

НАУЧНЫЙ ОБЗОР

УДК 004.67

DOI 10.17513/snt.39635

**МЕТОДИКА АНАЛИЗА ДАННЫХ О ЧРЕЗВЫЧАЙНЫХ СИТУАЦИЯХ
В СОЦИАЛЬНЫХ СЕТЯХ****Вострых А.В., Максимов А.В., Матвеев А.В., Смирнов А.С.***ФГБОУ ВО «Санкт-Петербургский университет Государственной противопожарной службы
МЧС России имени Героя РФ генерала армии Е.Н. Зиничева», Санкт-Петербург,
e-mail: a.vostrykh@list.ru*

При возникновении чрезвычайных ситуаций важная роль отводится точной и своевременной информации о происшествиях. Это дает возможность лицам, принимающим решения, составлять оперативные и адекватные планы действий в условиях неблагоприятно складывающейся обстановки. В последние годы одним из важных источников данных о происшествиях являются различные социальные сети, пользователи которых обмениваются информацией, выкладывая текстовые, фото- и видеоматериалы. В статье представлена новая оригинальная методика, позволяющая анализировать информацию, опубликованную в социальных сетях, с целью своевременного на нее реагирования. В отличие от существующих решений, предложенная методика учитывает многокритериальность распространения информации в реальном мире, легко масштабируется под любые социальные сети, учитывает взаимоотношения между пользователями в процессе распространения информации о чрезвычайных ситуациях. Также она обладает возможностью как низкоуровневого анализа отдельных пользователей, так и высокоуровневого анализа отдельных сообществ. При создании специализированного программного продукта на основе представленной методики появляется возможность оперативного доступа к структурированной информации о происшествиях экстренным службам, позволяя тем самым не только сократить время реагирования на чрезвычайные ситуации, но и своевременно предотвращать распространение ложной информации.

Ключевые слова: распространение информации, стохастическая игра, модель поведения пользователей, модель сообщества, социальная сеть

**METHODOLOGY FOR ANALYSIS OF DATA ON EMERGENCIES
IN SOCIAL NETWORKS****Vostrykh A.V., Maksimov A.V., Matveev A.V., Smirnov A.S.***Saint Petersburg University of the State Fire Service of the Ministry of Emergency Situations
of Russia named after the Hero of the Russian Federation General of the Army E.N. Zinichev,
Saint Petersburg, e-mail: a.vostrykh@list.ru*

Accurate and timely information and incidents play an important role in emergencies. This enables decision makers to draw up prompt and adequate action plans in an unfavorable situation. In recent years, one of the important sources of data on incidents has been various social networks, whose users exchange information by posting text, photo and video materials. The article presents a new original methodology that allows analyzing information published on social networks in order to respond to it in a timely manner. Unlike existing solutions, the proposed methodology takes into account the multicriteria of information dissemination in the real world, easily scales to any social networks, takes into account the relationship between users in the process of distributing information about emergencies. It also has the ability to both low-level analysis of individual users and high-level analysis of individual communities. When creating a specialized software product based on the presented methodology, it becomes possible to instantly access structured information about incidents to emergency services, thereby not only reducing the response time to emergency situations, but also preventing the spread of false information in a timely manner.

Keywords: information dissemination, stochastic game, user behavior model, community model, social network

Сегодня в мире цифровых технологий общество все чаще использует новые каналы связи для коммуникации. В последнее десятилетие активной средой для обмена информацией являются социальные сети («ВКонтакте», «Одноклассники» и др.), в которых люди разных возрастных и социальных категорий делятся своими взглядами, эмоциями и жизненными ситуациями. С помощью данных сетей ежедневно генерируется огромный объем гетерогенной информации, анализ которой может быть полезен для прогнозирования общественных волнений, социальной нестабильности, происшествий различного характера

и т.д. При наступлении каких-либо происшествий или чрезвычайных ситуаций (далее – ЧС) многие пользователи социальных сетей делятся текстовыми, фото- или видеоматериалами с другими пользователями, формируя тем самым огромный массив информации, которая потенциально может быть использована спасательными службами, позволяя им снизить степень информационной неопределенности и помочь в принятии адекватных управленческих решений при реагировании [1, 2].

В настоящей статье сделана попытка применения новостных данных из социальных сетей для снижения информационной

неопределенности, своевременного реагирования на ЧС и проведения превентивных мероприятий с целью минимизации рисков [3], предотвращения возникновения паники и недопущения ухудшения сложившейся ситуации.

