

УДК: 519.217, 32.019.52

Стохастическая модель числа сторонников политического лидера в цифровом публичном пространстве

К. В. Светлов^а, С. А. Иванов^б

Лаборатория изучения социально-экономических и политических процессов современного общества,
Санкт-Петербургский государственный университет (ФГБОУ ВО СПбГУ),
Россия, 191124, г. Санкт-Петербург, ул. Смольного, д. 1/3, 9-й подъезд

E-mail: ^а k.svetlov@spbu.ru, ^б s.a.ivanov@spbu.ru

Получено 12.05.2019, после доработки — 22.07.2019.

Принято к публикации 21.08.2019.

В представленной статье мы исследуем процесс изменения рейтинга одобрения политического лидера под влиянием процессов, протекающих в цифровом публичном пространстве. Драйвером указанных изменений служит взаимодействие пользователей онлайн-площадок (информационных и новостных ресурсов, блогов, социальных сетей), в результате которого они могут обмениваться друг с другом мнениями и формулировать свою позицию в отношении политика. Помимо межличностного взаимодействия мы рассмотрим такие факторы, как информационное воздействие, выражающееся в создании информационного потока, имеющего заданную мощность и тональность (положительную или отрицательную, в контексте влияния на имидж политического лидера), а также наличие группы агентов (лидеров мнений), оказывающих поддержку политике или же, наоборот, негативно влияющих на его представление в медийном пространстве.

Математической основой представленного исследования является модель Кирмана, имеющая истоки в биологии и первоначально нашедшая свое применение в экономике. В рамках данной модели считается, что каждый участник находится в одном из двух возможных состояний, а также задается скачкообразный марковский процесс, описывающий переходы между этими состояниями. Для рассматриваемой нами задачи данными состояниями являются 0 или 1, в зависимости от того, является ли конкретный агент сторонником политика и одобряет его деятельность или же нет. Пользуясь аппаратом теории марковских процессов, мы находим его диффузионное приближение, известное как процесс Якоби. При помощи спектрального разложения для инфинитезимального оператора данного процесса мы имеем возможность найти аналитическое представление для плотности переходных вероятностей.

Анализируя вероятности, полученные указанным образом, можно оценить влияние отдельных факторов модели: мощность и тональность новостных сообщений, доступных для пользователей онлайн-пространства и релевантных для задач формирования рейтинга, а также численности сторонников или противников политика. Далее, пользуясь найденными собственными функциями и значениями, мы выводим выражения для оценки условных математических ожиданий рейтинга политика, что может служить основой для построения прогнозов, важных для задач формирования стратегии представления политического лидера в онлайн-среде.

Ключевые слова: рейтинг одобрения, политическое лидерство, информационное воздействие, стадное поведение, марковский процесс

Исследование выполнено при финансовой поддержке РФФИ и АНО ЭИСИ в рамках научного проекта № 19-011-31651 «Коммуникативные стратегии российских политических лидеров в онлайн-пространстве».

UDC: 519.217, 32.019.52

Stochastic model of voter dynamics in online media

K. V. Svetlov^a, S. A. Ivanov^b

Laboratory for research of social-economic and political processes of modern society,
St. Petersburg State University,
9th entrance, 1/3 Smolny st., St. Petersburg, 191124, Russia

E-mail: ^a k.svetlov@spbu.ru, ^b s.a.ivanov@spbu.ru

Received 12.05.2019, after completion — 22.07.2019.

Accepted for publication 21.08.2019.

In the present article we explore the process of changing the level of approval of a political leader under the influence of the processes taking place in online platforms (social networks, forums, etc.). The driver of these changes is the interaction of users, through which they can exchange opinions with each other and formulate their position in relation to the political leader. In addition to interpersonal interaction, we will consider such factors as the information impact, expressed in the creation of an information flow with a given power and polarity (positive or negative, in the context of influencing the image of a political leader), as well as the presence of a group of agents (opinion leaders), supporting the leader, or, conversely, negatively affecting its representation in the media space.

The mathematical basis of the presented research is the Kirman model, which has its roots in biology and initially found its application in economics. Within the framework of this model it is considered that each user is in one of the two possible states, and a Markov jump process describing transitions between these states is given. For the problem under consideration, these states are 0 or 1, depending on whether a particular agent is a supporter of a political leader or not. For further research, we find its diffusional approximation, known as the Jacoby process. With the help of spectral decomposition for the infinitesimal operator of this process we have an opportunity to find an analytical representation for the transition probability density.

Analyzing the probabilities obtained in this way, we can assess the influence of individual factors of the model: the power and direction of the information flow, available to online users and relevant to the tasks of rating formation, as well as the number of supporters or opponents of the politician. Next, using the found eigenfunctions and eigenvalues, we derive expressions for the evaluation of conditional mathematical expectations of a politician's rating, which can serve as a basis for building forecasts that are important for the formation of a strategy of representing a political leader in the online environment.

Keywords: approval rating, political leadership, informational impact, herd behavior, Markov process

Citation: *Computer Research and Modeling*, 2019, vol. 11, no. 5, pp. 979–997 (Russian).

The reported study was funded by RFBR and EISR according to the research project No. 19-011-31651.

Введение

В данной статье мы исследуем вопрос формирования рейтинга одобрения (approval rating) политического лидера под влиянием процессов, протекающих в цифровом публичном пространстве. Под рейтингом одобрения мы будем понимать индикатор, значение которого оценивается в зависимости от ответов респондентов на вопрос вида «Одобряете ли вы деятельность... на его/ее посту?». Данный вопрос проводится по каждому политику отдельно и является закрытым, с заранее обозначенными вариантами ответов («да», «нет», «затрудняюсь ответить»). Доля положительных ответов показывает уровень одобрения действий и программ данного политика.

