ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2022 Управление, вычислительная техника и информатика № 60
Tomsk State University Journal of Control and Computer Science

Научная статья
УДК 519.663
doi: 10.17223/19988605/60/4

Анализ и моделирование процессов в сложных социальных сетевых структурах на основе уравнения Фоккера–Планка

Юлия Петровна Перова[1], Сергей Александрович Лесько[2], Дмитрий Олегович Жуков[3], Алексей Викторович Чечурин[4]

[1, 2, 3, 4] Российский технологический университет (МИРЭА), Москва, Россия
[1] jul-np@yandex.ru
[2] lesko@testor.ru
[3] zhukovdm@yandex.ru
[4] chechurin@testor.ru

Аннотация. Проведено исследование стационарных и динамических распределений новостей по числу комментариев и показано, что наблюдаемый на практике степенной закон зависимости стационарной плотности вероятности распределения новостей по числу комментариев (состояниям системы x) может быть получен из решения стационарного уравнения Фоккера–Планка, если при его выводе сделать ряд допущений: коэффициент $\mu(x)$, отвечающий в уравнении Фоккера–Планка за целенаправленное изменение состояния («снос») системы x (x – текущее число комментариев к новости) линейно зависит от состояния x, а коэффициент $D(x)$, отвечающий за случайное изменение («диффузия»), зависит от x квадратично.

Решение нестационарного дифференциального уравнения Фоккера–Планка при сделанных допущениях позволило получить уравнение для плотности вероятности переходов между состояниями системы в единицу времени, которая хорошо согласуется с наблюдаемыми данными с учетом влияния времени задержки между появлением комментария к новости и комментария к данному комментарию.

Ключевые слова: социальные сети; моделирование социальных процессов; уравнение Фоккера–Планка; мониторинг; управление; нелинейная динамика; степенной закон распределения

***Благодарности*:** Работа выполнена при финансовой 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».

***Для цитирования*:** Перова Ю.П., Лесько С.А., Жуков Д.О., Чечурин А.В. Анализ и моделирование процессов в сложных социальных сетевых структурах на основе уравнения Фоккера–Планка // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 60. С. 32–41. doi: 10.17223/19988605/60/4

Original article
doi: 10.17223/19988605/60/4

Analysis and modeling of processes in complex social network structures based on the Fokker-Planck equation

Julia P. Perova[1], Sergey A. Lesko[2], Dmitry O. Zhukov[3], Alexey V. Chechurin[4]

[1, 2, 3, 4] Russian Technological University (MIREA), Moscow, Russian Federation
[1] jul-np@yandex.ru
[2] lesko@testor.ru

[3] *zhukovdm@yandex.ru*
[4] *chechurin@testor.ru*

Abstract. The paper investigates stationary and dynamic distributions of news by the number of comments and shows that the power law of dependence of the stationary probability density of the distribution of news by the number of comments (states of the system x) observed in practice can be obtained from the solution of the stationary Fokker – Planck equation, if a number of assumptions are made during its derivation: the coefficient $\mu(x)$, responsible in the Fokker – Planck equation for purposeful state change (“demolition”) of system x (x is the current number of comments on the news) linearly depends on the state of x, and the coefficient $D(x)$ responsible for the random change (“diffusion”) depends on x quadratically.

The solution of the unsteady Fokker–Planck differential equation with the assumptions made made it possible to obtain an analytical equation for the probability density of transitions between the states of the system per unit of time, which is in good agreement with the observed data, taking into account the effect of the delay time between the appearance of a comment to the news and a comment to this comment.

Keywords: social networks; modeling of social processes; Fokker-Planck equation; monitoring; management; nonlinear dynamics; power law of distribution

***Acknowledgments*:** The work was carried out with the financial support of the Russian Science Foundation (RNF), grant No. 22-21-00109 “Development of models for predicting the dynamics of social moods based on the analysis of time series of text content of social networks using the Fokker-Planck equations and nonlinear diffusion”.

***For citation*:** Perova, J.P., Lesko, S.A., Zhukov, D.O., Chechurin, A.V. (2022) Analysis and modeling of processes in complex social network structures based on the Fokker-Planck equation . *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitelnaja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 60. pp. 32–41. doi: 10.17223/19988605/60/4

Описание поведения пользователей социальных сетей и информационных ресурсов является одним из важнейших направлений математической социологии. С практической точки зрения создание моделей, описывающих динамику проявления пользовательских мнений и предпочтений, позволяет разрабатывать системы автоматизированного мониторинга общественного настроения и тенденций его изменения. Преимуществом таких систем по сравнению с традиционными методами изучения общественного мнения является их технологичность в реализации.

Следует отметить, что динамику изменения мнений и настроений пользователей сети Интернет можно в значительной степени отнести к стохастическим процессам. Присутствие человеческого фактора (множество людей с различными мнениями, предпочтениями и характером поведения), с одной стороны, создает случайность изменений (в силу большого разнообразия поведенческих моделей пользователей), а с другой – вносит в динамику изменений элементы целенаправленности.

