, что подразумевает наблюдение индивида с его ориентацией на социальную группу, к которой он принадлежит в моменте дискуссии в социальной сети (что отличает этот метод от классических опросов о социальном самочувствии, когда индивид, получая вопрос, находится в анонимном контексте и отвечает скорее со своих личных позиций, без эффекта прямой коммутации со внешней средой). Таким образом, в качестве результата анализа текстов появляется возможность оценить общую эмоцию социальной группы: либо как категорию, описывающую общий эмоциональный фон района; либо с привязкой к социальной роли.

Итак, анализ настроений микротекстов (постов и комментариев к ним) становится ключевым инструментом для выявления трендов присутствующих в обществе настроений, это крайне важно для анализа городской среды. Однако анализ тональности таких текстов представляет собой сложную задачу из-за их краткости, неоднозначности и контекстуальной загруженности. В данном исследовании была поставлена задача найти такой способ анализа тональности микротекстов из социальных сетей, который бы учитывал контекст высказывания, на основании чего через соотношение фактов и эмоциональной части выносилось суждение про содержащуюся в тексте эмоцию.

Появление глубокого обучения значительно продвинуло область анализа настроений в сторону репрезентативности выводов. Такие модели, как BERT [22] и GPT [23] были успешно адаптированы для понимания нюансов языка, используемого в социальных сетях. Эти модели способны улавливать контекстные

нюансы, которые необходимы для точной интерпретации настроений. Например, в статье [24] продемонстрировано, как можно тонко настроить BERT для эффективного определения полярности настроений с учетом контекста, превзойдя традиционные модели на контрольных наборах данных.

Действительно, эти модели хорошо улавливают контекст, что улучшает их способность распознавать настроения в сложных ситуациях. Однако их способность связывать фактическую информацию с эмоциональным тоном небезупречна. Дополнительные проблемы создают такие языковые особенности, как сарказм и метафоры, когда выраженное настроение может резко противоречить буквальному значению слов. Например, саркастичное высказывание может быть формально положительным, но по сути подразумевать негатив, и только контекст и его правильное применение к дешифровке фразы может помочь правильно отнести текст к негативным (или наоборот).

Проблемы с возможностью «схватить» контекст короткого текста (факты, на основании которых пользователь пишет комментарий или пост) подразумевают, что подход к анализу тональности текста в социальных медиа должен быть основан на алгоритме, который бы объединял в себе несколько этапов обработки данных. Наши исследования направлены на тестирование оптимальности создаваемого пайплайна. В текущей реализации алгоритм состоит из следующих этапов анализа текстов: 1. предобработка текста (входной текст проходит предварительную обработку по стандартной схеме, включающей токенизацию, удаление стоп-слов и нормализацию текста); 2. выделение фактов (контекста) и эмоциональной составляющей; 3. вычисление соотношения контекста к эмоциональной части (FR), полученное число используется как коэффициент в итоговой формуле для оценки эмоции; 4. анализ эмоций с помощью Word2Vec (модель Word2Vec используется для выявления эмоциональной окраски текста на уровне отдельных слов, что позволяет учесть семантическое сходство слов с анкорными словами, характеризующими различные эмоциональные состояния.); 5. оценка тональности текста с использованием BERT (применяется для бинарной классификации тональности текста на позитивную или негативную, что дает возможность получить общее представление о тональности текста); 6. оценка эмоций с помощью ruGPT-3 (для получения вероятностей принадлежности текста к каждой из пяти категорий эмоций); 7. итоговая оценка эмоции (формула учитывает FR, W2V Score, Sentiment Score и GPT Score и определяет преобладающую эмоциональную составляющую текста).

Таким образом подход основан на учете как эмоциональной составляющей текста, так и его контекста при определении итоговой эмоции; формула для подсчета итоговой эмоции учитывает несколько ключевых факторов, позволяя более точно определить эмоциональную окраску текста. Итак, для

количественной оценки эмоциональной окраски текста используется формула, комбинирующая несколько компонентов:

1) Соотношение фактов к эмоциям (FR). Этот параметр определяет, насколько сильно содержание текста влияет на его эмоциональную окраску. Чем выше значение FR, тем больший вес имеют факты в определении итоговой эмоции.

