

Моделирование динамики изменения количества комментариев пользователей масс-медиа на основе уравнения Фоккера-Планка и параметров сетей их связей

Юлия П. Перова, Владимир Н. Калинин и Сергей А. Лесько

Аннотация: В ходе работы было проведено исследование динамики настроений пользователей интернет-медиа с использованием уравнения Фоккера-Планка и использованием параметров графов сетей комментариев. В статье вводится понятие базового вектора состояния графа, который включает среднюю величину коэффициента посредничества, среднее значение коэффициента кластеризации и долю пользователей в различных состояниях. Зависимость от времени косинуса угла между базовым и текущим вектором состояния графа формирует временной ряд, значения уровней которого можно интерпретировать как «блуждания точки» на отрезке возможных значений, что характеризует состояния сети. Текущее состояние графа комментариев может быть определено с помощью инструментов сетевого и текстового анализа. Решение уравнения Фоккера-Планка позволяет получить аналитическую зависимость для плотности вероятности обнаружения величины состояния сети в заданном значении на отрезке возможных значений. Этот подход может быть использован для создания алгоритма прогнозирования времени достижения заданного состояния графа сети комментариев пользователей с заданным уровнем вероятности. Анализ модели подтверждает ее адекватность и не противоречивость.

Ключевые слова: социальные сети, граф сети, характеристики графа социальной сети, моделирование социальных процессов, нелинейная динамика, уравнение Фоккера – Планка, мониторинг, управление

I. ВВЕДЕНИЕ

В ходе работы было проведено исследование динамики настроений пользователей интернет-медиа с использованием уравнения Фоккера-Планка и использованием параметров графов сетей комментариев. В статье вводится понятие базового вектора состояния графа, который включает среднюю

Статья получена 4 марта 2023. Исследование выполнено при поддержке Российского научного фонда (РНФ), грант № 22-21-00109 «Разработка моделей прогнозирования динамики социальных настроений на основе анализа временных рядов текстового контента социальных сетей с использованием уравнений Фоккера-Планка и нелинейной диффузии».

Ю. П. Перова, Институт кибербезопасности цифровых технологий, РГУ МИРЭА, Москва, Россия (e-mail: jul-np@yandex.ru).

В. Н. Калинин, Институт кибербезопасности цифровых технологий, РГУ МИРЭА, Москва, Россия (e-mail: vkalinin@mail.ru).

С. А. Лесько, Институт кибербезопасности цифровых технологий, РГУ МИРЭА, Москва, Россия (e-mail: sergey@testor.ru).

величину коэффициента посредничества, среднее значение коэффициента кластеризации и долю пользователей в различных состояниях. Зависимость от времени косинуса угла между базовым и текущим вектором состояния графа формирует временной ряд, значения уровней которого можно интерпретировать как «блуждания точки» на отрезке возможных значений, что характеризует состояния сети. Текущее состояние графа комментариев может быть определено с помощью инструментов сетевого и текстового анализа. Решение уравнения Фоккера-Планка позволяет получить аналитическую зависимость для плотности вероятности обнаружения величины состояния сети в заданном значении на отрезке возможных значений. Этот подход может быть использован для создания алгоритма прогнозирования времени достижения заданного состояния графа сети комментариев пользователей с заданным уровнем вероятности. Анализ модели подтверждает ее адекватность и не противоречивость.

II. АКТУАЛЬНОСТЬ ИССЛЕДОВАНИЯ И ПОСТАНОВКА ЗАДАЧИ

Анализ пользовательских настроений в социальных сетях может служить индикатором изменений в общественных процессах. Часто эти изменения происходят латентно, но начинают проявляться в социальных сетях заметно раньше, чем в других структурах. Для отслеживания таких явлений необходимо проводить анализ структуры сетевых структур и больших объемов текстовой информации в режиме реального времени, которую ежедневно генерируют пользователи социальных сетей. Это представляет собой сложную научную и технологическую задачу.

Одним из методов решения задач мониторинга и прогнозирования общественных настроений является анализ активности пользователей, оставляющих комментарии на блогах и новостных ресурсах по различным темам. С помощью средств текстовой аналитики, которые позволяют группировать тексты по выбранным темам, и инструментов сбора открытых данных из социальных сетей и новостных ресурсов можно определить настроение пользователей и построить графы их связей внутри выбранных тематических групп. Каждый граф имеет свои

характеристики, такие как плотность сети, коэффициент посредничества, коэффициент кластеризации и эластичность, которые могут изменяться ежедневно и образовывать многомерный временной ряд.

Исследуемая тема мониторинга и прогнозирования общественных настроений имеет высокую актуальность, так как ее решение способствует устойчивому экономическому развитию общества и стабильному развитию социальных институтов и отношений.

Можно рассмотреть ряд исследований, показывающий, что анализ социальных сетей является полезным инструментом для получения полной картины общественного настроения.

Например, для процесса поиска сходства между узлами в сети, который становится все более трудоемким с ростом сети, исследователи из работы [1] применили роевые алгоритмы. В данной статье методы оптимизации на основе роя, используемые для анализа социальных сетей, сравниваются с традиционными подходами к анализу сообщества и связей.