В статье [1] рассмотрена модель, описывающая пространственное и временное распространение информации в социальных сетях на основе стохастического дифференциального уравнения в частных производных. В работе была создана и исследована неавтономная диффузионная логистическая модель с граничными условиями Дирихле, которая показала, что на диффузию данных в социальных сетях сильно влияют коэффициент диффузии и внутренняя скорость роста (распространение информации или слухи можно рассматривать как своего рода вирусы, не обладающие физической формой).

В этой связи наиболее перспективными, на наш взгляд, для анализа динамики изменения общественного настроения являются модели, которые можно создать на основе стохастических дифференциальных уравнений, например уравнения Фоккера–Планка, которое учитывает как упорядоченные («снос»), так и случайные изменения («диффузия»). Уравнение Фоккера–Планка широко применяется для анализа и моделирования поведения временных рядов при описании процессов в сложных системах [2–5].

Следует отметить, что помимо уравнения Фоккера–Планка для моделирования на основе дифференциальных уравнений используются и другие подходы, например уравнения Лиувилля [5, 6], уравнения диффузии [4, 7] и др.

Помимо описания динамических процессов из уравнения Фоккера–Планка можно получить и стационарные решения, которые могут описывать состояние какой-либо системы в стационарном состоянии, когда, например, ее эволюция уже закончилась и изменения не происходят.

Исследование процессов, происходящих в сложных системах с участием человеческого фактора, показывает, что очень часто для наблюдаемых характеристик параметров этих процессов выполняется степенной закон распределения $\rho(x) \sim x^{-\gamma}$ (где γ – характеристическая степень) [8–13], но в то же время вопрос теоретического обоснования возможности его применения требует дальнейшего изучения. На наш взгляд, это обоснование очень важно. Выявление характера процессов, из которых возникает степенной закон, необходимо для более глубокого изучения поведения и анализа сложных социальных систем.

В связи с этим мы считаем перспективным исследование возможности применения уравнения Фоккера–Планка для разработки моделей динамики социальных процессов.

Результаты, полученные в работе:

1. Наблюдаемое на практике стационарное распределение новостей по числу комментариев к ним соответствует степенному закону: $\rho(x) = [\gamma - 1]x^{-\gamma}$, где $\rho(x)$ – доля новостей в общем их числе, имеющая x комментариев, а γ – показатель степени.

2. Динамика изменения с течением времени числа комментариев к новости или блогу может иметь как S-образный вид, так и двухступенчатый, что может быть связанно с существенным различием в среднем времени появления комментариев второго уровня (интервал времени между появлением комментария первого уровня и комментария к данному комментарию), т.е. величиной средней задержки.

3. Наблюдаемый на практике степенной закон зависимости стационарной плотности вероятности распределения новостей по числу комментариев (состояниям x) может быть получен из решения стационарного уравнения Фоккера–Планка, если при его выводе сделать ряд допущений: в частности, предположить, что коэффициент $\mu(x)$, отвечающий в уравнении Фоккера–Планка за целенаправленное изменение состояния системы x (x – текущее число комментариев к новости) линейно зависит от состояния x, а коэффициент $D(x)$, отвечающий за случайное изменение, зависит от x квадратично. Все это позволяет предположить, что уравнение Фоккера–Планка может быть использовано для описания процессов в сложных сетевых структурах.

4. Решение нестационарного уравнения Фоккера–Планка при допущениях о линейной зависимости $\mu(x)$ от состояния x и квадратичной зависимости $D(x)$ от состояния x позволяет получить уравнение для плотности вероятности переходов между состояниями системы в единицу времени, которая хорошо согласуется с наблюдаемыми данными с учетом влияния времени задержки между появлением комментария первого уровня и комментария к данному комментарию.

5. Разработанные на основе уравнения Фоккера–Планка модели хорошо согласуются с наблюдаемыми данными, что позволяет создать алгоритмы мониторинга и прогнозирования эволюции общественного мнения пользователей новостных информационных ресурсов.

В заключение отметим, что сложный характер динамики процессов в сложных социальных системах можно описывать не только на основе моделей, созданных на основе уравнения Фоккера–Планка. Например, в работах [14–17] представлены разработанные авторами модели описания стохастической динамики изменения состояний в сложных социальных системах, учитывающие процессы самоорганизации и наличие памяти (немарковские модели).

1. Сбор и обработка данных

Для наших исследований были выбраны несколько новостных порталов и одно из сетевых сообществ социальной сети «ВКонтакте», посвященное обсуждению новостей информационного ресурса РИА «Новости» (https://vk.com/ria). РИА «Новости» было выбрано, исходя из его узнаваемости и популярности в российском обществе, оно занимает первое место среди медиаресурсов (за март 2022 г.) по версии br-analytics (https://br-analytics.ru/mediatrends/media/?period=202203), входит в ТОП-3 самых цитируемых информационных агентств в СМИ и социальных медиа (за март 2022 г.), где занимает первое место (https://www.mlg.ru/ratings/media/federal/11110/#internet).