Основой для исследований в рассматриваемом нами случае являются модель Алана Кирмана [Kirman, 1993], а также ее обобщения [Alfarano et al., 2008; Kononovicius, Gontis, 2014]. Модель Кирмана базируется на фактах из биологии и описывает процесс формирования коллективного мнения среди членов муравьиной колонии. Всевозможные ее обобщения широко применяются в социально-экономическом моделировании процессов, связанных с феноменом стадного поведения. Так, например, в статье [Alfarano et al., 2008] данная модель используется для описания механизма формирования цены акции на рынке под влиянием смены настроений трейдеров (изменения соотношения между числом оптимистов и пессимистов, ожидающих роста или снижения цены соответственно). В статьях [Kononovicius, Gontis, 2014; Kononovicius, Gontis, 2013] указываются возможные направления в обобщении модели Кирмана, учитывающие такие факторы, как наличие контролируемых агентов и более чем двух состояний, в которых могут находиться агенты. В последнее время появляется большое число исследований в области формирования мнений и диффузии информации, опирающихся на рассматриваемый класс моделей. Среди исследований в области формирования мнений электората можно выделить работы [Jedrzejewski, Sznajd-Weron, 2018; Peralta et al., 2018], в которых рассматриваются нелинейные версии модели Кирмана в приложениях к формированию мнения граждан.

Основываясь на методах теории случайных процессов [Van Kampen, 1992; Гихман, Скороход, 1965; Karlin, Taylor, 1981], мы исследуем основные аспекты модели формирования рейтинга политического лидера, учитывающей внешнее информационное воздействие. В частности, исследуется влияние параметров модели на классификацию границ $\{0, 1\}$ интервала возможных значений рейтинга одобрения, а также приводятся выражения для определения переходных вероятностей между различными значениями рейтинговых показателей и выражения для расчета условных математических ожиданий. Также в статье представлены условия достижения состояния консенсуса в данной модели, то есть значений рейтинга одобрения, равных 0 или 1.

Мы будем предполагать, что в цифровом пространстве имеется набор агентов, у которых под влиянием ряда факторов формируется определенное мнение в отношении рассматриваемого политического лидера. Мнения агентов мы будем классифицировать в зависимости от того, являются ли данные агенты сторонниками данного политика и одобряют его деятельность или же, напротив, не готовы оказывать ему поддержку. Динамика рейтинга одобрения, то есть доли агентов, являющихся сторонниками политика и его программы, будет предметом исследования данной статьи. Важно подчеркнуть, что классификация конкретного агента, как сторонника рассматриваемого политического лидера, не является перманентной и может меняться с течением времени. Среди факторов, которые могут повлиять на то, что конкретный агент одобряет деятельность политика, мы будем выделять следующие: *межличностные взаимодействия* агентов, под влиянием которых они могут изменять свою точку зрения в отношении политика и становиться его сторонниками либо же выходить из их числа, а также *наличие выделенной категории агентов*, которые не меняют своего мнения в отношении данного политика, однако способны оказать влияние на других агентов (так называемые лидеры мнений). Представленная в данной статье модель предусматривает наличие дополнительного фактора, способного оказывать на численность популяции сторонников, — *новостного фона*, имеющего отношение к рассматриваемому

политику и доступного в публичном цифровом пространстве. Отметим, что существуют многочисленные публикации, например [Bollen et al., 2011; Barberá, Rivero, 2015], связывающие результаты анализа сообщений в социальных сетях методами NLP (Natural language processing, обработка естественного языка) и результаты социологических опросов по политической и социально-экономической тематике. Действительно, новостные сообщения, появляющиеся в онлайн-пространстве, представленном многочисленными информационными ресурсами и площадками, блогами и социальными сетями, являются одним из главных факторов, благодаря которому представители электората могут получить сведения о деятельности политика и сформировать свое отношение к нему. А следовательно, управление данным фактором и контроль за ним становятся неотъемлемыми условиями формирования успешной стратегии политического лидера. В работе [Gayo-Avello, 2012] делается обзор публикаций по предсказанию результатов опросов на основе таких характеристик новостей в социальной сети Twitter, как доля сообщений на заданную тему и их тональность. Статьи [O'Connor, 2010; Cummings et al., 2010] подтверждают наличие существенной корреляции (по данным работы [Cummings et al., 2010], корреляция может достигать до 0.696) между опросами по уровню одобрения политических лидеров и указанными показателями сообщений из социальных сетей. Отметим, что приведенные публикации касаются прежде всего микроблогингового сервиса Twitter ввиду его популярности в США и разработанных открытых средств получения информации из него. В данной статье мы не ставим перед собой целью анализ сообщений из конкретной социальной сети или новостного ресурса, поэтому используем широкий термин «публичное цифровое пространство». Включение конкретных информационных ресурсов для анализа сообщений из них следует осуществлять в зависимости от их популярности в том или ином регионе мира, например, на основании статистики, предоставляемой ресурсами [Alexa Internet] или [SimilarWeb].

Описание модели и постановка задачи

Пусть N — общее число агентов, обладающих доступом к цифровому публичному пространству и способных получать из него информацию, а также общаться между собою, формируя при этом собственное мнение. Будем считать, что каждый из агентов может находиться в одном из двух возможных состояний: $n_i = 0$ (не является сторонником рассматриваемого политика, не одобряет его деятельность или затрудняется с ответом) или $n_i = 1$ (является сторонником, одобряет деятельность), $i = 1, \dots, N$. В данной статье мы будем рассматривать случай, когда каждый агент i в состоянии общаться с оставшимися $N - 1$ агентами, что является допустимым предположением, в ситуации публичного цифрового пространства, где отдельный пользователь может свободно вступить в открытый диалог с остальными представителями сообщества. Также введем обозначение $n = \sum_{i=1}^N n_i$ для общего числа сторонников политического лидера в рамках данной аудитории.