Материалы и методы исследования

В целом анализ социальных сетей может быть использован для исследования информационного взаимодействия как отдельных агентов, так и определенных сообществ, образованных в той или иной социальной сети, прогнозирования их поведения, моделирования динамики распространения информации [4].

Ранее многие исследователи в различных научных областях уже проводили анализ социальных сетей. Для получения данных применялись различные подходы. Например, иностранными исследователями широко использовалась классическая модель Кермака – Маккендрика SIR, которая до этого применялась для анализа динамики распространения эпидемий [5], распределенные системы [6], линейные модели влияния [7], эвристико-жадный алгоритм [8], анализ отдельных пользователей [9]. Среди отечественных ученых значительный вклад в развитие теории анализа социальных сетей внесла научная школа Института проблем управления им. В.А. Трапезникова РАН [10]. Кроме того некоторые исследователи в решении данной задачи использовали аппарат теории нечетких множеств и нечеткого когнитивного моделирования [11], клеточные автоматы [12–14], методы имитационного моделирования [15], экспертно-стохастическое моделирование для измерения влияния в социальных сетях [16, 17], цепи Маркова [18] и др.

При использовании данных подходов и адаптации их к реальному миру у исследователей возник ряд вопросов, связанных с многокритериальностью распространения информации в реальном мире; сложностью масштабирования подходов; отсутствием учета взаимоотношений между пользователями в процессе распространения данных о ЧС в социальных сетях.

В статье предлагается новая методика, позволяющая анализировать социальные сети с целью своевременного реагирования на происшествя, а также предотвращения распространения ложной информации. В отличие от существующего подхода к анализу конкретных пользователей социальных сетей [9], в настоящей работе предлагается анализировать также и сообщества, которые объединяют определенное количество пользователей разных социальных

групп и взглядов. Это сделано по причине того, что в настоящее время социальными сетями пользуется огромное количество людей, и анализ каждого из них, в свою очередь, будет требовать значительных вычислительных мощностей и ресурсов операторов на обработку результатов. Также анализировать и полагаться на мнения отдельных пользователей имеет смысл только в тех случаях, когда они владеют достоверной информацией и имеют значительное влияние на общество. Целесообразнее анализировать сообщества, охватывающие не только широкие массы людей и имеющие соответствующую тематику направленности, но и генерирующие достаточно большое количество данных. Другими словами, в статье предлагается учитывать также поступающие от групп и сообществ данные. При необходимости более низкого уровня анализа и детализации возможен переход на учет показателей отдельных пользователей.

В наши дни в России широко распространена социальная сеть «ВКонтакте», в которой, по последним данным, зарегистрировано порядка 72,5 млн пользователей. Около половины из них ведут активную деятельность в сети, объединяясь в различные группы и сообщества. В результате поиска по группам тематики экстренных сообщений на территории Санкт-Петербурга и Ленинградской области было найдено порядка пятидесяти сообществ с количеством подписчиков от тысячи до полутора миллионов. Так, сообщество «ДТП и ЧП | Санкт-Петербург | Питер Онлайн | СПб» имеет 1 346 038 подписчиков, «ДТП и ЧП | Ленинградская область Онлайн | СПб» – 104 637 подписчиков, «ДТП и ЧП Санкт-Петербурга | Новости | Угоны» – 54 222 подписчика и т.д. Ежедневно в каждом из данных сообществ публикуется порядка 10–40 сообщений о происшествиях, таким образом, общее количество сообщений в сообществах может достигать до нескольких тысяч за сутки. Ручная проверка, анализ и реагирование на данные сообщения экстренными службами в настоящее время практически невозможны по двум основным причинам: отсутствие достаточного количества специалистов и отсутствие специального программного продукта, позволяющего автоматически выявлять достоверные сообщения, обобщать, ранжировать и предоставлять оператору результаты вычислений [19, 20]. Данные причины являются взаимоисключающими: имея специализированное программное обеспечение, необязательно содержать внушительный штат сотрудников, а можно обойтись существующей

в настоящее время дежурной сменой. Таким образом, разработка в будущем специализированного программного продукта (далее – ПП) на основе представленной в настоящей статье методики является перспективным направлением для работы. Данный ПП позволил бы решать следующие задачи:

- мгновенный сбор и обобщение информации о происшествиях в режиме реального времени;
- проверка достоверности информации за счет представленного ниже механизма;
- сокращение времени реагирования на происшествия экстренными службами;
- предотвращение распространения ложной информации, пресечение паники населения;
- выявление пользователей, умышленно способствующих распространению ложной информации [21];
- координация сил и средств для реагирования на происшествия и т.д.