Сначала мы с помощью специального программного приложения (парсера) скачали интересующий нас новостной контент начиная с 1 января 2019 г. по апрель 2022 г. с ресурса «ВКонтакте», используя разработанный нами парсер и API сети (https://dev.vk.com/guide). Внутри социальной сети у каждого поста есть свой уникальный адрес (https://vk.com/ria?w=wall{owner_id}_{post_id}), где {owner_id} – уникальный идентификатор сообщества (в случае РИА «Новости» это «-15755094», а {post_id} – уникальный идентификатор поста (новости). Каждый пост (новость) имеет ряд основных параметров: уникальный идентификатор поста в социальной сети; текст поста; дата и время публикации; количество просмотров и комментариев пользователей. Комментарии, в свою очередь, имеют следующие параметры: уникальный идентификатор в сообществе социальной сети; уникальный идентификатор пользователя; текст комментария; дата и время появления; уровень иерархии комментария; связь по уровню комментирования с родительским комментарием (кто из пользователей комментировал кого из других пользователей при обсуждении новости).

Поскольку комментарии могли оставлять чат-боты, спамеры и недобросовестные пользователи, которые пишут комментарии на профессиональной основе, необходимо было ввести правила отчистки данных. К недобросовестным были отнесены те, кто написал за год более 7 365 комментариев (в среднем более 20 за сутки) или писал с частотой более одного комментария в 5 минут.

При анализе полученных данных необходимо было определить, какому закону распределения подчиняется наблюдаемая плотность распределения. Были рассмотрены три наиболее часто наблюдаемых закона распределения: Гаусса $\rho(x) = e^{-\frac{x^2}{2\sigma^2}}/\sigma\sqrt{2\pi}$, экспоненциальный $\rho(x) = ae^{-\alpha x}$ и степенной $\rho(x) = \beta x^{-\gamma}$. При обработке собранных данных с помощью линеаризации в соответствующих координатах обнаружено, что лучшая линеаризация наблюдается для степенного закона распределения (рис. 1), для остальных законов линеаризация была плохой.

Прямая, проведенная на рис. 1, показывает, что линия тренда хорошо описывается выбранной нами линейной аппроксимацией $y = -0{,}76 - 1{,}48z$, где $y = \ln\{\rho(x)\}$, $z = \ln\{x\}$, $\ln\{\beta\} = -0{,}76$, а коэффициент корреляции равен 0,95.

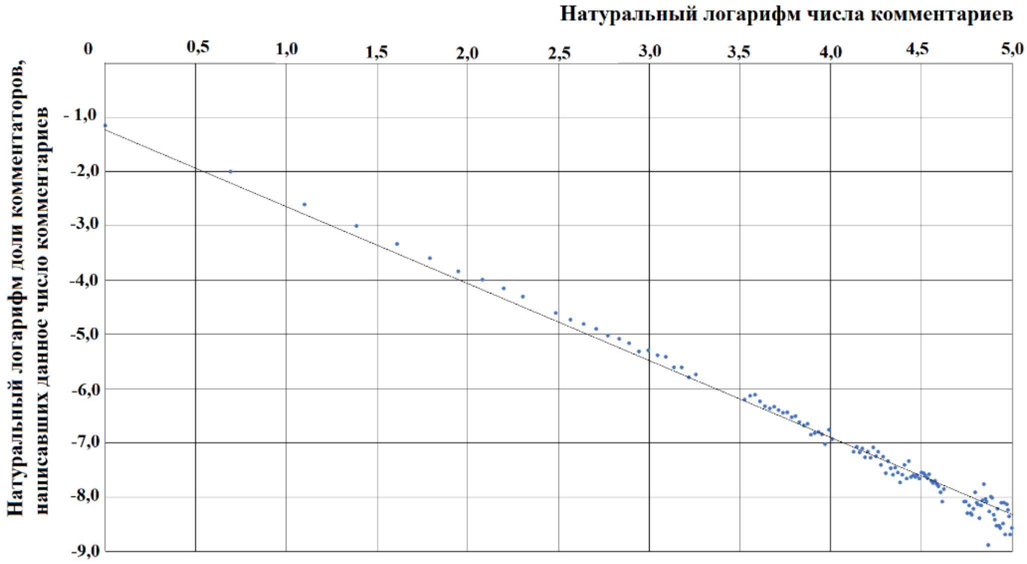


Рис. 1. Линеаризация наблюдаемых данных для степенного распределения доли комментаторов от числа сделанных ими комментариев
Fig. 1. Linearization of the observed data for the power distribution of the proportion of commentators from the number of comments they made