Динамика численности сторонников политического лидера моделируется при помощи задания для случайной величины n соответствующей цепи Маркова с непрерывным временем и описанием интенсивностей перехода между возможными состояниями $\{0, 1, \dots, N\}$.

Для интенсивностей переходов из состояния n в состояния $n - 1$ и $n + 1$ введем обозначения $\pi^-(n, t)$ и $\pi^+(n, t)$ соответственно. При этом отметим наличие зависимости данных интенсивностей от момента времени t . С учетом данных обозначений вероятности перехода между состояниями рассматриваемой цепи формулируются следующим образом:

$$\begin{aligned} P(n \pm 1, t + \Delta t | n, t) &= \pi^\pm(n, t)\Delta t + o(\Delta t), \\ P(n \pm m, t + \Delta t | n, t) &= o(\Delta t), \quad m > 1, \\ P(n, t + \Delta t | n, t) &= 1 - (\pi^-(n, t) + \pi^+(n, t))\Delta t + o(\Delta t). \end{aligned} \tag{1}$$

Таким образом, за малый промежуток времени Δt число агентов, одобряющих деятельность данного политика, может измениться на ± 1 или остаться неизменным.

Перейдем к описанию интенсивностей $\pi^\pm(n, t)$. В оригинальной модели Кирмана [Kirman, 1993], а также, в последствии, в модели Альфарано [Alfarano et al., 2008] они заданы как

$$\begin{aligned}\pi^+(n, t) &= (N - n)(a + bn), \\ \pi^-(n, t) &= n(a + b(N - n)).\end{aligned}$$

Смысл подобной функциональной зависимости состоит в следующем. Параметр $b > 0$ отвечает за изменение числа сторонников вследствие их межличностного взаимодействия и общения. Параметр $a > 0$, напротив, отвечает за смену мнений под влиянием внешних факторов, не связанных взаимодействием агентов между собой, например из-за деятельности СМИ по освещению событий, прямо или косвенно связанных с деятельностью данного политика. Отметим, что в данной форме рассматриваемая модель является идеологически близкой моделям [Михайлов, Маревцева, 2011; Михайлов и др., 2014], в которых рассматривается проблема динамики численности индивидов, владеющих информацией и распространяющих ее. Вместо аппарата теории марковских процессов в ней применяются методы теории дифференциальных уравнений. Аналог параметра a в данных работах связан со «скоростью внешней вербовки», а аналог параметра b влияет на «скорость внутренней вербовки». В представленной статье мы сформулируем один из возможных вариантов развития механизма воздействия информации на рост или снижение числа сторонников политического лидера, высказывающих одобрение проводимой им деятельности.

Следующий шаг, который возможно (см. [Kononovicius, Gontis, 2014]) сделать на пути построения модели распространения популярности политического лидера, состоит во введении в рассмотрение обособленной группы численностью M , представляющей множество агентов, не меняющих свое отношение к политику под влиянием каких бы то ни было факторов. При этом будем считать, что в данной группе находятся M_+ безусловных сторонников политического лидера и M_- безусловных противников. Как будет показано в ходе анализа данной модели, конфигурация данной группы агентов оказывает существенное значение на уровень поддержки политика, в том числе в предельном случае, при $N \rightarrow +\infty$. Поэтому при применении излагаемой модели на практике в данной группе необходимо учитывать лидеров мнений, имеющих большую аудиторию подписчиков, а также способных воздействовать на имидж политического лидера. С учетом изложенного интенсивности смены мнений будут моделироваться как

$$\begin{aligned}\pi^+(n, t) &= (N - n)(a + bn + cM_+), \\ \pi^-(n, t) &= n(a + b(N - n) + cM_-),\end{aligned}$$

где параметр $c > 0$ отвечает за степень влияния группы сторонников или противников политика на агентов, рассматриваемых в данной модели.

Завершающим элементом в построении нашей модели является описание механизма информационного воздействия, учитывающего возможную смену интенсивности потока сообщений и их тональности в цифровом публичном пространстве. Будем предполагать, что за интервал времени $(t - \Delta t, t]$ появляются $h_t \geq 0$ новых информационных сообщений, имеющих отношение к рассматриваемому политику и размещенных на крупных информационных порталах или крупных группах в социальных сетях. Иными словами, под информационными и новостными сообщениями мы будем подразумевать материалы, размещенные на ресурсах указанных типов. Длительность Δt выбирается в зависимости от особенностей цифрового ресурса (новостного сайта, блога или социальной сети) и промежутка времени, в течение которого актуальные информационные сообщения могут быть сравнительно просто получены пользователями. Каждое из сообщений $\{m_{1,t}, \dots, m_{h_t,t}\}$ имеет тональность (см. [De et al., 2014]) $e_{i,t} \in [-1, 1]$, $i = 1, \dots, h_t$.