В работах [2-4] была предложена методика KroMFac, которая использует регуляризованную неотрицательную матричную факторизацию (NMF) на основе графовой модели Кронекера для обнаружения сообществ. Этот подход объединяет методы анализа сети и обнаружения сообществ в единую унифицированную структуру. Методика KroMFac объединяет четыре области исследований: обнаружение сообществ на графах, обнаружение перекрывающихся сообществ, обнаружение сообществ в неполных сетях с отсутствующими краями и полных сетей.

Динамика общественных настроений пользователей интернета в значительной степени подчиняется стохастическим процессам. Присутствие человеческого фактора, представленного множеством людей с различными мнениями, предпочтениями и характером поведения, вносит случайность в изменения и одновременно элементы целенаправленности. Подробное описание использования стохастических методов для моделирования динамики социальных процессов можно найти в [5].

Широко применяемым инструментом для анализа и моделирования нестационарных процессов в различных сложных системах является уравнение Фоккера-Планка. Оно позволяет достичь хорошего согласования между прогнозируемым поведением и наблюдаемыми данными. Мы предполагаем, что это уравнение может быть использовано для анализа и моделирования появления комментариев к новостным статьям и блогам. Уравнение Фоккера-Планка имеет следующий вид:

$$\frac{\partial \rho(x,t)}{\partial t} = -\frac{\partial}{\partial x} [\mu(x) \cdot \rho(x,t)] + \frac{1}{2} \frac{\partial^2}{\partial x^2} [D(x) \cdot \rho(x,t)] \quad (1)$$

где $\rho(x,t)$ зависящая от времени t плотность вероятности распределения по состояниям x (в нашем случае состояние x это число комментариев, наблюдаемое в момент времени t), $D(x)$ – зависящий от состояния x коэффициент определяющий случайное изменение состояния x , $\mu(x)$ – зависящий от состояния x коэффициент определяющий целенаправленное

изменение состояния x . Это уравнение учитывает, как упорядоченные, так и случайные изменения.

Уравнение Фоккера-Планка широко используется для анализа и моделирования поведения временных рядов в описании процессов в сложных системах [6-9].

Необходимо отметить, что для моделирования на основе дифференциальных уравнений используются и другие подходы, помимо уравнения Фоккера-Планка, такие как уравнения Лиувилля [9, 10], уравнения диффузии [8, 11] и множество других, например, с применением теоретико-игровых подходов и методов принятия управленческих решений на их основе [12].

В данной статье мы предлагаем использовать векторное представление для описания состояний сети комментариев. Векторы будут состоять из элементов допустимых значений параметров сети, таких как плотность, коэффициент посредничества, коэффициент кластеризации, эластичность и другие. Кроме того, элементы векторов будут включать характеристики пользователей, такие как доля лоялистов, оппозиционеров, троллей и нейтральных пользователей на основе анализа текстов комментариев. Для характеристики желаемых или нежелательных состояний всей социальной сети в целом мы будем использовать базовые векторы.

Величина косинусного расстояния между базовым вектором и вектором текущего состояния изменяется со временем и может быть рассмотрена как «блуждание точки» на отрезке от L_{\min} до L_{\max} , либо как случайный временной ряд, поведение которого можно описать на основе уравнения Фоккера - Планка. Определенное значение этого расстояния (состояние, при котором необходимо принимать управленческие решения) можно рассматривать как ловушку или точку допустимого порога, куда может попасть "блуждающая точка" со временем. Это позволяет создавать вероятностные социодинамические модели для прогнозирования динамики общественных настроений.

III. СБОР И ОБРАБОТКА ДАННЫХ

Для исследования временных рядов активности пользователей был выбран информационный ресурс «РИА-новости», который был выбран из-за его популярности и узнаваемости в российском обществе.

На момент 2022 года этот новостной ресурс занимает первое место среди российских медиа по версии «br-analytics» и входит в ТОП-3 самых цитируемых информационных агентств в СМИ.

Для сбора базы данных и коллекции текстов, новостей и их метаданных использовались программно-технические средства, такие как «парсер», разработанный для сайта ria.ru.

Собираемые данные:

1) статья (article): id статьи, заголовок, ссылка, дата и время.

2) комментарий (comment): id комментария, id статьи, id профиля пользователя, текст комментария, дата и время публикации комментария, уровень комментария, ссылка на родительский комментарий.

3) профиль пользователя (profile): id профиля, имя

пользователя, ссылка.

Структура таблиц базы данных парсера представлена на рисунке 1.

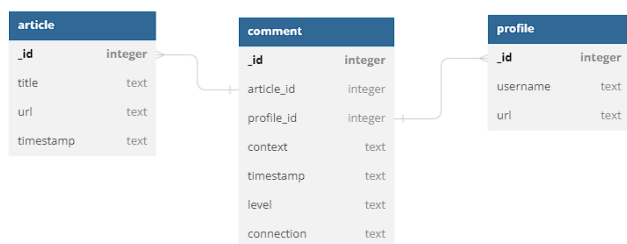


Рисунок 1 – Структура базы данных

Для проведения исследования по анализу структуры графов комментариев была выбрана одна из новостей, опубликованных на портале «РИА Новости». Были собраны все комментарии пользователей к этой новости и получены доступные данные. Всего было обработано 633 комментария, которые были помечены на основе анализа текстов как относящиеся к одному из четырех типов: лоялист, оппозиционер, тролль или не определенные.