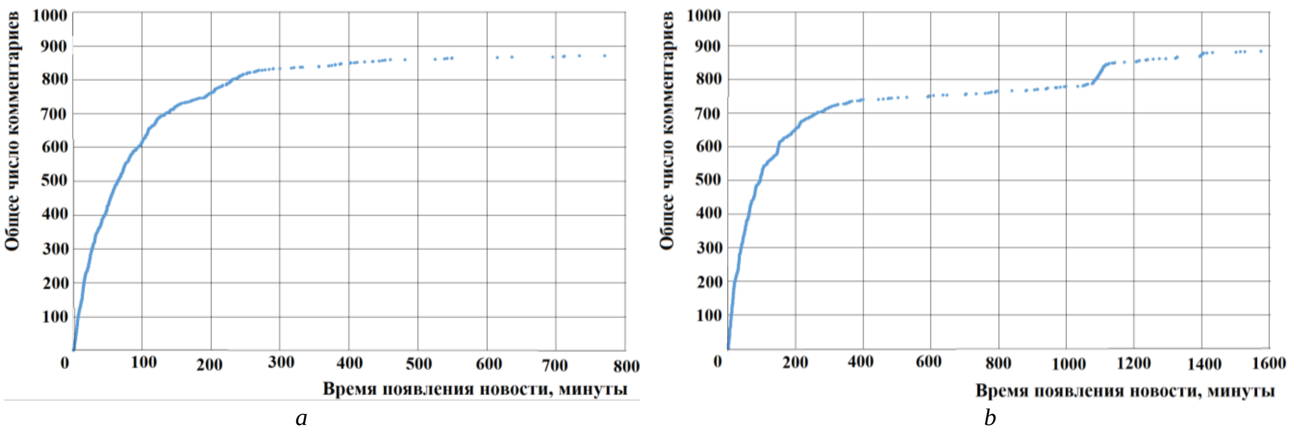
Для подтверждения вывода о линейной аппроксимации можно исследовать поведение остатков и проверить гипотезу о том, что они нормально распределены со средним значением, равным нулю, и имеют однородную дисперсию. Вычисление остатков можно провести на основе реально наблюдаемых значений натурального логарифма доли комментаторов, давших данное число комментариев, и полученного нами уравнения. Рассчитанная величина математического ожидания для распределения остатков равна 0,25, а дисперсия 0,13. Проверка гипотезы о наклоне (двухвыборочный F-тест для дисперсий) показывает, что дисперсия остатков, рассчитанная относительно линии тренда, существенно меньше, чем дисперсия отклонения точек линейной регрессии от среднего значения величины наблюдаемых данных ($\Sigma y_i/n = \Sigma\{\rho(\ln x_i)\}/n$): Она равна 2,11 (0,13 << 2,11). Таким образом, из полученных данных можно сделать вывод, что распределение остатков очень близко к нормальному и выявленая регрессия является значимой, что подтверждает вывод о том, что натуральный логарифм доли комментаторов, написавших данные комментарии, линейно зависит от натурального логарифма числа комментариев, что подтверждает выполнение степенного закона.

При проведении исследования также представляло интерес рассмотрение динамики изменения числа комментариев к новостям, привлекающим большое общественное внимание (за время просмотра такие новости или блоги набирают сотни комментариев), с течением времени.

Наблюдение динамики изменения числа комментариев к новости показывает, что она может иметь как S-образный, так и двухступенчатый характер. В качестве примера такой динамики комментирования пользователями «ВКонтакте» новостей ресурса РИА «Новости» выберем несколько публикаций:

1. «Зеленский покинул Украину и переехал в Польшу, – заявил Володин» (динамика комментирования имеет S-образный характер; рис. 2, *a*; https://vk.com/ria?w=wall-15755094_34243579; https://ria.ru/20220304/zelenskiy-1776545154.html). Дата и время появления: 2022-03-04 16:13:27 UTC +03:00). Общее число комментариев составило 894. Число комментариев первого уровня (комментарий самой новости) составило 433, второго (комментарии комментариев первого уровня) 461. Общее количество просмотров – 118 764. Среднее время появления комментариев первого уровня 73 минуты, а второго уровня – 74.

2. «Скоро город будет освобожден, – заявил он» (динамика комментирования имеет двухступенчатый характер; рис. 2, *b*; https://vk.com/ria?w=wall-15755094_35202266; https://ria.ru/20220410/ukraina-1782778315.html). Дата и время появления: 2022-04-10 17:14:40 UTC +03:00). Общее число комментариев составило 901. Число комментариев первого уровня (комментарий самой новости) составило 173, второго (комментарии комментариев первого уровня) 728. Общее количество просмотров – 173 607. Среднее время появления комментариев первого уровня – 75 минуты, а второго уровня – 82.



a *b*

Рис. 2. Наблюдаемая динамика изменения числа комментариев к новостям
Fig. 2. The observed dynamics of changes in the number of comments to the news

На наш взгляд, это может быть связано как с различием в среднем времени появления комментариев второго уровня (интервал времени между появлением комментария первого уровня и комментария к данному комментарию), так и с соотношением между числом комментариев первого и второго уровней. Если для первой новости интервалы среднего времени появления комментариев первого и второго уровней практически совпадают, то для второй наблюдается небольшое увеличение интервала среднего времени появления комментариев второго уровня (происходит запаздывание по времени). Кроме того, для второй новости их число существенно превосходит число комментариев первого уровня.

Для дальнейшего изучения можно сформулировать следующую задачу теоретического исследования: какова природа процессов комментирования новостей и блогов и какие особенности этих сложных социальных систем приводят к тому, что для зависимости плотности вероятности распределения комментариев по их числу выполняется степенной закон, а динамика имеет во многих случаях сложный двухступенчатый характер?