Для ее определения можно использовать один из многочисленных методов оценки тональности текстов (sentiment analysis) [Mälntyläl et al., 2018; Nakov et al., 2018]. Тональность -1 соответствует абсолютно негативному (по эмоциональной окраске) новостному сообщению, и, напротив, оценка $+1$ соответствует абсолютно позитивному сообщению. Если тональность сообщения не удастся определить, считается, что оно имеет нейтральную (нулевую) тональность. Введем обозначение для среднего значения тональностей сообщений:

$$s_t = \frac{\sum_{i=1}^{h_t} e_{i,t}}{h_t}, \quad (2)$$

и обозначение для показателя мощности информационного воздействия:

$$q_t = \tanh(\delta h_t). \quad (3)$$

Использование монотонно возрастающих, ограниченных функций, таких как \tanh или сигмоида, является распространенной практикой (например, [Chen et al., 2016]) при работе с неограниченными величинами. Величина $\delta \geq 0$ играет роль параметра чувствительности и отвечает за то, какое число сообщений требуется для роста мощности воздействия от 0 до 1. При $0 < \delta < +\infty$ минимальная мощность ($q_t = 0$) соответствует отсутствию каких-либо сообщений, относящихся к данному политику, а максимальная мощность ($q_t = 1$) соответствует ситуации бесконечно большого потока сообщений. При $\delta = 0$ величина q_t будет равняться 0, вне зависимости от числа сообщений, таким образом делая рассматриваемую нами систему нечувствительной к внешней информации. В ситуации, когда $\delta = +\infty$, наличие хотя бы одного сообщения будет соответствовать максимальной мощности. Для того чтобы учесть, что даже максимальное информационное воздействие может иметь ограниченный эффект при формировании мнения агентов в отношении политического лидера, введем параметр $\bar{a} \geq 0$, как и прежде отражающий степень влияния информации на число сторонников политика, одобряющих его деятельность, и величину

$$a_t = \bar{a} q_t.$$

Далее под мощностью информационного воздействия будем понимать именно величину a_t . Также нами будет сделано допущение о том, что тональность s_t и мощность q_t информационного воздействия являются функциями времени t . Данное допущение, безусловно, накладывает ограничения на рассматриваемую нами модель, делая ее менее «гибкой» и реалистичной, однако позволяет понять принципиальные зависимости между объемом и характером информации и динамикой рейтинговых показателей.

Теперь применим наше знание о средней тональности информационных сообщений, для этого зададим интенсивности переходов в рассматриваемой модели как

$$\begin{aligned} \pi^+(n, t) &= (N - n) \left(a_t \frac{1 + s_t}{2} + bn + cM_+ \right) = (N - n) \left(\bar{a} \cdot \tanh(\delta h_t) \cdot \frac{1 + s_t}{2} + bn + cM_+ \right), \\ \pi^-(n, t) &= n \left(a_t \frac{1 - s_t}{2} + b(N - n) + cM_- \right) = n \left(\bar{a} \cdot \tanh(\delta h_t) \cdot \frac{1 - s_t}{2} + b(N - n) + cM_- \right). \end{aligned} \quad (4)$$

Подобная функциональная форма задания интенсивностей соответствует тому, что мощность информационного воздействия a_t раскладывается на две составляющие $a_t \frac{1 + s_t}{2}$ и $a_t \frac{1 - s_t}{2}$, значения которых определяется в зависимости от показателя тональности. Первая из частей положительно влияет на интенсивность роста числа сторонников и увеличивает рейтинг одобрения, а вторая, соответственно, повышает интенсивность их оттока и негативно сказывается на динамике рейтинга.

Мощность информационного воздействия, влияющего на рост числа сторонников политика, будет достигать максимального значения при абсолютно положительной тональности новостей ($s_t = 1$) и снижаться до нулевого значения при абсолютно отрицательных сообщениях ($s_t = -1$). Противоположным образом сказанное относится к тем, кто не причисляет себя к сторонникам политика. Нейтральная по тональности информация $s_t = 0$ будет с одинаковой мощностью $a_t/2$ действовать как на интенсивность роста, так и на интенсивность снижения числа сторонников, поскольку с одинаковой вероятностью может быть интерпретирована как в положительном, так и в отрицательном ключе. Также отметим, что даже отрицательная средняя тональность новостей $-1 < s_t < 0$ имеет ненулевое влияние (при $a > 0$) на интенсивность роста числа сторонников, поскольку, как и в предыдущем случае, возможны различные варианты интерпретации новостных сообщений, в том числе не совпадающие с результатами алгоритмов определения тональности.

Для дальнейшего изложения модели введем процесс

$$X_t = \frac{n_t}{N}, \quad t \geq 0,$$

равный доле агентов на момент времени t , одобряющих деятельность политика. Именно данный процесс (рейтинг одобрения) будет предметом нашего дальнейшего рассмотрения.

Пространством состояний процесса X является отрезок множество $\left\{0, \frac{1}{N}, \frac{2}{N}, \dots, 1\right\}$. Для описания динамики процесса процесса X введем следующие вспомогательные величины:

$$\begin{aligned} \pi^+(x, t) &= (1 - x) \left(\frac{a_t \frac{(1 + s_t)}{2} + cM_+}{N} + bx \right), \\ \pi^-(x, t) &= x \left(\frac{a_t \frac{(1 - s_t)}{2} + cM_-}{N} + b(1 - x) \right). \end{aligned} \tag{5}$$

Связь между ранее заданными интенсивностями (4) и только что определенными величинами (5) задается соотношениями

$$\pi^\pm(n, t) = N^2 \pi^\pm(x, t).$$

С использованием (5) переходные вероятности в терминах процесса X можно записать как

$$P(x \pm \Delta x, t + \Delta t \mid x, t) = N^2 \pi^\pm(x, t) \Delta t + o(\Delta t),$$

где $\Delta x = \frac{1}{N}$.