Результаты исследования и их обсуждение

Перейдем к описанию предлагаемой методики. Примем следующие условные обозначения: R_i – репутация пользователя A_i , которая отражает доверие других пользователей к мнению, выраженному этим пользователем. Если $R_i = 0$, другие пользователи не доверяют пользователю A_i , напротив, если $0 < R_i < 1$, то другие пользователи доверяют субъективным суждениям пользователя A_i относительно поступающих от него сообщений. Если $R_i = 1$, другие пользователи полностью доверяют мнению пользователя A_i . Обозначим уровень интереса других пользователей к сообщениям пользователя A_i через I_i с аналогичным характером его оценки. Так, при $I_i = 0$ другим пользователям не интересны публикации анализируемого пользователя, при $0 < I_i < 1$ другие пользователи проявляют некоторый интерес к опубликованным суждениям, при $I_i = 1$ – проявляют максимальную заинтересованность.

Характер распространения информации о ЧС в социальной сети зависит от параметров определенного пользователя. При составлении шаблонов поведения можно использовать следующие составляющие:

- скорость прочтения и обработки информации (количество постов, прочитанных за единицу времени):

$$V_r = (Q_s / T_r) \times K_k, \quad (1)$$

где Q_s – число знаков в тексте (объем); T_r – время, затраченное на чтение текста (в минутах); K_k – коэффициент понимания;

- количество полученных откликов в виде оценок (в частности, в сети «ВКонтакте» можно подсчитать количество отметок «нравится», полученных определенным пользователем за публикации):

$$L_k = M_l / T_s, \quad (2)$$

где M_l – количество отметок «нравится», полученных определенным пользователем; T_s – время пользовательской сессии;

- количество поставленных откликов в виде оценок определенного пользователя другим пользователям:

$$L_{ks} = K_{ks} / T_s, \quad (3)$$

где K_{ks} – количество отметок «нравится», поставленных определенным пользователем за все время существования аккаунта; T_s – время пользовательской сессии;

- публикационная активность характеризуется количеством оставленных постов определенным пользователем:

$$P_s = K_{ps} / T_s, \quad (4)$$

где K_{ps} – количество оставленных постов определенным пользователем; T_s – время пользовательской сессии;

- активность распространения характеризуется количеством вторичных публикаций, сделанных определенным пользователем:

$$R_{ps} = K_{rs} / T_s, \quad (5)$$

где K_{rs} – количеством вторичных публикаций, сделанных определенным пользователем; T_s – время пользовательской сессии;

- степень одобрения характеризуется количеством положительных комментариев, оставленных под записью определенного пользователя:

$$K_p = K_{pt} / T_s, \quad (6)$$

где K_{pt} – количество положительных комментариев, оставленных под записью определенного пользователя; T_s – время пользовательской сессии;

- авторитетность пользователя характеризуется количеством активных подписчиков у определенного пользователя:

$$A_f = \frac{\sum A_p | P_s(A_p) > k}{T_{er}}, \quad (7)$$

где A_p – активный пользователь; P_s – публикационная активность; T_{er} – временной интервал (неделя, месяц, год – зависит от целей исследования); k – коэффициент, определяющий нижнюю границу понятия «активный пользователь» в используемом контексте;

– обратная активность распространения характеризуется количеством вторичных публикаций определенного автора другими пользователями:

$$R_{pt} = \frac{\sum P_{tw} | P_{tw}(A_p) > k}{T_{er}}, \quad (8)$$

где A_p – активный пользователь; P_{tw} – количество вторичных публикаций; T_{er} – временной интервал (неделя, месяц, год – зависит от целей исследования); k – коэффициент, определяющий нижнюю границу понятия «активный пользователь» в используемом контексте.