2) Оценка эмоций с помощью Word2Vec (W2V Score). Модель Word2Vec используется для выявления эмоциональной окраски текста на уровне отдельных слов. W2V Score представляет собой сумму оценок эмоциональной значимости слов в тексте, что позволяет учесть эмоциональные контексты на уровне слов.

3) Оценка тональности текста с помощью Sentiment Score. В формуле учитывается оценка общей тональности текста с помощью модели классификации тональности (в текущей реализации используется BERT). Sentiment Score представляет числовую оценку, отражающую общую эмоциональную окраску текста, что позволяет корректировать итоговую эмоцию в соответствии с общим эмоциональным контекстом.

4) Оценка эмоций с помощью ruGPT-3 (GPT Score). Модель генерации текста (в текущей реализации используется ruGPT-3) дает возможность оценки эмоциональной окраски текста на более высоком уровне семантики. GPT Score представляет собой сумму вероятностей различных эмоций, предсказанных моделью, что дополняет оценку, предоставляемую Word2Vec и Sentiment Score.

Итоговая формула выглядит следующим образом:

$$\text{Final Emotion Score} = (\text{FR} \times \text{W2V Score} \times \text{Sentiment Score}) - (\text{GPT Score})$$

Эта формула учитывает как объективные, так и субъективные аспекты текста, комбинируя фактическое содержание с эмоциональной и тональной окраской. Она учитывает как конкретные слова и их эмоциональное значение, так и общий контекст и тональность текста. Такой комплексный подход позволяет получить более полное представление о эмоциональной составляющей текста и улучшить точность его анализа.

Итоговый вывод программы представляет собой комбинированную оценку тональности текста по пятибальной шкале с использованием двух весовых коэффициентов. Сочетание двух нейросетевых моделей в аналитическом пайплайне позволяет не только точно классифицировать текст по эмоциональной окраске, но и учитывать фактическую информацию, что делает анализ более ориентированным на контекст высказывания.

Стоит отметить, что для GPT Score использовалась разработанная модель базовых эмоций [25], основанная на исследовании [26].

Для того, чтобы протестировать созданный пайплайн оценки тональности текстов был собран датасет публичных постов и комментариев к ним из социальной сети ВКонтакте (были выбраны сообщества, имеющие привязку к району). Поиск групп, из которых брались текстовые данные проводился вручную: каждый район

отдельно изучался с точки зрения присутствующей социальной активности, далее группы заносились в отдельный датасет, с присваиванием маркера, к какой социальной роли относится паблик. Далее каждая группа, вошедшая в выборку, анализировалась на предмет публикационной активности, а также на динамику вербальной и невербальной реакции участников сообщества на эти посты. Таким образом был произведен отсев групп, с низким уровнем социального взаимодействия состоящих в них пользователей. Это было сделано из соображений отсева шума, который мог помешать дальнейшему анализу. В итоговую выборку вошли 412 публичных групп, в среднем по 25 групп на каждый район. Численность подписчиков в среднем варьировалась от 10 тыс. до 80 тыс. (в зависимости от тематической направленности группы); посты в выбранных группах появляются с частотой 1-2 раза в сутки, динамика комментариев к ним варьируется от 3 до 20 (в среднем). Мы собирали данные за 2018-2023 годы. Средняя длина комментариев в итоговой выборке — 525 слов (на пост оставляют 10–15 комментариев); средняя длина постов — 41 слово.

Собранные текстовые данные были проанализированы с использованием тематического моделирования (для чего был разработан метод, выделяющий признаки социальных ролей в комментариях пользователей), затем посты и комментарии к ним исследовались на предмет наличия эмоциональной окраски комментариев (разработан подход к оценке общего эмоционального градуса поста, описано ниже). Выделение принадлежности текстов к социальным ролям позволило проанализировать не только общий эмоциональный фон, соотносящийся с локациями районов, но также посмотреть возникающие взаимосвязи тем, на которые пишут пользователи, и эмоций, которые они испытывают при этом. Таким образом возникла результирующая сущность «социальная сфера – эмоция – район».