Пусть на начальный момент времени t_0 имеется $n_0 = \lceil x_0 \cdot N \rceil$ сторонников политика. Вероятность $P(n, t) = P(n_t = n \mid n_{t_0} = n_0)$ наличия n сторонников в момент времени $t > t_0$ удовлетворяет уравнению Колмогорова [Гихман, Скороход, 1965]:

$$\frac{\partial P}{\partial t}(n, t) = \pi^+(n - 1, t)P(n - 1, t) + \pi^-(n + 1, t)P(n + 1, t) - (\pi^-(n, t) + \pi^+(n, t))P(n, t).$$

В терминах переменной $x \left(x = \frac{n}{N}\right)$, отвечающей процессу X , данное уравнение имеет вид

$$\frac{\partial P}{\partial t}(x, t) = N^2 (\pi^+(x - \Delta x, t)P(x - \Delta x, t) + \pi^-(x + \Delta x, t)P(x + \Delta x, t) - (\pi^-(x, t) + \pi^+(x, t))P(x, t)).$$

Используя операторы шага $Ef(x) = f(x + \Delta x)$, $E^{-1}f(x) = f(x - \Delta x)$ (см. [Van Kampen, 1992]), можно переписать данное уравнение как

$$\frac{\partial P}{\partial t}(x, t) = N^2((E - 1)(\pi^-(x, t)P(x, t)) + (E^{-1} - 1)(\pi^+(x, t)P(x, t))). \quad (6)$$

Перепишем уравнение (6) для плотности $p(x, t)$, определяемой как $p(x, t) = \frac{1}{\Delta x}P(x, t) = \frac{1}{\Delta x}P\left(x \leq X_t < x + \Delta x \mid X_{t_0} = \frac{n_0}{N}\right)$:

$$\frac{\partial p}{\partial t}(x, t) = N^2((E - 1)(\pi^-(x, t)p(x, t)) + (E^{-1} - 1)(\pi^+(x, t)p(x, t))).$$

Для дважды непрерывно дифференцируемой функции f действие оператора E можно описать приближенно:

$$\begin{aligned} Ef(x) &= f(x + \Delta x) = f(x) + \Delta x \frac{\partial f}{\partial x} + \frac{1}{2}(\Delta x)^2 \frac{\partial^2 f}{\partial x^2} + o(\Delta x^2), \quad \Delta x \rightarrow 0, \\ E^{-1}f(x) &= f(x - \Delta x) = f(x) - \Delta x \frac{\partial f}{\partial x} + \frac{1}{2}(\Delta x)^2 \frac{\partial^2 f}{\partial x^2} + o(\Delta x^2), \quad \Delta x \rightarrow 0. \end{aligned}$$

Подставим данное разложение в предыдущее уравнение:

$$\frac{\partial p}{\partial t}(x, t) = N^2\left(-\Delta x \frac{\partial}{\partial x}((\pi^+ - \pi^-)p(x, t)) + \frac{1}{2}(\Delta x)^2 \frac{\partial^2}{\partial x^2}((\pi^+ + \pi^-)p(x, t)) + o(\Delta x^2)\right)$$

и осуществим предельный переход $N \rightarrow +\infty$.

При этом

$$N(\pi^+ - \pi^-) = a_t \left(\frac{1 + s_t}{2} - x\right) + c(M_+(1 - x) - M_-x)$$

и

$$\pi^+ + \pi^- = 2bx(1 - x) + O\left(\frac{1}{N}\right) \rightarrow 2bx(1 - x).$$

Таким образом, имеет место следующее предложение о переходных вероятностях диффузионного приближения рассматриваемого процесса.

Предложение 1. *Плотность переходной вероятности $p(x, t \mid x_0, t_0)$ процесса X_t , $t > t_0$, удовлетворяет следующему уравнению Фоккера–Планка:*

$$\frac{\partial p}{\partial t}p(x, t \mid x_0, t_0) = -\frac{\partial}{\partial x}(\mu(x, t)p(x, t \mid x_0, t_0)) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(\sigma^2(x, t)p(x, t \mid x_0, t_0)), \quad (7)$$

$$\lim_{t \downarrow t_0} p(x, t \mid x_0, t_0) = \delta(t - t_0), \quad (8)$$

где функции сноса и диффузии определены как

$$\mu(x, t) = a_t \left(\frac{1 + s_t}{2} - x\right) + c(M_+(1 - x) - M_-x), \quad (9)$$

$$\sigma^2(x, t) = 2bx(1 - x). \quad (10)$$

Известно (см. [Gardiner et al., 1985]), что полученное уравнение Фоккера–Планка соответствует следующему стохастическому дифференциальному уравнению (СДУ):

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dW_t. \quad (11)$$

Действительно, возьмем тестовую функцию $f(x) \in C^2(R)$, такую, что $f(x) = 0$ при $x \in R \setminus [0, 1]$, $f(0) = 0 = f(1)$ и $\frac{\partial f}{\partial x}(0) = 0 = \frac{\partial f}{\partial x}(1)$. Вычислим выражение для $\frac{d}{dt}Ef(X_t | X_{t_0} = x_0)$, пользуясь формулой Ито [Øksendal, 2003] и уравнением (11):

$$\begin{aligned} \frac{d}{dt}Ef(X_t | X_{t_0} = x_0) &= \int_0^1 \left(\frac{\partial f}{\partial x} \mu(x, t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma^2(x, t) \right) p(x, t | x_0, t_0) dx = \\ &= \int_0^1 \left(-\frac{\partial}{\partial x} (\mu(x, t) p(x, t | x_0, t_0)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(x, t) p(x, t | x_0, t_0)) \right) f(x) dx. \end{aligned} \tag{12}$$

С другой стороны, имеет место следующее выражение:

$$\frac{d}{dt}Ef(X_t | X_{t_0} = x_0) = \frac{d}{dt} \int_0^1 p(x, t | x_0, t_0) f(x) dx = \int_0^1 \frac{\partial}{\partial t} p(x, t | x_0, t_0) f(x) dx. \tag{13}$$

Приравняв правые части (12) и (13), получаем в силу произвольного выбора тестовой функции f

$$\frac{\partial}{\partial t} p(x, t | x_0, t_0) = -\frac{\partial}{\partial x} (\mu(x, t) p(x, t | x_0, t_0)) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(x, t) p(x, t | x_0, t_0)).$$

Таким образом, плотность вероятности перехода случайного процесса, являющегося решением уравнения (11), удовлетворяет уравнению Фоккера – Планка (8).