Группа «не определенные» была создана, потому что в большинстве комментариев было недостаточно информации, чтобы определить принадлежность пользователя к одной из трех других групп. Некоторые комментарии (77 штук) были удалены модераторами сайта за нарушения правил, но информация о их существовании была сохранена без текстов. В итоге общее число узлов графа составляет 710. Пользователи распределены следующим образом: лоялисты - 10,28% (73 узла), оппозиционеры - 43,8% (311 узлов), тролли - 30,42% (216 узлов), «не определенные» - 4,65% (33 узла), удаленные - 10,85% (77 узлов).

На рисунке 2 изображена структура графа, полученного в результате анализа комментариев к новостям. При использовании цветовой кодировки, узлы графа могут быть раскрашены в разные цвета в зависимости от их состояния (принадлежности к одному из четырех типов). Связи на рисунке 2 показывают взаимодействие пользователей друг с другом через комментарии. Таким образом, по цвету узлов можно сделать вывод о их состоянии, а по ребрам графа - о взаимодействии.

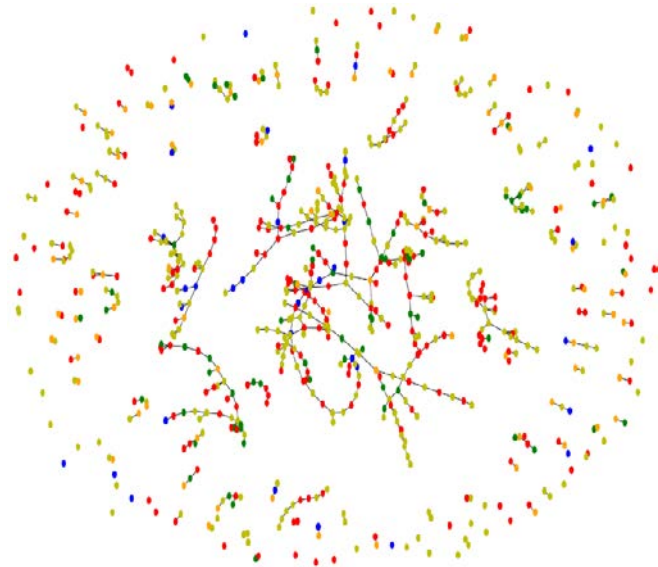


Рисунок 2 – Структура графа комментариев к рассматриваемой новости

На рисунке 2 можно увидеть, что в данной структуре присутствует множество несвязанных одиночных вершин. На рисунке 3 отдельно показана связанная компонента графа. Замкнутые овалы на рисунке 3 указывают на комментарии, оставленные пользователями о самих себе.

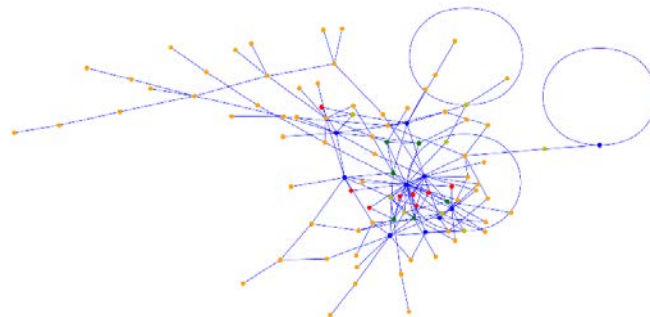


Рисунок 3 – Граф взаимосвязей пользователей по комментарием

Рассмотрим элементы вектора состояния сети, которые мы будем использовать в нашей модели:

Доля узлов, которые имеют определенное состояние, например, тех, кто отрицательно настроен по отношению к какому-либо событию в общественной жизни;

Коэффициент кластеризации — это мера плотности связей данной вершины со своими соседями. Он вычисляется как отношение реального количества связей между ближайшими соседями узла i к максимально возможному количеству связей, при котором все ближайшие соседи данного узла были бы соединены непосредственно друг с другом. Значение коэффициента кластеризации узла находится в диапазоне от 0 до 1. Чем выше значение коэффициента, тем более значимой является данная вершина в обмене информацией;

Степень посредничества показывает, какое количество кратчайших путей между всеми парами узлов сети проходит через данный узел, в отношении общего числа всех кратчайших путей в сети. Значение степени посредничества находится в диапазоне от 0 до 1. Чем выше значение степени посредничества, тем более значимая роль данной вершины в обмене информацией.

Окраска узлов на рисунке 3 показывает значения их коэффициентов кластеризации:

- оранжевые – 0;
- синие – от 0,01 до 0,1;
- зеленые – от 0,11 до 0,3;
- желтые – от 0,31 до 0,6;
- красные – от 0,61 до 1.