Для оценки качества алгоритма анализа настроений в коротких текстах социальных сетей, использовалась метрика F1-score. Эта метрика является гармоническим средним между точностью и полнотой и позволяет оценить сбалансированность алгоритма, учитывая как ошибки false positives, так и false negatives. Итоговое значение F1-score = 0.86, что указывает на высокую точность и полноту классификации.

IV. СКЛОННОСТЬ ГОРОЖАН К ИСПОЛЬЗОВАНИЮ ИНТЕРНЕТ-КОММУНИКАЦИИ

В период с сентября 2023 по август 2024 года были проведены две волны опросов жителей Санкт-Петербурга (N=1000), направленные на выявление маркеров социального самочувствия и анализ

вовлеченности горожан во взаимодействие с городской инфраструктурой посредством интернет-коммуникаций. Исследование было сосредоточено на оценке того, как жители города воспринимают городскую среду, есть ли у них доступ к цифровым сервисам и насколько они удовлетворены качеством взаимодействия с административными структурами через онлайн-платформы (или предпочитают использовать оффлайн-коммуникации).

Данные о маркерах социального самочувствия были сравнены с результатами анализа текстов, собранных в социальной сети.

В большинстве случаев зафиксированное в собранных из социальной сети текстовых данных недовольство горожан было связано с невозможностью вступить в коммуникацию с той или иной административной городской структурой. Это проявляется в затруднениях, с которыми сталкиваются жители при попытке зарегистрировать запросы к городским властям, получить необходимую информацию на официальных ресурсах и оставить обратную связь. Такой разрыв в коммуникации влияет на уровень раздражения индивида, и именно с этой эмоцией он «приходит» в беседу к своей социальной группе (жители района, работники района и т. п.), постепенно распространяя свое настроение на всех участников дискуссии (эффект циркулярной реакции [27]).

Данные опросов показали, что жители нескольких районов активно используют систему городских цифровых сервисов в своих повседневных практиках, именно в этих районах зафиксирована наиболее позитивная палитра эмоций. Таким образом прослеживается прямая корреляция между налаженным взаимодействием города и жителей через цифровые сервисы и социальным настроением, а также актуализируется вопрос выстраивания моделей мотивации, которые бы привлекали остальных жителей в онлайн-пространство.

V. МОДЕЛИ МОТИВАЦИИ К ИСПОЛЬЗОВАНИЮ ГОРОДСКИХ ЦИФРОВЫХ СЕРВИСОВ

Выявленные в ходе исследования закономерности позволили разработать мотивационную модель, которая обобщает ключевые факторы, определяющие желание использовать цифровые платформы как способ коммуникации с администрацией города. Логика модели строится на взаимодействии трех компонентов: доступности и удобства сервисов, позитивного эмоционального и социального опыта пользователей при их использовании, а также предсказательной мотивации, которая позволяет городским службам адаптировать свои услуги для повышения удовлетворенности граждан.

Итак, предлагаемая модель мотивации (рис.1) к использованию городских цифровых сервисов основывается на трех ключевых элементах.

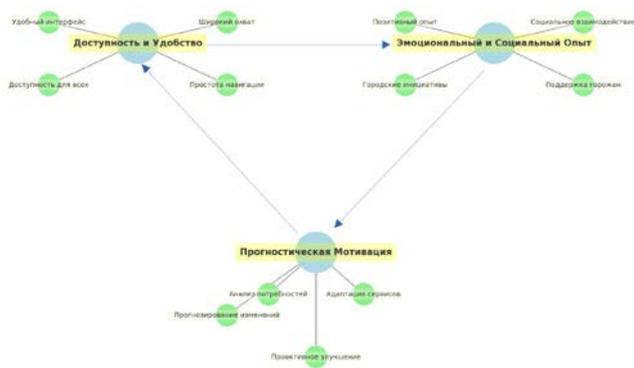


Рис. 2. Модель мотивации к использованию городских цифровых сервисов

Доступность и удобство (Access and Convenience). Ключевыми факторами являются простота доступа к сервисам и удобство их использования. Чем проще сервисы в использовании, тем больше вероятность, что жители будут ими активно пользоваться. Это включает в себя удобные интерфейсы, широкое распространение сервисов по разным районам города и обеспечение их доступности для всех слоев населения.