Значения показателей пользователей: личные познания C_p , интересность I_i и репутация R_i – определяются следующим образом:

$$\begin{cases} C_i = (V_r, L_{ks}, P_s, R_{ps}) \\ I_i = (A_f, L_k) \\ R_i = (K_p, R_{pt}) \end{cases} \quad (9)$$

Нормализация уровня интереса к пользователю, его репутации и осведомленности вычисляется с помощью следующего выражения:

$$\begin{cases} c_i = \frac{C_i}{F}, c_i < 1 \\ i_i = \frac{I_i}{F}, i_i < 1 \\ r_i = \frac{R_i}{F}, r_i < 1, F \rightarrow +\infty \end{cases} \quad (10)$$

Цель пользователя, предоставляющего информацию, заключается в максимизации своей полезности u_i в социальной сети, которая определяется как комбинация личных познаний C_p , интересности I_i и репутации R_i :

$$u_i(C, I, R) = \varphi * C_i + \omega * I_i + \tau * R_i, \quad 0 \leq \varphi, \omega, \tau \leq 1; \quad \varphi + \omega + \tau = 1. \quad (11)$$

Набор весовых коэффициентов φ , ω , τ предназначен для идентификации конкретного типа личности пользователя, а также возможна корректировка в соответствии с приоритетами вычислений, когда, напри-

мер, необходимо подчеркнуть значимость для исследования одного из параметров.

Относительно составления шаблонов сообществ можно выделить следующие показатели:

– количество активных пользователей (пользователи, оставляющие комментарии к постам, отметки «нравится», делающие вторичные публикации):

$$A_{ci} = \frac{\sum A_p}{T_{er}(A_{ci})}; \quad (12)$$

– количество постов в сутки на стене сообщества:

$$Mk_{ps} = \frac{K_{ps}}{T_{er}(Mk_{ps})}; \quad (13)$$

– среднее количество просмотров одного поста:

$$\bar{V}_{is} = \frac{V_{is}}{T_{er}(V_{is})}; \quad (14)$$

– среднее количество комментариев, оставленных к одному посту:

$$\bar{K}_{om} = \frac{K_{om}}{T_{er}(K_{om})}; \quad (15)$$

– среднее количество вторичных публикаций анализируемого сообщества в других сообществах:

$$\bar{D}_{pos} = \frac{D_{pos}}{T_{er}(D_{pos})}; \quad (16)$$

– среднее количество отметок «нравится» в других сообществах записи, изначально опубликованной в анализируемом сообществе и вторично опубликованном в других:

$$\bar{D}_{lk} = \frac{D_{lk}}{T_{er}(D_{lk})}. \quad (17)$$

Количество данных параметров может быть увеличено в зависимости от целей исследования и функциональных возможностей социальной сети.

С помощью представленных выше параметров можно выразить актуальность C_p , интересность I_p и репутацию R_p сообществ в следующем виде:

$$\begin{cases} C_p = (A_{ci}, Mk_{ps}) = \frac{\sum A_p}{T_{er}(A_{ci})} * k_{A_{ci}} + \frac{K_{ps}}{T_{er}(Mk_{ps})} * k_{Mk_{ps}} \\ I_p = (\bar{V}_{is}, \bar{K}_{om}) = \frac{V_{is}}{T_{er}(V_{is})} * k_{V_{is}} + \frac{K_{om}}{T_{er}(K_{om})} * k_{K_{om}} \\ R_p = (\bar{D}_{pos}, \bar{D}_{lk}) = \frac{D_{pos}}{T_{er}(D_{pos})} * k_{D_{pos}} + \frac{D_{lk}}{T_{er}(D_{lk})} * k_{D_{lk}} \end{cases} \quad (18)$$

Нормализация уровня интереса к сообществу, его репутации и актуальности вычисляется с помощью следующего выражения:

$$\begin{cases} c_p = \frac{C_p}{F}, c_p < 1 \\ i_p = \frac{I_p}{F}, i_p < 1 \\ r_p = \frac{R_p}{F}, r_p < 1, F \rightarrow +\infty \end{cases} \quad (19)$$

Так же как и для отдельных пользователей, цель сообщества состоит в максимизации своей полезности u_p в социальной сети, которая представляется в виде аддитивной свертки трех показателей (актуальности C_p , интересности I_p и репутации R_p):

$$u_p(C, I, R) = \varphi * C_p + \omega * I_p + \tau * R_p, \quad 0 \leq \varphi, \omega, \tau \leq 1; \quad \varphi + \omega + \tau = 1. \quad (20)$$

Рассмотренные показатели напрямую влияют на оценку достоверности публикуемой информации в анализируемом сообществе, так как из-за имеющейся конкуренции администраторы данных групп стремятся получить как можно большую аудиторию с помощью качественных и интересных материалов, обладающих определенной ценностью. Это позволяет привлекать в сообщество больше людей и подниматься в рейтинге, становясь более популярным. Чем выше рейтинг сообщества, тем большая активность предоставления и передачи информации в нем присутствует.