2. Решение стационарного уравнения Фоккера–Планка

В общем виде уравнение Фоккера–Планка имеет вид:

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

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 («диффузия»), $\mu(x)$ – зависящий от состояния x коэффициент, определяющий целенаправленное изменение состояния x («снос»). Применительно к нашей модели $D(x)$ можно трактовать как действия пользователя, вызванные спонтанным импульсом, возникшим при прочтении новости или комментариев к ней других пользователей, когда описываемое в новости или блоге событие не является существенно важным, но пользователь готов потратить время на комментарий или ответить другому комментатору (у пользователя возникло спонтанное желание отреагировать на данную новость). Коэффициент $\mu(x)$ можно интерпретировать как целенаправленные действия, вызванные желанием отреагировать на существенно важную для пользователя новость или блог, а также дать комментарий на комментарий другого пользователя, если он затронул важную с точки зрения данного пользователя тему (пользователь постоянно интересуется данной темой). Для построения модели необходимо сделать предположение о зависимости $D(x)$ и $\mu(x)$ от состояния x и рассмотреть два условия: во-первых, учтем размерность членов, входящих в уравнение (1), а во-вторых, можно сделать предположение, что с ростом состояния x (ростом числа возможных комментариев (значимости новости или блога) величины $D(x)$ и $\mu(x)$ также должны увеличиваться. Логика подсказывает, что все члены уравнения (1) должны иметь одинаковую размерность, которую имеет $\rho(x)$. Оба условия будут выполнены, если зависимости $D(x)$ и $\mu(x)$ от состояния x будут иметь вид: $\mu(x) = \mu_0 x$ и $D(x) = D_0 x^2$.

Решение стационарного уравнения Фоккера–Планка $-\frac{d}{dx}[\mu(x)\rho(x)] + \frac{1}{2}\frac{d^2}{dx^2}[D(x)\rho(x)] = 0$ при допущениях $\mu(x) = \mu_0 x$ и $D(x) = D_0 x^2$ имеет вид: $\rho(x) = [\gamma - 1]x^{-\gamma}$, что соответствует наблюдаемому на практике степенному закону распределения. По полученным из анализа наблюдаемых данных результатам $\gamma = 1{,}48$, $\gamma - 1 = 0{,}48$, натуральный логарифм $(\gamma - 1)$ равен $-0{,}73$, что с достаточно высокой точностью равно $\ln\{\beta\} = -0{,}76$ (см. полученное уравнение линеаризации: $y = -0{,}76 - 1{,}48z$, где $y = \ln\{\rho(x)\}$, $z = \ln\{x\}$, $\ln\{\beta\} = \ln\{\gamma - 1\} - 0{,}76$, а коэффициент корреляции равен 0,95). В целом это указывает на адекватность разработанной модели.

3. Решение нестационарного уравнения Фоккера–Планка и анализ модели

Полученное нами решение нестационарного уравнения Фоккера–Планка при сделанных для $\mu(x)$ и $D(x)$ допущениях имеет вид:

$$\rho(x,t) = \int \frac{\left[\frac{[\ln(x)]^2}{D_0 t} + \left[\frac{1}{2} - \frac{\mu_0}{D_0}\right]\ln(x) - 1\right]}{\sqrt{2\pi D_0 t^3}} e^{-\left[\frac{[\ln(x)]^2}{2D_0 t} + \left[\frac{3}{2} - \frac{\mu_0}{D_0}\right]\ln(x) + \left[\frac{1}{2} - \frac{\mu_0}{D_0}\right]^2 \frac{D_0 t}{2}\right]} dt. \qquad (2)$$

Вероятность того, что число комментариев к моменту времени t достигнет некоторого числа L, можно найти по формуле $P(L,t) = 1 - \int_0^L \rho(x,t)dx$. Зависимость числа комментариев $N(t)$ от времени t будет описываться уравнением $N(t) = P(L, t)L$.

Для анализа полученного решения проведем имитационное моделирование. В качестве примера выберем $L = 100$ и три набора значений μ_0 и D_0 ($\mu_0 = 0{,}45$ и $D_0 = 0{,}50$ условных единиц ($\mu_0 < D_0$; см.

кривая *1* на рис. 3), $\mu_0 = 0{,}50$ и $D_0 = 0{,}50$ условных единиц ($\mu_0 = D_0$; см. кривая *2* на рис. 3) и $\mu_0 = 0{,}55$ и $D_0 = 0{,}50$ условных единиц ($\mu_0 > D_0$; см. кривая *3* на рис. 3). Расчеты показывают, что с ростом μ_0 относительно D_0 скорость роста кривых для числа комментариев $N(t)$ при выбранных значениях параметров модели μ_0, D_0 и L увеличивается (см. рис. 3).

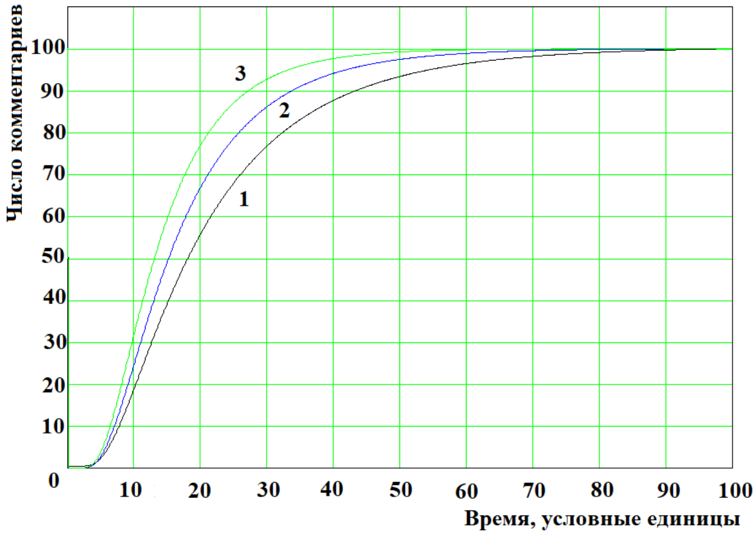


Рис. 3. Динамика изменения с течением времени числа комментариев к новости в имитационной модели на основе уравнения Фоккера–Планка
Fig. 3. Dynamics of changes over time in the number of comments to the news in the simulation model based on the Fokker-Planck equation