Именно процесс X , являющийся решением уравнения (11), будет основным предметом нашего внимания. В рамках данной статьи мы представим выражения для вычисления переходных вероятностей данного процесса, а также исследуем задачу построения прогнозов для данного процесса для произвольных моментов будущего T , в том числе в долгосрочной перспективе, т. е. при $T \rightarrow +\infty$.

Анализ и обсуждение модели

В данном и следующем разделах для простоты мы будем полагать, что мощность и тональность информационного потока являются фиксированными, то есть $a_t \equiv a$ и $s_t \equiv s$ соответственно. В соответствии с данным допущением далее в функциях σ и μ мы будем пропускать зависимость от переменной t , поскольку мощность и тональность информационного воздействия фиксированы и не зависят от времени.

Следуя технике, изложенной в [Бородин, 2013], введем функцию

$$s(x) = e^{-\int_x^1 \frac{2\mu(z)}{\sigma^2(z)} dz} \propto \frac{1}{(1-x)^{\frac{a}{b} \frac{1-s}{2} + \frac{c}{b} M_-} x^{\frac{a}{b} \frac{1+s}{2} + \frac{c}{b} M_+}}, \tag{14}$$

называемую *плотностью меры шкалы*. Далее, $S(x) = \int_x^1 s(z) dz$ — *шкала диффузии*, и соответствующая ей мера $S[x, y] = S(x) - S(y) = \int_x^y s(z) dz$ — *мера шкалы*. Также введем функцию

$$m(x) = \frac{1}{\sigma^2(x)s(x)} \propto (1-x)^{\frac{a}{b} \frac{1-s}{2} + \frac{c}{b} M_- - 1} x^{\frac{a}{b} \frac{1+s}{2} + \frac{c}{b} M_+ - 1}, \tag{15}$$

называемую *плотностью меры скорости*, и меру, индуцированную ею, $M[x, y] = \int_x^y m(z) dz$, называемую *мерой скорости*. Отсутствие нижнего индекса интегрирования в данных определениях означает, что требуется вычислить произвольную первообразную подынтегральной функции.

Таким образом, введенные выше функции определены с точностью до некоторого постоянно-го слагаемого или множителя. Однако нас будут интересовать не конкретные значения данных функций, а возможность того, могут ли данные функции принимать значения, равные $\pm\infty$.

Поскольку рассматриваемый нами процесс принимает значения в ограниченном интервале $[l, r]$, $l = 0$, $r = 1$, необходимо определить характер его поведения в граничных точках 0 и 1 в зависимости от сочетания параметров модели. Для решения этой задачи дополнительно введем следующие функции и рассчитываемые на их основе величины, которые будут использованы при установлении характеристик процесса X :

$$\begin{aligned}\Sigma[x, y] &= \int_x^y (S(z) - S(x)) dM(z), & N[x, y] &= \int_x^y (S(y) - S(z)) dM(z), \\ \Sigma(l) &= \lim_{z \downarrow l} \Sigma(z, c), & N(l) &= \lim_{z \downarrow l} N(z, c), \\ \Sigma(r) &= \lim_{z \uparrow r} \Sigma(c, z), & N(r) &= \lim_{z \uparrow r} N(c, z),\end{aligned}$$

где c — произвольная точка из интервала (l, r) , выбор которой не принципиален для определения того, конечны или же бесконечны величины $\Sigma(l)$, $\Sigma(r)$, $N(l)$, $N(r)$. Указанные величины имеют следующую эвристическую интерпретацию [Karlin, Taylor, 1981]: $\Sigma(l)$ измеряет время, которое требуется для достижения l диффузией, начинающейся во внутренней точке $c \in (l, r)$; $N(l)$ измеряет время, которое требуется для достижения точки $c \in (l, r)$ диффузией, начинающейся с границы l . Аналогичную интерпретацию имеют величины $\Sigma(r)$, $N(r)$.

Пусть $H_z = \inf\{s: X_s = z\}$ — момент первого достижения процессом X уровня z . Если уровень z не достигается, полагаем $H_z = \infty$. Граница l называется *границей входа*, если она не может быть достигнута из любой точки интервала (l, r) . Но при этом можно рассматривать диффузии, начинающиеся на этой границе. Такие диффузии перемещаются внутрь интервала и никогда не возвращаются к входной границе. Необходимым и достаточным условием для того, чтобы l была границей входа, является $\Sigma(l) = \infty$ и $N(l) < \infty$. Если же граница может как заходить на границу, так и выходить из нее, то такая граница называется *регулярной*. Граница l является регулярной тогда и только тогда, когда $\Sigma(l) < \infty$ и $N(l) < \infty$. Еще одним пунктом классификации границ являются *границы выхода*: l является границей выхода, если

$$\lim_{z \downarrow l} \lim_{x \downarrow l} P_x(H_z < t) = 0.$$

Т.е. для диффузии, начинающейся в точке l (или в точке x , приближающейся к l), невозможно достигнуть никакого другого внутреннего состояния z , независимо от того, как близко z к l . Граница является выходом тогда и только тогда, когда $\Sigma(l) < \infty$ и $N(l) = \infty$. Далее, граница, которая не может быть достигнута за конечное время и из которой диффузия не может начинаться, называется *естественной* (по Феллеру). Граница l является естественной тогда и только тогда, когда $\Sigma(l) = \infty$ и $N(l) = \infty$. Все приведенные характеристики аналогично переносятся и на правую границу r . Пользуясь выражениями (14), (15) и повторяя рассуждения, сделанные в [Karlin, Taylor, 1981] для диффузионного процесса в модели Райта–Фишера, приходим к следующему предложению о классификации границ интервала возможных значений рейтинга одобрения.