Эмоциональный и социальный опыт (Emotional and Social Experience). Положительные эмоции и социальная вовлеченность играют важную роль в мотивации. Жители, которые получают положительный опыт взаимодействия с городскими сервисами (например, оперативность и эффективность решения их вопросов), более склонны продолжать их использование. Социальная вовлеченность, возникающая через взаимодействие с другими жителями и участие в городских инициативах, также усиливает мотивацию.

Прогностическая мотивация (Predictive Motivation). Прогнозирование потребностей пользователей с использованием методов анализа данных. Когда городские сервисы адаптируются к изменяющимся потребностям жителей, это повышает их удовлетворенность и укрепляет мотивацию к дальнейшему использованию сервисов. Предиктивные модели помогают выявлять будущие запросы и потребности, что способствует проактивному улучшению сервисов.

Разработанная модель поддерживает целостное управление мотивацией пользователей, подстраивая городской сервис под реальные потребности и интересы жителей.

VI. ЗАКЛЮЧЕНИЕ

Предложенный метод анализа тональности текстов из социальных сетей продемонстрировал эффективность при определении эмоционального тона в дискуссиях, касающихся городской среды. Метрики качества показывают, что формула является перспективной для использования в заявленных целях.

В рамках исследования была проанализирована социальная сеть ВКонтакте с целью мониторинга эмоционального фона горожан. Также было проведено две волны опроса жителей Санкт-Петербурга, которые уточнили маркеры социального самочувствия, а также

подтвердили гипотезу о корреляции социального настроения и успешности коммуникативного взаимодействия городских администраций с жителями. Был сделан вывод о том, что онлайн-среда является перспективным пространством для выстраивания непрерывного взаимодействия с горожанами. Итогом исследования стало создание мотивационной модели к использованию экосистемы городских сервисов. Предложенная мотивационная модель служит основой для выстраивания эффективного контура взаимодействия между горожанами и этими сервисами. На её основе можно разрабатывать стратегии, которые будут учитывать три ключевых компонента: доступность и удобство сервисов, эмоциональный и социальный опыт пользователей, а также предсказание потребностей горожан. Разработанная модель позволяет городским администрациям не просто устранять текущие недостатки, такие как ограниченный доступ к сервисам или негативный пользовательский опыт, но и проактивно создавать условия для повышения вовлеченности граждан. Благодаря прогнозированию потребностей можно заранее адаптировать цифровые платформы под изменяющиеся запросы населения, что позволит минимизировать риски низкой удовлетворенности. Таким образом, модель становится инструментом для выравнивания уровня доступности и качества сервисов по всему городу.

БЛАГОДАРНОСТИ

Исследование выполнено за счет гранта Российского научного фонда и Санкт-Петербургского научного фонда № 23-28-10069 «Прогнозирование социального самочувствия с целью оптимизации функционирования экосистемы городских цифровых сервисов Санкт-Петербурга» (<https://rscf.ru/project/23-28-10069/>).