Представленный на рисунке 3 граф обладает следующими характеристиками:

Доля узлов имеющих определенное состояние («оппозиционеры»): 0,44

Средний коэффициент кластеризации: 0,11

Средняя степень посредничества: 0,15

Определим значения элементов базового вектора состояния сети, обозначенного как « θ ». Они определяют пороговые значения, которые считаются нежелательными при переходе через них в контексте

управления состоянием. Учитывая, что в любом сообществе всегда есть доля участников (от 0,10 до 0,15), которые не согласны с любым решением, мы можем принять долю тех, кто отрицательно относится к рассматриваемому событию, равной 0,12. Желаемое среднее значение коэффициента кластеризации и средней степени посредничества узлов в такой сети также будет небольшим и равным 0,05. Таким образом, базовый вектор будет иметь вид $\theta = (0,12; 0,05; 0,05)$.

При расчетах косинусного расстояния между базисным вектором и вектором текущего состояния сети, его величина может изменяться в диапазоне от 0 до 1. Можно ввести индикаторы состояний, т.е. пороги значения косинуса, сигнализирующие о приближении к нежелательному состоянию. К примеру, на рисунке 4 обозначены три порога: зеленый, желтый и, а синяя точка – текущее значение косинуса.

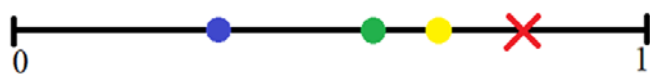


Рисунок 4 – блуждание точки

Заметим, что количество параметров, необходимых для описания состояния сети, может быть больше, чем то, которое мы выбрали. Мы отобрали только те параметры, которые на наш взгляд наиболее значимы. Кроме того, выбранные параметры были нормированы и лежат в диапазоне от 0 до 1, что позволяет им одинаково влиять на расчет метрики расстояний.

Мы предлагаем объединить различные графы комментирования новостей на выбранном ресурсе по определенным тематикам в единую структуру за сутки через связи между узлами пользователей. Это позволит выделить большой граф, описывающий активность пользователей сетевого информационного ресурса за сутки. Затем можно определить элементы вектора текущего состояния, который описывает его характеристики.

За определенное время изменения компонентов данного вектора в течение каждого дня будут формировать временной ряд многомерного типа.

IV. ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

Чтобы построить модель на основе уравнения Фоккера - Планка, необходимо сделать предположения о зависимости $D(x)$ и $\mu(x)$ от состояния x и рассмотреть два условия. Во-первых, нужно учитывать размерность членов, входящих в уравнение (1). Во-вторых, можно предположить, что при росте состояния x (росте числа возможных комментариев или значимости новости или блога) величины $D(x)$ и $\mu(x)$ также должны увеличиваться. Логично предположить, что все члены уравнения (1) должны иметь одинаковую размерность, которую имеет $\rho(x)$. И первое, и второе условие будет выполнено если зависимости $D(x)$ и $\mu(x)$ от состояния x будут иметь вид: $\mu(x)=\mu_0 \cdot x$ и $D(x)=D_0 x^2$. В таком виде это с одной стороны обеспечен рост $D(x)$ и $\mu(x)$ при увеличении состояния x , а с другой стороны будет выполнено условие сохранения размерности.

Относительно нашей модели $D(x)$ можно объяснить, что он отражает действия пользователя, возникшие в результате спонтанного импульса, вызванного

прочтением новостей или комментариев других пользователей. Если событие, описываемое в новости или блоге, не является критически важным, но пользователь все же готов потратить время на комментарий или ответить другому комментатору, это свидетельствует о его спонтанном желании отреагировать на данную новость. С другой стороны, $\mu(x)$ можно рассматривать как целенаправленные действия, вызванные желанием пользователя отреагировать на критически важную для него новость или блог, а также дать комментарий на комментарий другого пользователя, если он затрагивает важную тему для данного пользователя. Это свидетельствует о постоянном интересе пользователя к данной теме.

Решение стационарного уравнения Фоккера-Планка:

$$-\frac{d}{dx} [\mu(x) \cdot \rho(x)] + \frac{1}{2} \frac{d^2}{dx^2} [D(x) \cdot \rho(x)] = 0 \quad (2)$$

при сделанных допущениях имеет вид:

$$\rho(x) = [\gamma - 1] x^{-\gamma} \quad (3)$$

Уравнение (3) является степенной зависимостью и соответствует наблюдаемому на практике закону распределения комментаторов по числу комментариев, который имеет степенную зависимость. Следовательно, можно предположить, что уравнение Фоккера-Планка может быть применено для описания социальных процессов.

Для моделирования динамики общественных настроений можно использовать решение нестационарного уравнения Фоккера-Планка, которое описывает изменение косинусного расстояния между текущим вектором состояния и базовым вектором с течением времени. Для этого необходимо сформулировать краевую задачу, которая позволит решить этот процесс для графа сети комментариев.