Двухступенчатую кривую можно получить, если использовать функцию плотности распределения, учитывающую время задержки τ:

$$\rho\,(x, t-\tau) = \int \frac{\left[\frac{[ln(x)]^2}{D_0[t-\tau]} + \left[\frac{1}{2} - \frac{\mu_0}{D_0}\right] ln(x) - 1\right]}{\sqrt{2\pi D_0[t-\tau]^3}}\, e^{-\left[\frac{[ln(x)]^2}{2D_0[t-\tau]} + \left[\frac{3}{2} - \frac{\mu_0}{D_0}\right] ln(x) + \left[\frac{1}{2} - \frac{\mu_0}{D_0}\right]^2 \frac{D_0[t-\tau]}{2}\right]} dt. \tag{3}$$

Соответствие теоретической модели и наблюдаемых данных (см. рис. 2, 3) можно получить, если предположить, что могут одновременно протекать два процесса с различными μ_0 и D_0. Причем сумма парциальных долей процессов должна быть равна 1, т.е. $P_{общ}(L,t) = \alpha_1 P_1(L,t) + \alpha_2 \cdot P_2(L,t)$.

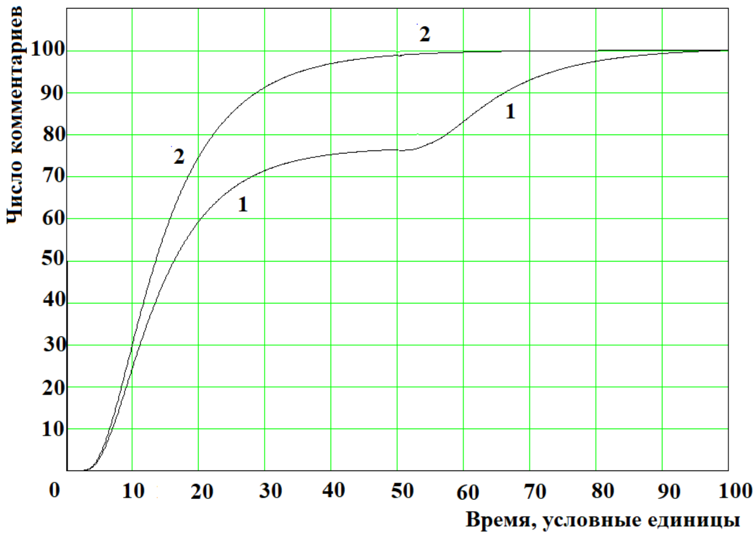


Рис. 4. Динамика изменения с течением времени числа комментариев к новости в имитационной модели на основе уравнения Фоккера–Планка с учетом двух параллельных процессов
Fig. 4. Dynamics of changes over time in the number of comments on the news in the simulation model based on the Fokker-Planck equation, taking into account two parallel processes

В качестве примера моделирования выберем для процесса комментирования самой новости или блога следующие параметры модели: $\mu_{0,1} = 0{,}55$, $D_{0,1} = 0{,}50$, а для процесса комментирования комментариев $\mu_{0,2} = 0{,}50$, $D_{0,2} = 0{,}50$, $\tau = 50$ условных единиц, $\alpha_1 = 0{,}75$, $\alpha_2 = 0{,}25$ ($\alpha_1 + \alpha_2 = 1$), $L = 100$ ($\mu_{0,1} > \mu_{0,2}$ было выбрано исходя из предположения, что комментирование новости является более первичным процессом для пользователей, чем комментирование комментариев).

На рис. 4 представлены результаты моделирования динамики изменения с течением времени числа комментариев $N(t)$ с учетом того, что могут параллельно протекать два процесса. Как видно из результатов моделирования, представленных на рис. 4, наблюдается хорошее совпадение реальных данных и теоретических расчетов.

Заключение

В работе проведено исследование стационарных и динамических распределений новостей по числу комментариев. Обработка данных, полученных с информационного портала, показала, что статическое распределение новостей по числу комментариев к ним подчиняется степенному закону $\rho(x) \sim x^{-\gamma}$ (где γ – характеристическая степень), а динамическое распределение (изменение числа комментариев с течением времени) в ряде случаев имеет S-образный характер, а иногда и более сложный – двухступенчатый; это зависит от величины среднего времени появления комментариев второго уровня (интервал времени между появлением комментария первого уровня и комментария к данному комментарию), т.е. величины средней задержки.