БИБЛИОГРАФИЯ

- [1] Стратегия цифровой экономики Российской Федерации на 2017-2030 годы: утверждена протоколом заседания президиума Совета при Президенте Российской Федерации по стратегическому развитию и национальным проектам от 4 июня 2019 г. URL: <http://government.ru/docs/36458/> (дата обращения: 26.10.2024).
- [2] Estecahandy H. The Democratic Illusion through the Technological Illusion: a Case Study of the Implementation of a Blockchain to Support an E-voting Platform in Moscow (Active Citizen) // arXiv.org. 2023. URL: <https://arxiv.org/abs/2301.03954> (дата обращения: 26.10.2024).
- [3] Qinghong Cui, Run Chen, Ruirui Wei, Xiancun Hu, Guangbin Wang. Smart Mega-City Development in Practice: A Case of Shanghai, China // MDPI. 2023. DOI: <https://doi.org/10.3390/su15021591>.
- [4] Civitta. Vilnius Municipality — Modernising E-Services // Civitta. URL: <https://civitta.com> (дата обращения: 26.10.2024).
- [5] Gartner. 59 AI customer service statistics for 2024 // Zendesk. 2024. URL: <https://www.zendesk.com/blog/ai-customer-service-statistics/> (дата обращения: 26.10.2024).
- [6] Gartner. Top Customer Service Trends and Priorities To Watch In 2024 // Gartner. 2024. URL: <https://www.gartner.com/en/articles/top-customer-service-trends-and-priorities-to-watch-in-2024> (дата обращения: 26.10.2024).
- [7] Oliveira T. A., Oliver M., Ramalinho H. Challenges for Connecting Citizens and Smart Cities: ICT, E-Governance and Blockchain // Sustainability. 2020. T. 12, № 7. С. 2926. DOI: <https://doi.org/10.3390/su12072926>.
- [8] Lin Y. R. et al. Voices of victory: A computational focus group framework for tracking opinion shift in real time // Proceedings of the 22nd international conference on World Wide Web. 2013. P. 737-748.

- [9] Цветкова И.В. Факторы социального самочувствия горожан // Карельский научный журнал. Социальные науки. 2017. Т. 6. № 1(18). С. 113-117.
- [10] Реутов Е. В. Городская среда как фактор социального самочувствия населения и риски, сопутствующие её изменению // Управление городом: теория и практика. 2021. С. 51-54.
- [11] Ochoa Rico M. S. et al. Study of citizen satisfaction and loyalty in the urban area of Guayaquil: Perspective of the quality of public services applying structural equations // PloS one. 2022. Vol. 17. No. 2.
- [12] Prahara S., Han J. H., Hawken S. Innovative civic engagement and digital urban infrastructure: Lessons from 100 smart cities mission in India // Procedia Engineering. 2017. Vol. 180. P. 1423-1432.
- [13] Secinaro S. et al. Does citizen involvement feed on digital platforms? // International Journal of Public Administration. 2022. Vol. 45. No. 9. P. 708-725.
- [14] Mwangi S. C. A survey of civic engagement tools and capabilities of city and county government web sites // Online Journal of Communication and Media Technologies. 2021. Vol. 11. No. 4.
- [15] Tang J. et al. Social Influence Analysis in Large-Scale Networks // Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2012. P. 807–815. DOI: 10.1145/2339530.2339636.
- [16] Peng H. et al. Social Recommendation Using Graph Neural Networks // IEEE Transactions on Neural Networks and Learning Systems. 2016. Vol. 27. No. 8. P. 2107-2119. DOI: 10.1109/TNNLS.2016.2518933.
- [17] Ляйич, А., Проданович, Н., Медвецки, Д., Башарагин, Б., & Митрович Е. Uncovering the Reasons Behind COVID-19 Vaccine Hesitancy in Serbia: Sentiment-Based Topic Modeling // Journal of Medical Internet Research. 2022. No. 24(11). e42261. DOI: <https://doi.org/10.2196/42261>.
- [18] Li C., Ma J., Guo X., Mei Q. Sentiment Analysis of Weibo Texts Based on Deep Learning // IEEE Transactions on Computational Social Systems. 2020. Vol. 7. No. 1. P. 24–35. DOI: 10.1109/TCSS.2020.2964923
- [19] Ekman P. Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life. New York: Holt, 2007.
- [20] Hanny D., Resch B. Understanding Environmental Posts: Sentiment and Emotion Analysis of Social Media Data // IEEE Journals & Magazine. 2023. URL: <https://ieeexplore.ieee.org/document/10453566> (дата обращения: 27.10.2024).
- [21] Liu B. Sentiment analysis and opinion mining. Springer Nature, 2022.
- [22] Devlin J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding // arXiv preprint. 2018. arXiv:1810.04805.
- [23] Radford A. et al. Improving language understanding by generative pre-training. 2018.
- [24] Sun C. et al. How to fine-tune bert for text classification? // Chinese computational linguistics: 18th China national conference, CCL 2019, proceedings 18. Springer International Publishing, 2019. P. 194-206.
- [25] Чижик А. В. Исследование динамики общественного настроения в социальных сетях с использованием методов тематического моделирования // International Journal of Open Information Technologies. 2021. Т. 9, №. 12. С. 21–29.
- [26] Ekman P. Methods for measuring facial action // Handbook of methods in nonverbal behavior research. 1982. P. 45-90.
- [27] Чижик А. В. Социокультурный потенциал социальных медиа: создание пространства общего опыта через эмоциональное взаимодействие членов сети // Новые информационные технологии в автоматизированных системах. 2019. №. 22. С. 130-134.