Первое граничное условие будет выбрано на основе следующих соображений: состояние $x=L_{\min}$, которое является левой границей отрезка возможных состояний, определяет состояние, которое необходимо избежать, переходя через него (область, находящаяся на отрезке левее данного состояния, является нежелательной). Вероятность обнаружения такого состояния системы может быть не равна нулю. Плотность вероятности, определяющая поток в состоянии $x=L_{\min}$, должна быть равна 0, так как состояния не должны выходить за эту границу (это условие отражения).

$$\rho(x, t)|_{x=L_{\min}} = 0 \quad (a)$$

Второе граничное условие заключается в ограничении области возможных состояний справа до некоторой величины $x=L_{\max}$. При расчетах используется метрика, которая не может превышать величину вектора, элементы которого имеют максимальные значения в выбранных координатах. Вероятность обнаружения такого состояния со временем будет отлична от нуля, однако плотность вероятности, определяющая поток в состоянии $x=L_{\max}$, должна быть равна нулю. Это связано с тем, что расстояние между текущим и базовым вектором состояний ограничено максимальными значениями возможных координат в используемом векторном

пространстве, что реализует условие отражения от границы.

$$\rho(x, t)|_{x=L_{max}} = 0 \tag{b}$$

Для формулирования краевой задачи требуется определить начальное условие. Если в момент времени $t=0$ система находится в состоянии, где расстояние между базовым вектором и текущим вектором равно значению x_0 , то начальное условие может быть выражено следующим образом:

$$\rho(x, t = 0) = \begin{cases} \int \delta(x - x_0) dx = 1, & x = x_0 \\ 0, & x \neq x_0 \end{cases} \tag{c}$$

При заданных начальном и граничных условиях и предположениях о зависимости плотности вероятности обнаружения состояния системы от времени $D(x)$ и $\mu(x)$, решение этого уравнения будет иметь следующий вид:

При $L_{min} \leq x \leq x_0$:

$$\rho_1(x, t) = -2 \frac{x_0^\alpha \cdot x^{-[1+\alpha]} \cdot e^{-\frac{D_0 \alpha^2 t}{2}}}{\ln\left(\frac{L_{max}}{L_{min}}\right)} \sum_{n=1}^M \frac{\sin\left(\pi n \frac{\ln\left(\frac{x_0}{L_{min}}\right)}{\ln\left(\frac{L_{max}}{L_{min}}\right)}\right) \sin\left(\pi n \frac{\ln\left(\frac{x}{L_{min}}\right)}{\ln\left(\frac{L_{max}}{L_{min}}\right)}\right)}{\cos(\pi n)} e^{-\frac{\pi^2 n^2 D_0 t}{2 \left[\ln\left(\frac{L_{max}}{L_{min}}\right)\right]^2}}$$

При $x_0 \leq x \leq L_{max}$:

$$\rho_2(x, t) = 2 \frac{x_0^\alpha \cdot x^{-[1+\alpha]} \cdot e^{-\frac{D_0 \alpha^2 t}{2}}}{\ln\left(\frac{L_{max}}{L_{min}}\right)} \sum_{n=1}^M \frac{\sin\left(\pi n \frac{\ln\left(\frac{x_0}{L_{min}}\right)}{\ln\left(\frac{L_{max}}{L_{min}}\right)}\right) \sin\left(\pi n \frac{\ln\left(\frac{x}{L_{max}}\right)}{\ln\left(\frac{L_{max}}{L_{min}}\right)}\right)}{\cos(\pi n)} e^{-\frac{\pi^2 n^2 D_0 t}{2 \left[\ln\left(\frac{L_{max}}{L_{min}}\right)\right]^2}}$$

где $\alpha = \frac{1}{2} - \frac{\mu_0}{D_0}$

Для определения вероятности того, что система к моменту времени t не достигнет порогового состояния (θ) и будет находиться в пределах от L_{min} до L_{max} , можно использовать следующую формулу:

$$P(\theta, t) = \int_{L_{min}}^{x_0} \rho_2(x, t) dx + \int_{x_0}^{L_{max}} \rho_1(x, t) dx \tag{4}$$

Вероятность $Q(\theta, t)$ того, что пороговое состояние θ к моменту времени t окажется достигнутым или превзойденным, вычисляется по формуле:

$$Q(\theta, t) = 1 - P(\theta, t) \tag{5}$$

Определение границ отрезка возможных состояний, начиная от L_{min} и заканчивая L_{max} мы рассмотрим в разделе «Анализ полученной модели».

V. АНАЛИЗ ПОЛУЧЕННОЙ МОДЕЛИ

Для графа, изображенного на рисунке 2, можно вычислить его характеристики и элементы вектора текущего состояния в момент времени t (принимая за $t=0$): $X(t)=(0,44; 0,11; 0,15)$. Расстояние между заданным базовым вектором желаемого состояния $\theta=(0,12; 0,05; 0,05)$ и вектором текущего состояния $X(t)$ в момент времени $t=0$ будет равно $x_0=0,34$. Проведя анализ динамики временного ряда изменения состояния сети за несколько предыдущих дней и используя уравнения модели, можно решить обратную задачу и определить значения параметров модели μ_0 и D_0 . В данном случае $\mu_0=0,0003$ и $D_0=0,007$.

Правую границу отрезка возможных состояний L_{max} можно определить, как расстояние между базовым вектором (θ) и вектором максимально возможных значений параметров сети $X(t) = (1; 1; 1)$. В данном случае L_{max} равен 1,61. Левую границу для обеспечения страховки можно определить, например, как половину длины заданного базового вектора (в

данном случае $|\theta| = 0,14$), таким образом L_{min} будет равно 0,07.