Предложение 2. Для рассматриваемого процесса X , в зависимости от значения величин

$$\gamma_l = \frac{a}{b} \frac{1+s}{2} + \frac{c}{b} M_+, \quad (16)$$

$$\gamma_r = \frac{a}{b} \frac{1-s}{2} + \frac{c}{b} M_-, \quad (17)$$

имеет место следующая классификация границ: $l = 0$ и $r = 1$ (таблица 1).

Таблица 1. Классификация по Феллеру границы $i = l, r$ в зависимости от значений параметров модели

Случай	$\Sigma(i)$	$N(i)$	Тип границы i
$\gamma_i = 0$	$< \infty$	$= \infty$	Граница выхода
$\gamma_i \in (0, 1)$	$< \infty$	$< \infty$	Регулярная граница
$\gamma_i \geq 1$	$= \infty$	$< \infty$	Граница входа

Дадим некоторые комментарии по представленной классификации в приложении к исследуемому вопросу. Ситуация, при которой левая граница ($l = 0$) интервала возможных значений рейтинга будет границей выхода, означает, что реалистичным (и достижимым) вариантом развития событий является полная утрата поддержки у сторонников и отсутствие возможности продолжать политическую карьеру. Данная ситуация возможна при абсолютно негативном ($s = -1$) потоке новостей (либо полном их отсутствии ($a = 0$)) вместе с отсутствием «группы поддержки» из числа лидеров мнений ($M_+ = 0$). При наличии хотя бы одного лидера мнений ($M_+ \geq 1$), безусловно поддерживающего данного политика, граница $l = 0$ уже не будет границей выхода. А при наличии $M_+ > \frac{b}{c}$ лиц из числа лидеров мнений, безусловно поддерживающих данного политика, граница $l = 0$ станет границей входа и перестанет быть достижимой из внутренних состояний интервала (l, r) , вне зависимости от тональности информационных сообщений. Ситуация, когда правая граница $r = 1$ будет границей выхода, соответствует тому, что отсутствуют безусловные политические противники ($M_- = 0$), способные оказать влияние на рейтинг одобрения данного политика, а информационные сообщения либо имеют абсолютно положительную тональность ($s = 1$), либо же отсутствуют вообще ($a = 0$).

Безусловно, приведенные выше варианты, при которых границы интервала возможных значений рейтинга являются границами выхода, являются исключительно модельными, поскольку в цифровом публичном пространстве имеет место плюрализм мнений, приводящий к тому, что тональность сообщений никогда не достигает предельных значений ± 1 . Поэтому далее мы будем считать, что $-1 < s < 1$, и предполагать, что границы интервала возможных значений рейтинга являются либо регулярными границами с отражением, либо границами входа.

Обратимся теперь к вопросу отыскания плотности $\pi(x)$ стационарного распределения процесса X , такой, что выполняется равенство $\pi(y) = \int_0^1 \pi(x)p(y, t | x, 0) dx$ для любого $t > 0$. Известно (см. [Karlin, Taylor, 1981]), что выражение для нее может быть представлено через определенные ранее меру шкалы S и плотность меры скорости m :

$$\pi(x) = m(x) (C_1 S(x) + C_2), \tag{18}$$

где константы C_1 и C_2 выбираются таким образом, чтобы $0 \leq \pi(x) < +\infty, x \in [0, 1]$ и $\int_0^1 \pi(x) dx = 1$.

Если $\gamma_r = \frac{a(1-s)}{b} + \frac{c}{b}M_- \geq 1$ или $\gamma_l = \frac{a(1+s)}{b} + \frac{c}{b}M_+ \geq 1$, то интеграл в определении $S(x)$ будет расходящимся, поэтому $C_1 = 0$ и, таким образом, $\pi(x) = C_2 m(x)$:

$$\pi(x) = \frac{\Gamma(\gamma_r + \gamma_l)}{\Gamma(\gamma_r)\Gamma(\gamma_l)} (1-x)^{\gamma_r-1} x^{\gamma_l-1}. \tag{19}$$

Таким образом, стационарным распределением будет бета-распределение $Be(\gamma_l, \gamma_r)$. Если же и $\gamma_l < 1$ и $\gamma_r < 1$, то формула (18) определяет параметрическое семейство распределений. С учетом сделанных замечаний, для того, чтобы плотность стационарного распределения процесса X

определялась в соответствии с формулой (19), достаточно выполнение условия

$$b \leq \frac{1}{2}(a + cM).$$

Рассмотрим инфинитезимальный оператор A случайного процесса X :

$$A = \frac{1}{2}\sigma^2(x)\frac{d}{dx^2} + \mu(x)\frac{d}{dx} = bx(1-x)\frac{d^2}{dx^2} + \left(a\frac{1+s}{2} + cM_+ - x(a+cM)\right)\frac{d}{dx} = \frac{1}{2}\frac{d}{dM}\left(\frac{df(x)}{dS}\right), \quad (20)$$

заданный в $L_2([0, 1], m)$ — пространстве функций, квадраты которых на отрезке $[0, 1]$ суммируемы с весом $m(x)$. В данном пространстве скалярное произведение $\langle \cdot, \cdot \rangle_m$ определяется как

$$\langle u, v \rangle_m = \int_0^1 u(x)v(x)m(x)dx.$$

Областью определения оператора A является (см. [McKean, 1956])

$$D(A) = \{f \in L_2([0, 1], m) : Af \in C_b((0, 1)), \text{ граничные условия в точках } 0 \text{ и } 1\},$$

где граничные условия выбираются в зависимости от типа границы: если точка $l = 0$ является границей выхода или же регулярной границей с поглощением, то условие задается как $\lim_{x \downarrow 0} f(x) = 0$; если же $l = 0$ — это граница входа или регулярная граница с отражением, то усло-

вие формулируется как $\lim_{x \downarrow 0} \frac{f'(x)}{s(x)} = 0$. Аналогично задаются условия для правой границы $r = 1$.