В статье показано, что наблюдаемый на практике степенной закон зависимости стационарной плотности вероятности распределения новостей по числу комментариев (состояниям x) может быть получен из решения стационарного уравнения Фоккера–Планка $-\frac{d}{dx}[\mu(x)\rho(x)] + \frac{1}{2}\frac{d^2}{dx^2}[D(x)\rho(x)] = 0$), если при его выводе сделать ряд допущений: в частности, предположить, что коэффициент $\mu(x)$, отвечающий в уравнении Фоккера–Планка за целенаправленное изменение («снос») состояния системы x (x – текущее число комментариев к новости) линейно зависит от состояния x, а коэффициент $D(x)$, отвечающий за случайное изменение («диффузия»), зависит от x квадратично.

Решение нестационарного дифференциального уравнения Фоккера–Планка при сделанных допущениях позволило получить аналитическое уравнение для плотности вероятности переходов между состояниями системы в единицу времени

$$\rho(x,t) = \int \frac{\left[\frac{[ln(x)]^2}{D_0 t} + \left[\frac{1}{2} - \frac{\mu_0}{D_0}\right] ln(x) - 1\right]}{\sqrt{2\pi D_0 t^3}} e^{-\left[\frac{[ln(x)]^2}{2D_0 t} + \left[\frac{3}{2} - \frac{\mu_0}{D_0}\right] ln(x) + \left[\frac{1}{2} - \frac{\mu_0}{D_0}\right]^2 \frac{D_0 t}{2}\right]} dt,$$

которое хорошо согласуется с наблюдаемыми данными с учетом влияния времени задержки между появлением комментария первого уровня и комментария к данному комментарию.

Список источников

1. Du B., Lian X., Cheng X. Partial differential equation modeling with Dirichlet boundary conditions on social networks // Boundary Value Problems. V. 2018, is. 1. Art. 50.
2. Lux T. Inference for systems of stochastic differential equations from discretely sampled data: a numerical maximum likelihood approach // Annals of Finance. 2012. V. 9 (2), P. 217–248.
3. 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 // Financial Econometrics Handbook / G. Gregoriou, R. Pascalau (eds.). London : Palgrave, 2010
4. Elliott R.J., Siu T.K., Chan L.A PDE approach for risk measures for derivatives with regime switching // Annals of Finance. 2007. V. 4(1). P. 55–74.
5. Орлов Ю.Н., Федоров С.Л. Генерация нестационарных траекторий временного ряда на основе уравнения Фоккера–Планка // Труды МФТИ. 2016. № 8 (2). С. 126–133.
6. 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. V. 44 (4). P. 397–443.
7. Savku E., Weber G.-W. Stochastic differential games for optimal investment problems in a Markov regime-switching jump-diffusion market // Annals of Operations Research. 2022. V. 312. P. 1171–1196. doi: 10.1007/s10479-020-03768-5

8. Dorogovtsev S.N., Mendes J.F.F. Evolution of networks // Adv. Phys. 2002. V. 51. P. 1079–1187.
9. Newman M.E.J. The structure and function of complex networks // SIAM Rev. 2003. V. 45. P. 167–256.
10. Dorogovtsev S.N., Mendes J.F.F., Samukhin A.N. Generic scale of the scale-free growing networks // Phys. Rev. E. 2001. V. 63. Art. 062101.
11. Golder S.A., Wilkinson D.M., Huberman B.A. Rhythms of social interaction: messaging within a massive online network // Communities and Technologies : Proc. of the 3rd Communities and Technologies Conference, Michigan State University. 2007. P. 41–66.
12. Kumar R., Novak J., Tomkins A. Structure and evolution of online social networks // Proc. of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD'06. 2006. P. 611–617.
13. Mislove A., Marcon M., Gummadi K.P., Druschel P., Bhattacharjee B. Measurement and analysis of online social networks // Proc. of the 7th ACM SIGCOMM Conference on Internet Measurement. IMC'07. 2007. P. 29–42.
14. Sigov A.S., Zhukov D.O., Khvatova T.Yu., Andrianova E.G. A Model of Forecasting of Information Events on the Basis of the Solution of a Boundary Value Problem for Systems with Memory and Self-Organization // Journal of Communications Technology and Electronics. 2018. V. 18, № 2. P. 106–117.
15. Zhukov D., Khvatova T., Millar C., Zaltcman A. Modelling the stochastic dynamics of transitions between states in social systems incorporating self-organization and memory // Technological Forecasting and Social Change. 2020. V. 158. Art. 120134.
16. Zhukov D.O., Lesko S.A. 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
17. Zhukov D.O., Zaltcman A.D., Khvatova T.Yu. 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. Art. 8928295. P. 149–153.