В свою очередь пользователи, находящиеся в данном сообществе по личным или иным целям, осуществляют некоторую

деятельность, которая влияет на все сообщество в целом. Предположим, что в определенном сообществе состоят два пользователя, целью которых, является максимизация ожидаемой полезности. Их отношения между собой рассмотрим в виде стохастической игры с общей суммой:

$$S_k = (N, W, D, Q, \pi, V, \beta), \quad (21)$$

где $N = \{1, 2, \dots, n\}$ – множество пользователей; W – набор характеристик пользователей; $D = \{D_1, D_2, \dots, m\}, m \in N$ – множество наборов доступных действий пользователя N ; Q – функция перехода $Q: W \times D \times W \rightarrow [0, 1]$; π – стратегия пользователей, демонстрирующая вероятность выбора конкретного перехода $\pi: T \rightarrow [0, 1]$; $V = \{V_1, V_2, \dots, V_r\}, V_r \in (-\infty; +\infty), r \in N$ – функция вознаграждения для пользователей, осуществляющих переход по ссылке; β ($0 < \beta < 1$) – коэффициент дисконтирования будущего вознаграждения.

Допустим, что пользователь i и пользователь j являются соседями в социальной сети, и их игра ведется следующим образом. В момент времени t пользователь i принимает действие d_t^i из множества D_i , а пользователь j принимает действие d_t^j из множества D_j в состоянии w_t . Пользователь i получает вознаграждение $r_t^i = V^i(w_t, d_t^i, d_t^j)$, а пользователь j получает вознаграждение $r_t^j = V^j(w_t, d_t^i, d_t^j)$. Затем игра переходит в новое состояние w_{t+1} с условной вероятностью:

$$Prob(w_{t+1} | w_t, a_t^i, a_t^j) = Q(w_t, a_t^i, a_t^j, w_{t+1}). \quad (22)$$

Ожидаемое вознаграждение определим как вектор:

$$\begin{cases} U_{\pi^i, \pi^j}^k = [U_{\pi^i, \pi^j}^k(w_1) \dots U_{\pi^i, \pi^j}^k(w_N)]^T, K = \{i, j\} \\ U_{\pi^i, \pi^j}^k(w) = E_{\pi^i, \pi^j}^k \{r_t^k + \beta r_{t+1}^k + \dots + (\beta)^N r_{t+N}^k\} = E_{\pi^i, \pi^j}^k \left\{ \sum_{n=0}^N (\beta)^N r_{t+n}^k \right\}. \end{cases} \quad (23)$$

Оператор ожидания $E_{\pi^i, \pi^j}^k \{ \cdot \}$ характеризует выполнение действия пользователя k , используя распределение вероятностей $\pi^k(w_{t+n})$ в w_{t+n} , после чего он получает вознаграждение:

$$r_{t+n}^k = \pi^i(w_{t+n})^T V^k(w_{t+n}) \pi^j(w_{t+n}), n \geq 0, V^k(w) = [V^k(w, d^i, d^j)], K = \{i, j\}, \quad (24)$$

где $K = \{i, j\}$ – матрица вознаграждения пользователя в состоянии w .

При игре двух агентов равновесие Нэша (π^* , π^*) обеспечивается при следующем условии:

$$U^i(\pi^{i*}, \pi^{j*}) \geq U^i(\pi^i, \pi^{j*}); \quad U^j(\pi^{i*}, \pi^{j*}) \geq U^j(\pi^i, \pi^j). \quad (25)$$

Для обеспечения равновесия агенты следуют стратегиям π^* и π^* . Отклонение от стратегий приводит к снижению полезности.

Выразим множество вероятностных векторов длины n через

$$\mathcal{G}^n = \{q \in \varepsilon^n \mid \sum_{j=1}^n q_j = 1, q_j \geq 0\}. \quad (26)$$