Чижик Анна Владимировна, к. культ., старший научный сотрудник Центра технологий электронного правительства Института дизайна и урбанистики, Университет ИТМО (<http://itmo.ru/>), Санкт-Петербург, email: afrancuzova@mail.ru, elibrary.ru: authorid=708001, scopus.com: authorId=57222136821, ORCID: orcidID=0000-0002-4523-5167

Urban Emotions and Digital Engagement: Modeling Motivation for Using Urban Services Based on Sentiment Analysis of Social Media Posts and Comment

Anna V. Chizhik

Abstract— The article explores the possibilities of analyzing the social sentiment of urban residents through the tone of posts and comments on social media, with the aim of further evaluating the effectiveness of urban environment functionality and developing a motivation model for using digital urban services. The paper describes a developed method for assessing the sentiment of texts from social networks, taking into account their contextual content. The study demonstrates that topic modeling combined with sentiment analysis is well-suited for this task, establishing the relationship between "emotions — social sphere — urban area." However, accurately identifying emotions remains a significant challenge, as social media texts often contain colloquial expressions, sarcasm, and cultural nuances. This highlights the necessity of considering the context of residents' statements to properly evaluate the sentiment of their messages. The outcome of the research is a method and a developed formula that enables the identification of emotions based on various factors while accounting for context. The study revealed a connection between social sentiment and the propensity of urban residents to engage with municipal services via online communication. The article also describes the created motivation model for the use of digital urban services.

Keywords— social sentiment, sentiment analysis, topic modeling, language models, motivation models.