На рисунке 4 представлены результаты моделирования зависимости от времени вероятности достижения заданного порогового состояния сети.

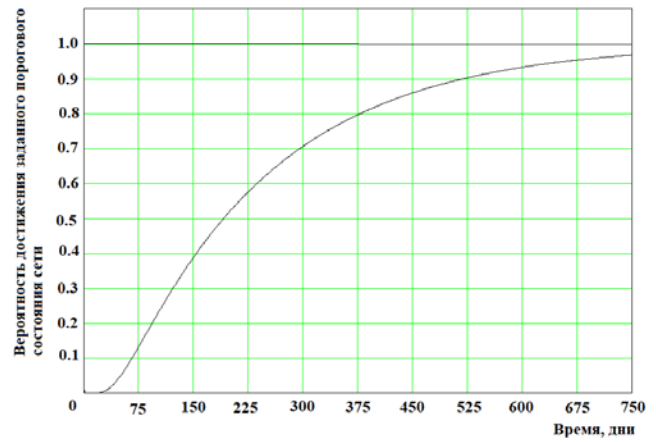


Рисунок 4 – зависимость вероятности достижения заданного порогового состояния сети от времени для рассмотренного примера

Результаты исследования показывают, что при отсутствии воздействия на сеть, требуемое состояние может быть достигнуто с вероятностью 0,8 на 375 день и с вероятностью 0,9 на 525 день. Эти результаты являются возможными, однако требуют дополнительных исследований для оценки их точности.

VI. АЛГОРИТМ ПРОГНОЗИРОВАНИЯ ДОСТИЖЕНИЯ ЗАДАННОГО СОСТОЯНИЯ ГРАФА СЕТИ КОММЕНТАРИЕВ ПОЛЬЗОВАТЕЛЕЙ НОВОСТНЫХ МАСМЕДИА

Для прогнозирования динамики настроений пользователей интернет-медиа можно использовать уравнение Фоккера-Планка и изменять параметры графов сетей комментариев. Для этого применяется следующий алгоритм:

1. Необходимо собрать текстовые комментарии и метаданные пользователей по определенной тематике с новостных сетевых медиаресурсов с указанием даты и времени.

2. После сбора данных, их необходимо обработать с помощью инструментов текстовой аналитики и сентимент-анализа, чтобы получить граф комментариев пользователей по выбранной тематике. Затем рассчитываются характеристики этого графа, такие как плотность сети, средняя величина коэффициента посредничества, среднее значение коэффициента кластеризации, эластичность и доля пользователей с определенным настроением.

3. Далее необходимо задать значения элементов базового вектора, который будет определять достижение желаемого или не желаемого состояния (θ), и на основе обработанных данных и заданного вектора временных рядов изменений графа комментариев пользователей по выбранной тематике с течением времени формировать прогноз.

4. Задается длительность шага τ (час, день, неделя и т.д.), и на нескольких шагах для заданного τ определяются параметры модели μ_0 и D_0 с помощью численных расчетов и уравнений (4) и (5).

5. Последнее среднее значение метрики расстояний между базовым вектором и вектором текущего состояния сети принимается за начальное состояние x_0 . Затем, используя полученные значения μ_0 и D_0 и уравнения (4) и (5), проводятся расчеты, чтобы получить зависимость вероятности достижения желаемого или не желаемого состояния от времени. Для оценки времени достижения заданного уровня вероятности можно задать величину вероятности (например, 0,95).

VII. ЗАКЛЮЧЕНИЕ

В заключение следует отметить, что сложную динамику процессов в сложных социальных системах можно описывать не только с помощью моделей, основанных на уравнении Фоккера-Планка. Некоторые работы [13-16] представляют модели, учитывающие стохастическую динамику изменения состояний в таких системах, а также процессы самоорганизации и память. Это позволяет учитывать не только Марковские, но и не Марковские процессы. В этих исследованиях было выведено нелинейное дифференциальное уравнение второго порядка, которое позволяет формулировать и решать граничные задачи для определения функции плотности вероятности амплитуды отклонений параметров, описывающих наблюдаемые процессы нестационарных временных рядов, в зависимости от интервала времени и глубины учета памяти. Это отличает его от уравнения Фоккера-Планка.

БЛАГОДАРНОСТИ

Исследование выполнено при поддержке Российского научного фонда (РНФ), грант № 22-21-00109 «Разработка моделей прогнозирования динамики социальных настроений на основе анализа временных рядов текстового контента социальных сетей с использованием уравнений Фоккера-Планка и нелинейной диффузии».