Для рассматриваемой игры существует по крайней мере одна смешанная стратегия, приводящая к равновесию Нэша, которая для стохастической игры может быть найдена путем решения задачи нелинейного программирования. Для стохастической игры из двух игроков получаем следующее выражение:

$$\left\{ \begin{array}{l} \min_{u^i, u^j, \pi^i, \pi^j} \{1^T [u^k - R^k(\pi^i, \pi^j) - \beta Q(\pi^i, \pi^j) u^k]\} \\ n p_i R^i(w_l) \pi^i(w_l) + \beta T(w_l, \mu^i) \pi^j(w_l) \leq \mu^i(w_l), \quad l = 1, \dots, N; \\ \pi^i(w_l)^T R^j(w_l) + \beta \pi^i(w_l, \mu^j) T(w_l, \mu^i) \leq \mu^j(w_l) 1^T, \quad l = 1, \dots, N. \\ T(w, \mu^k) = [q(w_1 | w, d^i, d^j) \dots q(w_N | w, d^i, d^j)]^T \mu^k \\ d^i \in D_i, d^j \in D_j, w_i \in W, \mu^i \in \varepsilon^N, \pi^k \in \mathcal{G}^{1^T} \end{array} \right. \quad (27)$$

Глобальный минимум целевой функции определяет требуемые условия оптимальности, решение задачи нелинейного программирования (π^i, π^j, u^i, u^j) будет формировать равновесие по Нэшу.

Методику анализа данных о ЧС и ее распространения в социальных сетях можно представить следующей последовательностью шагов:

Шаг 1 – Выбор социальной сети для анализа.

Шаг 2 – Выявление значимых кандидатов для оценки сообществ из пользователей сети.

Шаг 2.1 – Оценка репутации кандидатов по формулам (6) и (8).

Шаг 2.2 – Оценка интересности кандидатов по формулам (2) и (7).

Шаг 2.3 – Оценка личных познаний кандидатов по формулам (1), (3), (4), (5).

Шаг 2.4 – Нормализация показателей пользователей (10).

Шаг 2.5 – Ранжирование пользователей с помощью формулы (11).

Шаг 3 – Анализ сообществ по заданной тематике.

Шаг 3.1 – Поиск сообществ по их названиям в соответствии с ключевыми словами.

Шаг 3.2 – Оценка актуальности кандидатов по формулам (12) и (13).

Шаг 3.3 – Оценка интересности кандидатов по формулам (14) и (15).

Шаг 3.3 – Оценка репутации кандидатов по формулам (16) и (17).

Шаг 3.4 – Нормализация показателей сообществ с помощью формулы (19).

Шаг 3.5 – Ранжирование сообществ с помощью формулы (20).

Шаг 4 – Анализ данных по заданной тематике из выбранных сообществ.

Шаг 5 – Вывод результатов оператора в виде сообщений по заданной тематике, ранжированных по авторитетности пользователей.

Приведем пример использования предложенной методики. Допустим, в сети находится $N = 1000$ пользователей, коэффициент неопределенности $\lambda = 0,5$, исходные значения репутации r_i выбраны случайным образом, коэффициент дисконтирования будущего вознаграждения $\beta = 0,5$.

В примере используем два типа пользователей: p_1 – пользователь, пересылающий сообщения, основываясь на личной заинтересованности (p_{1+} – истинные сообщения, p_{1-} – ложные сообщения), поэтому $\varphi = \tau = 0$, $\psi = 1$; пользователь p_2 (p_{2+} – истинные сообщения, p_{2-} – ложные сообщения) имеет высокую репутацию, поэтому его коэффициенты будут равны $\varphi = \tau = 0,5$, $\psi = 0$.

Используя формулу (11), получаем, что пользователи p_1 будут распространять информацию с вероятностью $s = 0,71$, пользователи p_2 с вероятностью $s = 0,45$. С помощью среды разработки PyCharm, языка программирования Python и библиотеки matplotlib построены кривые зависимости динамики распространения информации от параметров пользователей (рис. 1).

Из представленной диаграммы можно заметить, что ложные сообщения в начальной фазе быстро распространяются в широкие массы, но затем с течением времени их популярность падает. Это связано с тем, что после того, как в социальной сети появляется информация, например о ЧС, пользователи не сразу понимают, где истинные новости, но с течением времени получают подтверждение из других источников, и происходит отсеивание ложных сообщений.