REFERENCES

- [1] *Strategy of the Digital Economy of the Russian Federation for 2017–2030*: approved by the minutes of the meeting of the Presidium of the Presidential Council of the Russian Federation for Strategic Development and National Projects dated June 4, 2019 [Online]. Available: <http://government.ru/docs/36458/>. Accessed: Oct. 26, 2024.
- [2] H. Estecahandy, "The Democratic Illusion through the Technological Illusion: a Case Study of the Implementation of a Blockchain to Support an E-voting Platform in Moscow (Active Citizen)," *arXiv.org*, 2023. Available: <https://arxiv.org/abs/2301.03954>. Accessed: Oct. 26, 2024.
- [3] Q. Cui, R. Chen, R. Wei, X. Hu, and G. Wang, "Smart Mega-City Development in Practice: A Case of Shanghai, China," *MDPI*, 2023. doi: <https://doi.org/10.3390/su15021591>.
- [4] Civitta, "Vilnius Municipality — Modernising E-Services," Civitta, <https://civitta.com>.
- [5] Gartner, "59 AI customer service statistics for 2024," Zendesk, 2024, <https://www.zendesk.com/blog/ai-customer-service-statistics>.
- [6] Gartner, "Top Customer Service Trends and Priorities To Watch In 2024," Gartner, 2024, <https://www.gartner.com/en/articles/top-customer-service-trends-and-priorities-to-watch-in-2024>.
- [7] T. A. Oliveira, M. Oliver, and H. Ramalhinho, "Challenges for Connecting Citizens and Smart Cities: ICT, E-Governance and Blockchain," *Sustainability*, vol. 12, no. 7, p. 2926, 2020. DOI: <https://doi.org/10.3390/su12072926>.
- [8] Y. R. Lin et al., "Voices of victory: A computational focus group framework for tracking opinion shift in real time," in *Proceedings of the 22nd International Conference on World Wide Web*, 2013, pp. 737–748.
- [9] I. V. Tsvetkova, "Factors of social well-being of citizens," *Karelian Scientific Journal, Social Sciences*, vol. 6, no. 1(18), pp. 113–117, 2017.
- [10] E. V. Reutov, "Urban environment as a factor in social well-being of the population and the risks associated with its change," in *City Management: Theory and Practice*, pp. 51–54, 2021.
- [11] M. S. Ochoa Rico et al., "Study of citizen satisfaction and loyalty in the urban area of Guayaquil: Perspective of the quality of public services applying structural equations," *PLoS One*, vol. 17, no. 2, 2022.
- [12] S. Prahara, J. H. Han, and S. Hawken, "Innovative civic engagement and digital urban infrastructure: Lessons from 100 Smart Cities Mission in India," *Procedia Engineering*, vol. 180, pp. 1423–1432, 2017.
- [13] S. Secinaro et al., "Does citizen involvement feed on digital platforms?," *International Journal of Public Administration*, vol. 45, no. 9, pp. 708–725, 2022.
- [14] S. C. Mwangi, "A survey of civic engagement tools and capabilities of city and county government web sites," *Online Journal of Communication and Media Technologies*, vol. 11, no. 4, 2021.
- [15] J. Tang et al., "Social Influence Analysis in Large-Scale Networks," in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012, pp. 807–815. DOI: 10.1145/2339530.2339636.
- [16] H. Peng et al., "Social Recommendation Using Graph Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 2107–2119, 2016. DOI: 10.1109/TNNLS.2016.2518933.
- [17] A. Ljajić, N. Prodanović, D. Medvečki, B. Basharagin, and E. Mitrović, "Uncovering the Reasons Behind COVID-19 Vaccine Hesitancy in Serbia: Sentiment-Based Topic Modeling," *Journal of Medical Internet Research*, vol. 24, no. 11, e42261, 2022. DOI: <https://doi.org/10.2196/42261>.
- [18] C. Li, J. Ma, X. Guo, and Q. Mei, "Sentiment Analysis of Weibo Texts Based on Deep Learning," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 1, pp. 24–35, 2020. DOI: 10.1109/TCSS.2020.2964923.
- [19] P. Ekman, "Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life," New York: Holt, 2007.
- [20] D. Hanny and B. Resch, "Understanding Environmental Posts: Sentiment and Emotion Analysis of Social Media Data," *IEEE Journals & Magazine*, 2023. <https://ieeexplore.ieee.org/document/10453566>.
- [21] B. Liu, *Sentiment Analysis and Opinion Mining*, Springer Nature, 2022.
- [22] J. Devlin et al., "BERT: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint*, arXiv:1810.04805, 2018.
- [23] A. Radford et al., "Improving language understanding by generative pre-training," 2018.
- [24] C. Sun et al., "How to fine-tune BERT for text classification?" in *Chinese Computational Linguistics: 18th China National Conference, CCL 2019, Proceedings 18*, Springer International Publishing, pp. 194–206, 2019.
- [25] A. V. Chizhik, "Study of the dynamics of public mood in social networks using topic modeling methods," *International Journal of Open Information Technologies*, vol. 9, no. 12, pp. 21–29, 2021.
- [26] P. Ekman, "Methods for measuring facial action," *Handbook of Methods in Nonverbal Behavior Research*, pp. 45–90, 1982.

[27] A. V. Chizhik, "Sociocultural potential of social media: Creating a space of common experience through emotional interaction of network members," *New Information Technologies in Automated Systems*, no. 22, pp. 130–134, 2019.

Anna V. Chizhik, Ph.D in Cultural Studies, Senior Researcher of E-Governance Center, Institute of Design and Urban Studies, ITMO University (<http://itmo.ru/>), Saint-Petersburg, email: afrancuzova@mail.ru, elibrary.ru: authorid=708001, scopus.com: authorId=57222136821, ORCID: [orcidID=0000-0002-4523-5167](http://orcid.org/0000-0002-4523-5167).