БИБЛИОГРАФИЯ

- [1] Pulipati S. Somula R. Parvathala B.R. Nature inspired link prediction and community detection algorithms for social networks: a survey International Journal of Systems Assurance Engineering and Management 2021 ISSN 09756809 DOI 10.1007/s13198-021-01125-8
- [2] Cong Tran, Won-Yong Shin, Andreas Spitz Community Detection in Partially Observable Social Networks ACM Transactions on Knowledge Discovery from Data <https://doi.org/10.1145/3461339>
- [3] Zhengdao Chen, Lisha Li, and Joan Bruna. Supervised community detection with line graph neural networks. In Proceedings of the 7th International Conference on Learning Representations. ACM. 2019.
- [4] Till Hoffmann, Leto Peel, Renaud Lambiotte, and Nick S. Jones. Community detection in networks without observing edges. 2020. Science Advances 6, 4 (2020), 1–11.
- [5] C. Gardiner. Stochastic Methods: A Handbook for the Natural and Social Sciences, Springer-Verlag, 2009
- [6] Lux, T., Inference for systems of stochastic differential equations from discretely sampled data: a numerical maximum likelihood approach. Annals of Finance, 2012, 9(2), 217–248, doi 10.1007/s10436-012-0219-9.
- [7] Hum, A., Jeisman, J., Lindsay, K. Teaching an old dog new tricks: improved estimation of the parameters of stochastic differential equations by numerical solution of the Fokker-Planck equation. In: Gregoriou, G., Pascual, R. (eds.) Financial Econometrics Handbook. Palgrave, London, 2010

- [8] Elliott, R. J., Siu, T. K., Chan, L. A PDE approach for risk measures for derivatives with regime switching. Annals of Finance, 2007, 4(1), 55–74. doi: 10.1007/s10436-006-0068-5
- [9] Orlov Yu.N., Fedorov, S.L.: Generation of non-stationary trajectories of a time series based on Fokker-Planck equation. MFTI Proceedings. 8, 2, 126 – 133 (2016) (in Russian: Орлов, Ю.Н., Федоров, С.Л. Генерация нестационарных траекторий временного ряда на основе уравнения Фоккера – Планка. ТРУДЫ МФТИ. 8, 2, 126 – 133 (2016))
- [10] Chen, Y., Cosimano, T.F., Himonas, A.A., Kelly, P. An Analytic Approach for Stochastic Differential Utility for Endowment and Production Economies. Computational Economics, 2013, 44(4), 397–443. doi: 10.1007/s10614-013-9397-4; 2) Orlov and Fedorov, 2016
- [11] Savku, E., Weber, G.-W. Stochastic differential games for optimal investment problems in a Markov regime-switching jump-diffusion market. Annals of Operations Research. 2020 doi: 10.1007/s10479-020-03768-5
- [12] Krasnikov K.E. Mathematical modeling of some social processes using game-theoretic approaches and making managerial decisions based on them. Russian Technological Journal. 2021;9(5):67-83. <https://doi.org/10.32362/2500-316X-2021-9-5-67-83> (In Russ.).
- [13] Zhukov, D., Khvatova, T., Millar, C., and Zaltzman, A. Modelling the stochastic dynamics of transitions between states in social systems incorporating self-organization and memory. Technological Forecasting and Social Change, 2020, 158, 120134. doi:10.1016/j.techfore.2020.120134
- [14] Dmitry O. Zhukov, Anastasia D. Zaltzman, Tatiana Yu. Khvatova. Forecasting Changes in States in Social Networks and Sentiment Security Using the Principles of Percolation Theory and Stochastic Dynamics. Proceedings of the 2019 IEEE International Conference «Quality Management, Transport and Information Security, Information Technologies»; IT and QM and IS 2019, Article number 8928295, pp. 149-153.
- [15] D.O. Zhukov, S.A. Lesko Stochastic self-organisation of poorly structured data and memory realisation in an information domain when designing news events forecasting models. The 2nd IEEE International Conference on Big Data Intelligence and Computing. 2016, August 8-12, Auckland, New Zealand, DOI: 10.1109/DASC-PICom-DataCom-CyberSciTec.2016.153
- [16] Zhukov D.O., Khvatova T.Yu., Lesko S.A., Zaltsman A.D. The influence of the connections' density on clusterization and percolation threshold during information distribution in social networks. Informatika i ee Primeneniya, 2018, Volume 12, Issue 2, Pages 90 - 97.

Об авторах:

Перова Юлия Петровна, старший преподаватель, институт радиоэлектроники и информатики, МИРЭА - Российский технологический университет

Калинин Владимир Николаевич, лаборант, институт радиоэлектроники и информатики, МИРЭА - Российский технологический университет

Лесько Сергей Александрович, кандидат технических наук, доцент, институт кибербезопасности цифровых технологий, МИРЭА - Российский технологический университет

Modeling the dynamics of changes in the number of comments of mass media users based on the Fokker-Planck equation and parameters of networks of their connections

J.P. Perova, V.N. Kalinin, S.A. Lesko

Abstract: In the course of the study, the dynamics of the moods of Internet media users were carried out using the Fokker-Planck equation and using the parameters of comment network graphs. The article introduces a general graph state vector, which includes a high intermediation coefficient, an average value of the clustering coefficient, and the proportion of users in various states. The time dependence of the cosine of the angle between the base and current state vector graphs forms a time series, the values of the levels, which can be interpreted as “wandering points” over a selection of parameters that characterize the state of the network. The current state of the comment graph can be determined using network and text analysis tools. The solution of the Fokker-Planck equation makes it possible to obtain an analytical dependence for the probability density of detecting the magnitude of the state of the network at a given value on the interval of possible values. This approach can be used to create an algorithm for predicting the time to reach a given state of the user comments network graph with a given level of probability. Analysis of the model confirms its adequacy and consistency.