References

1. Du, B., Lian, X. & Cheng, X. (2018) Partial differential equation modeling with Dirichlet boundary conditions on social networks. *Boundary Value Problems*. 2018(1). No 50. DOI: 10.1186/s13661-018-0964-4
2. Lux, T. (2012) Inference for systems of stochastic differential equations from discretely sampled data: a numerical maximum likelihood approach. *Annals of Finance*. 9(2). pp. 217–248. DOI: 10.1007/s10436-012-0219-9
3. Hurn, A., Jeisman, J. & Lindsay, K. (2010) 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. & Pascalau, R. (eds) *Financial Econometrics Handbook*. London: Palgrave.
4. Elliott, R.J., Siu, T.K. & Chan, L. (2007) A PDE approach for risk measures for derivatives with regime switching. *Annals of Finance*. 4(1). pp. 55–74. DOI: 10.1007/s10436-006-0068-5
5. Orlov, Yu.N. & Fedorov, S.L. (2016) Generation of nonstationary time series trajectories based on the Fokker-Planck equation. *Trudy MFTI*. 8(2). pp. 126–133.
6. Chen, Y., Cosimano, T.F., Himonas, A.A. & Kelly, P. (2013) An Analytic Approach for Stochastic Differential Utility for Endowment and Production Economies. *Computational Economics.* 44(4). pp. 397–443. DOI: 10.1007/s10614-013-9397-4
7. Savku, E. & Weber, G.-W. (2020) Stochastic differential games for optimal investment problems in a Markov regime-switching jump-diffusion market. *Annals of Operations Research*. 312. pp. 1171–1196. DOI: 10.1007/s10479-020-03768-5
8. Dorogovtsev, S.N. & Mendes, J.F.F. (2002) Evolution of networks. *Advances of Physics*. 51. pp. 1079–1187.
9. Newman, M.E.J. (2003) The structure and function of complex networks. *SIAM Review*. 45. pp. 167–256. DOI: 10.1137/S003614450342480
10. Dorogovtsev, S.N., Mendes, J.F.F. & Samukhin, A.N. (2001) Generic scale of the scale-free growing networks. *Physical Review E.* 63. 062101. DOI: DOI: 10.1103/PhysRevE.63.062101
11. Golder, S.A., Wilkinson, D.M. & Huberman, B.A. (2007) Rhythms of social interaction: messaging within a massive online network. In: Steinfield, C., Pentland, B.T., Ackerman, M., Contractor, N. (eds) *Communities and Technologies*. London: Springer. pp. 41–66. DOI: 10.1007/978-1-84628-905-7_3
12. Kumar, R., Novak, J. & Tomkins, A. (2006) Structure and evolution of online social networks. *Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining. KDD '06*. pp 611–617.
13. Mislove, A., Marcon, M., Gummadi, K.P., Druschel, P. & Bhattacharjee, B. (2007) Measurement and analysis of online social networks. *Proceedings of the 7th ACM SIGCOMM conference on internet measurement. IMC '07*. pp. 29–42.
14. Sigov, A.S., Zhukov, D.O., Khvatova, T.Yu. & Andrianova, E.G. (2018) A Model of Forecasting of Information Events on the Basis of the Solution of a Boundary Value Problem for Systems with Memory and Self-Organization. *Journal of Communications Technology and Electronics*. 18(2). pp.106–117. DOI: 10.1134/S1064226918120227
15. Zhukov, D., Khvatova, T., Millar, C. & Zaltcman, A. (2020) Modelling the stochastic dynamics of transitions between states in social systems incorporating self-organization and memory. *Technological Forecasting and Social Change*. 158. 120134. DOI: 10.1016/j.techfore.2020.120134
16. Zhukov, D.O. & Lesko, S.A. (2016) 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*. August 8–12, Auckland, New Zealand. DOI: 10.1109/DASC-PICom-DataCom-CyberSciTec.2016.153

17. Zhukov, D.O., Zaltcman, A.D. & Khvatova, T.Yu. (2019) 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 No 8928295. pp. 149–153.

***Информация об авторах*:**
Перова Юлия Петровна – старший преподаватель кафедры телекоммуникаций Института радиоэлектроники и информатики Российского технологического университета – МИРЭА (Москва, Россия). E-mail: jul-np@yandex.ru.
Лесько Сергей Александрович – кандидат технических наук, доцент кафедры моделирования систем Института кибербезопасности и цифровых технологий Российского технологического университета – МИРЭА (Москва, Россия). E-mail: lesko@testor.ru
Жуков Дмитрий Олегович – профессор, доктор технических наук, профессор кафедры информационного противоборства Института кибербезопасности и цифровых технологий Российского технологического университета – МИРЭА (Москва, Россия). E-mail: zhukovdm@yandex.ru.
Чечурин Алексей Викторович – кандидат технических наук, доцент кафедры моделирования систем Института кибербезопасности и цифровых технологий Российского технологического университета – МИРЭА (Москва, Россия). E-mail: chechurin@testor.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

***Information about the authors*:**
Perova Julia P. (Senior lecturer of the Department of Telecommunications, Institute of Radio Electronics and Computer Science, MIREA – Russian Technological University, Moscow, Russian Federation). E-mail: jul-np@yandex.ru
Lesko Sergey. A. (Candidate of Technical Sciences, Associate Professor, Department of Systems Modeling, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University, Moscow, Russian Federation). E-mail: lesko@testor.ru
Zhukov Dmitry O. (Doctor of Technical Sciences, Professor, Department of Information Warfare, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University, Moscow, Russian Federation). E-mail: zhukovdm@yandex.ru
Chechurin Alexey V. (Candidate of Technical Sciences, Associate Professor, Department of Systems Modeling, Institute of Cybersecurity and Digital Technologies, MIREA – Russian Technological University, Moscow, Russian Federation). E-mail: chechurin@testor.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 29.03.2022; принята к публикации 30.08.2022

Received 29.03.2022; accepted for publication 30.08.2022