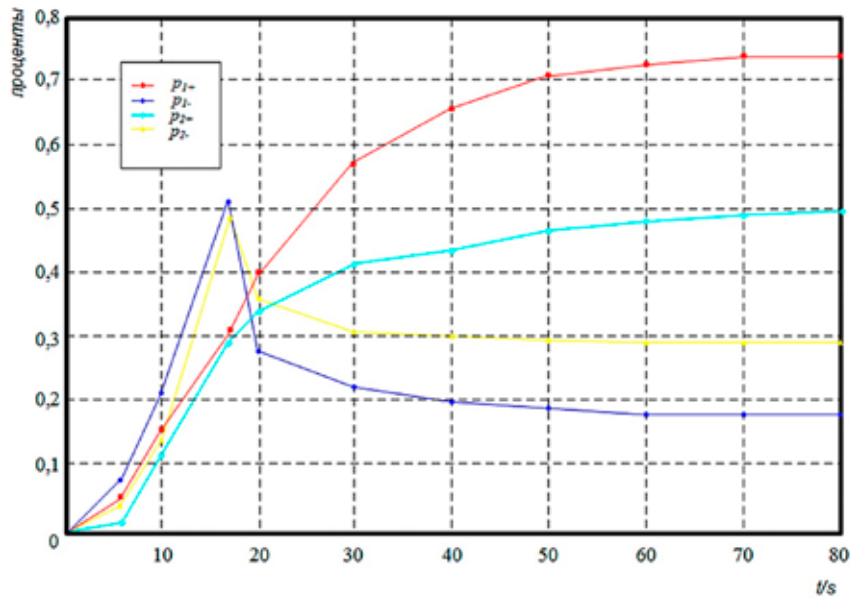


Рис. 1. Зависимости динамики распространения информации от параметров пользователей

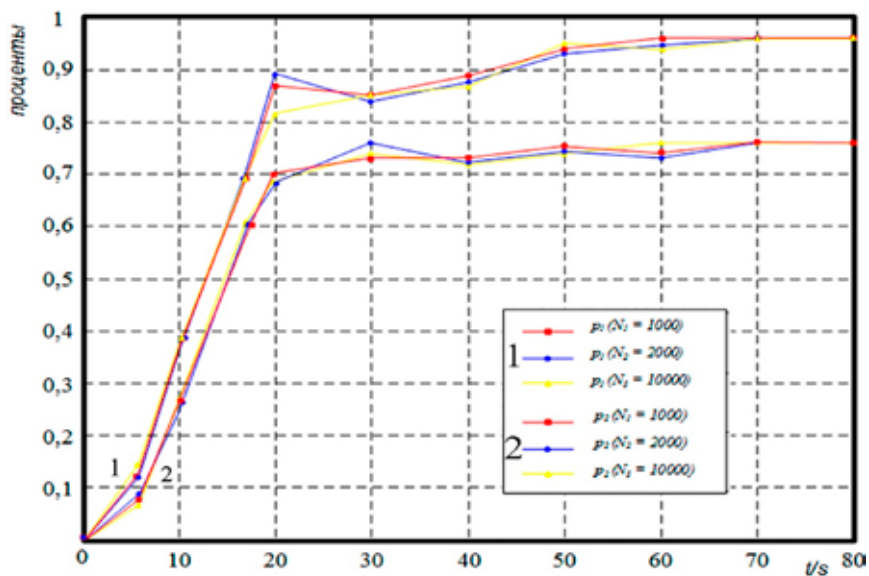


Рис. 2. Зависимость динамики распространения информации от размеров соцсети

Также из рис. 1 можно заметить, что пользователи, обладающие различными параметрами, имеют разный масштаб и скорость распространения их информации в социальной сети. Хотя распространение информации пользователями p_1 быстрее, чем p_2 , это не означает, что пользователи p_2 менее активные, а скорее это свидетельствует о том, что они делятся информацией с большей осторожностью.

В процессе исследования были сделаны выводы о том, что масштаб социальной сети практически не влияет на скорость

распространения информации. На рис. 2 представлена динамика распространения информации в социальных сетях с численностью пользователей: $N_1 = 1000$, $N_2 = 2000$, $N_3 = 10000$.

Также из рис. 2 можно выделить стадию стремительного распространения информации. Обладая достоверной информацией и имея авторитетный источник распространения, можно предотвратить популяризацию ложной информации на начальной стадии или снизить эффект ее воздействия на пользователей.

Заключение

Таким образом, в статье представлен процесс распространения информации по социальным сетям, выявлены параметры, влияющие на скорость ее популяризации, представлены два уровня анализа данных (уровень отдельных пользователей и уровень сообществ), представлена стохастическая игровая модель взаимодействия пользователей при обмене информацией, предложена новая методика, позволяющая анализировать информацию, опубликованную в социальных сетях о происходящих ЧС, с целью своевременного реагирования на них.

При разработке на основе представленной методики специализированного ППП появляется возможность мгновенного доступа к структурированной информации о происшествиях экстренным службам, что позволит сократить время реагирования на ЧС, а также своевременно предотвращать распространение ложной информации.