Keywords: social networks, network graph, characteristics of a social network graph, modeling of social processes, nonlinear dynamics, Fokker-Planck equation, monitoring, control.

REFERENCES

- [1] Pulipati S. Somula R. Parvathala B.R. Nature inspired link prediction and community detection algorithms for social networks: a survey International Journal of Systems Assurance Engineering and Management 2021 ISSN 09756809 DOI 10.1007/s13198-021-01125-8
- [2] Cong Tran, Won-Yong Shin, Andreas Spitz Community Detection in Partially Observable Social Networks ACM Transactions on Knowledge Discovery from Data <https://doi.org/10.1145/3461339>
- [3] Zhengdao Chen, Lisha Li, and Joan Bruna. Supervised community detection with line graph neural networks. In Proceedings of the 7th International Conference on Learning Representations. ACM. 2019.
- [4] Till Hoffmann, Leto Peel, Renaud Lambiotte, and Nick S. Jones. Community detection in networks without observing edges. 2020. Science Advances 6, 4 (2020), 1–11.
- [5] C. Gardiner. Stochastic Methods: A Handbook for the Natural and Social Sciences, Springer-Verlag, 2009
- [6] Lux, T., Inference for systems of stochastic differential equations from discretely sampled data: a numerical maximum likelihood approach. Annals of Finance, 2012, 9(2), 217–248, doi 10.1007/s10436-012-0219-9.
- [7] Hurn, A., Jeisman, J., Lindsay, K. Teaching an old dog new tricks: improved estimation of the parameters of stochastic differential equations by numerical solution of the Fokker-Planck equation. In: Gregoriou, G., Pascualau, R. (eds.) Financial Econometrics Handbook. Palgrave, London, 2010
- [8] Elliott, R. J., Siu, T. K., Chan, L. A PDE approach for risk measures for derivatives with regime switching. Annals of Finance, 2007, 4(1), 55–74. doi: 10.1007/s10436-006-0068-5
- [9] Orlov Yu.N., Fedorov, S.L.: Generation of non-stationary trajectories of a time series based on Fokker-Planck equation. MFTI Proceedings. 8, 2, 126 – 133 (2016) (in Russian: Орлов, Ю.Н., Федоров, С.Л. Генерация нестационарных траекторий временного ряда на основе уравнения Фоккера – Планка. ТРУДЫ МФТИ. 8, 2, 126 – 133 (2016))
- [10] Chen, Y., Cosimano, T.F., Himonas, A.A., Kelly, P. An Analytic Approach for Stochastic Differential Utility for Endowment and Production Economics. Computational Economics, 2013, 44(4), 397–443. doi: 10.1007/s10614-013-9397-4; 2) Orlov and Fedorov, 2016
- [11] Savku, E., Weber, G.-W. Stochastic differential games for optimal investment problems in a Markov regime-switching jump-diffusion market. Annals of Operations Research. 2020 doi: 10.1007/s10479-020-03768-5
- [12] Krasnikov K.E. Mathematical modeling of some social processes using game-theoretic approaches and making managerial decisions based on them. Russian Technological Journal. 2021;9(5):67-83. <https://doi.org/10.32362/2500-316X-2021-9-5-67-83> (In Russ.).
- [13] Zhukov, D., Khvatova, T., Millar, C., and Zaltzman, A. Modelling the stochastic dynamics of transitions between states in social systems incorporating self-organization and memory. Technological Forecasting and Social Change, 2020, 158, 120134. doi:10.1016/j.techfore.2020.120134
- [14] Dmitry O. Zhukov, Anastasia D. Zaltzman, Tatiana Yu. Khvatova. Forecasting Changes in States in Social Networks and Sentiment Security Using the Principles of Percolation Theory and Stochastic Dynamics. Proceedings of the 2019 IEEE International Conference «Quality Management, Transport and Information Security, Information Technologies»; IT and QM and IS 2019, Article number 8928295, pp. 149-153.
- [15] D.O. Zhukov, S.A. Lesko Stochastic self-organisation of poorly structured data and memory realisation in an information domain when designing news events forecasting models. The 2nd IEEE International Conference on Big Data Intelligence and Computing. 2016, August 8-12, Auckland, New Zealand, DOI: 10.1109/DASC-PICOM-DataCom-CyberSciTec.2016.153
- [16] Zhukov D.O., Khvatova T.Yu., Lesko S.A., Zaltsman A.D. The influence of the connections' density on clusterization and percolation threshold during information distribution in social networks. Informatika i ee Primeneniya, 2018, Volume 12, Issue 2, Pages 90 - 97.

About the authors:

Perova Yulia Petrovna, Senior Lecturer, Institute of Radio Electronics and Informatics, MIREA - Russian Technological University

Kalinin Vladimir Nikolaevich, laboratory assistant, Institute of Radioelectronics and Informatics, MIREA - Russian Technological University

Lesko Sergey Aleksandrovich, Candidate of Technical Sciences, Associate Professor, Institute of Cybersecurity of Digital Technologies, MIREA - Russian Technological University