Список литературы

1. Максимов А.В., Матвеев А.В. Перспективы использования коллективных знаний при реагировании на чрезвычайные ситуации // Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России. 2019. № 4. С. 89–97.
2. Еникеева К.Р., Абдуллин А.Х., Христуло О.И., Исаева (Юсупова) Ю.И. О роли сервисов социальных сетей для поддержки принятия решений в чрезвычайных ситуациях // Проблемы анализа риска. 2016. Т. 13, № 1. С. 36–45.
3. Зенина Е.А., Матвеев А.В. Актуальные вопросы моделирования информационного противодействия терроризму // Национальная безопасность и стратегическое планирование. 2016. № 1 (13). С. 57–60.
4. Батура Т.В. Методы анализа компьютерных социальных сетей // Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2012. Т. 10, № 4. С. 13–28.
5. Gruhl D., Guha R., Nowell D.L., Tomkins A. Information Diffusion Through Blogspace // Proceedings of the 13th international conference on World Wide Web. 2004. Vol. 6. P. 43–52.
6. Cinelli M. et al. The COVID-19 social media infodemic // Scientific reports. 2020. Vol. 10, Is. 1. P. 1–10.
7. Yang J., Leskovec J. Modeling Information Diffusion in Implicit Networks. Data Mining // 2010 IEEE 10th International Conference on Data Mining. 2010. P. 599–608.
8. Aghaee Z., Kianian S. Efficient influence spread estimation for influence maximization // Social Network Analysis and Mining. 2020. Vol. 10. P. 1–21.
9. Ding J., Liu L., Wang Y. Stochastic Game Model for Information Dissemination of Emergency Events in Social Network // Intelligence and Security Informatics (ISI), IEEE International Conference. 2013. P. 166–168.
10. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Модели влияния в социальных сетях // Управление большими системами: сборник трудов. 2009. № 27. С. 205–281.
11. Ажмухамедов И.М., Мачуева Д.А., Жолобов Д.А. Моделирование процесса распространения информации в социальных сетях // Фундаментальные исследования. 2017. № 5. С. 9–14.
12. Горковенко Д.К. Сравнительный анализ моделей эпидемии и клеточного автомата при моделировании распространения информации в социальных сетях // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Информатика. Телекоммуникации. Управление. 2017. Т. 10, № 3. С. 103–113. DOI: 10.18721/JCSTCS.10309.
13. Гончаров И.В., Паринов П.А., Сирота А.А. Моделирование процессов информационно-психологического воздействия в социальных сетях // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. 2018. № 2. С. 93–104.
14. Ломакин С.Г., Федотов А.М. Анализ модели передачи информации в сети клеточных автоматов // Вестник Новосибирского государственного университета. Серия: Информационные технологии. 2014. Т. 12, № 3. С. 86–99.
15. Бадрызов В.А., Сидельцев В.В. Оценка эффективности распространения информации в социальных сетях с использованием имитационного моделирования // Креативная экономика. 2018. Т. 12, № 9. С. 1359–1372. DOI: 10.18334/ce.12.9.39389.
16. Ломакин М.И., Докукин А.В., Соседов Г.А. Модель измерения влияния в социальных сетях // Компетентность. 2014. № 7 (118). С. 34–39.
17. Бреев В.В. Стохастические модели социальных сетей // Управление большими системами: сборник трудов. 2009. № 27. С. 169–204.
18. Губанов Д.А., Новиков Д.А., Чхартишвили А.Г. Модели информационного влияния и информационного управления в социальных сетях // Проблемы управления. 2009. № 5. С. 28–35.
19. Буйневич М.В., Максимов А.В., Вострых А.В. Анализ результатов аудита сетевых информационных ресурсов МЧС России // Вестник Санкт-Петербургского университета Государственной противопожарной службы МЧС России. 2020. № 1. С. 101–110.
20. Вострых А.В. Анализ информационных систем, используемых в МЧС России для мониторинга и прогнозирования чрезвычайных ситуаций // Сервис безопасности в России: опыт, проблемы, перспективы. Мониторинг, предотвращение и ликвидация чрезвычайных ситуаций природного и техногенного характера: материалы международной научно-практической конференции. Санкт-Петербург, 2021. С. 257–260.
21. Наумова Т.Е. Проблемы верификации в социальных сетях сообщений об угрозе чрезвычайной ситуации // Столыпинский вестник. 2023. Т. 5. № 2.