В работах [Elliott, 1955; McKean, 1956] доказано, что в ситуации отсутствия естественных границ у процесса X оператор A имеет дискретный спектр, состоящий из собственных значений $\{\lambda_n\}_{n=0}^{\infty}$, $0 \geq \lambda_0 > \dots > \lambda_n > \dots$, $\lim_{n \rightarrow \infty} \lambda_n = -\infty$. Пусть ψ_n — собственные функции, отвечающие данным собственным значениям, такие, что $\|\psi_n\|_m^2 = 1$ и

$$A\psi_n = \lambda_n\psi_n, \quad n \geq 0. \quad (21)$$

Указанные функции будут ортогональными:

$$\langle \psi_i, \psi_j \rangle_m = \delta_{ij}, \quad (22)$$

где δ_{ij} — символ Кронекера. Ортогональность функций легко проверяется за счет использования канонического представления оператора A :

$$\begin{aligned} \lambda_i \langle \psi_i, \psi_j \rangle_m &= \lambda_i \int_0^1 \psi_i(x)\psi_j(x)m(x)dx = \int_0^1 \frac{1}{2} \frac{1}{m(x)} \frac{d}{dx} \left(\frac{1}{s(x)} \frac{d\psi_i(x)}{dx} \right) \psi_j(x)m(x)dx = \\ &= \int_0^1 \psi_i(x) \frac{1}{2} \frac{1}{m(x)} \frac{d}{dx} \left(\frac{1}{s(x)} \frac{d\psi_j(x)}{dx} \right) m(x)dx = \lambda_j \int_0^1 \psi_i(x)\psi_j(x)m(x)dx = \lambda_j \langle \psi_i, \psi_j \rangle_m, \end{aligned}$$

внеинтегральные члены при этом обращаются в 0 в силу заданных граничных условий. Таким образом, при $i \neq j$ функции ψ_i и ψ_j ортогональны друг другу.

Выпишем уравнение для нахождения собственных функций и собственных значений:

$$bx(1-x)\psi_n'' + \left(a \cdot \frac{1+s}{2} + cM_+ - x(a+cM)\right)\psi_n' = \lambda_n\psi_n, \quad (23)$$

а также соответствующие граничные условия:

$$\lim_{x \downarrow 0} \psi'_n(x)x^{\gamma_l} = 0, \tag{24}$$

$$\lim_{x \uparrow 1} \psi'_n(x)(1-x)^{\gamma_r} = 0. \tag{25}$$

Проделив замену $y = 2x - 1$, $\varphi_n(y) = \psi_n(x)$, преобразуем (23) к виду

$$(1-y^2)\varphi'_n + \left(\frac{a}{b}s + \frac{c}{b}(M_+ - M_-) - y\left(\frac{a}{b} + \frac{c}{b}M\right)\right)\varphi'_n - \frac{\lambda_n}{b}\varphi_n = 0. \tag{26}$$

Решением (26) является (см. [Abramowitz, Stegun, 1966]) следующая функция:

$$\begin{aligned} \varphi_n(y) &= P_n^{(\gamma_r-1, \gamma_l-1)}(y), \quad n \geq 0, \\ \lambda_n &= -(bn^2 + (a + cM)n - 1), \quad n \geq 0, \end{aligned} \tag{27}$$

где $P_n^{(\alpha, \beta)}(x)$ — многочлен Якоби:

$$P_n^{(\alpha, \beta)}(x) = \frac{\Gamma(\alpha + n + 1)}{n! \Gamma(\alpha + \beta + n + 1)} \sum_{m=0}^n \binom{n}{m} \frac{\Gamma(\alpha + \beta + n + m + 1)}{\Gamma(\alpha + m + 1)} \left(\frac{x-1}{2}\right)^m,$$

а величины γ_l, γ_r определены как (16) и (17) соответственно. Возвращаясь к исходным переменным, получим $\psi_n(x) = P_n^{\gamma_r-1, \gamma_l-1}(2x - 1)$. Поскольку ψ_n являются непрерывными на отрезке $[0, 1]$ функциями и, по предположению, $\gamma_l > 0$ и $\gamma_r > 0$, то граничные условия (24), (25) очевидным образом выполняются. Для функций $\psi_n(x)$, заданных таким образом, выполнено (см. [Abramowitz, Stegun, 1966]) условие $\int_0^1 (\psi_n(x))^2 m(x) dx = h_n$, где

$$h_n = \frac{1}{2n + \gamma_l + \gamma_r - 1} \frac{\Gamma(n + \gamma_l)\Gamma(n + \gamma_r)}{n! \Gamma(n + \gamma_l + \gamma_r - 1)}.$$

Поэтому для выполнения тождества $\|\psi_n\|^2 = 1$ переопределим ψ_n как

$$\psi_n(x) = \frac{1}{\sqrt{h_n}} P_n^{(\gamma_r-1, \gamma_l-1)}(2x - 1), \quad n \geq 0. \tag{28}$$

Для вывода уравнения для плотности $p(y, t | x, t_0)$ воспользуемся тем, что функция

$$u(x, t) = E(g(X(t)) | X_{t_0} = x) = \int_0^1 g(y)p(y, t | x, t_0)dy, \quad t > t_0, \tag{29}$$

является (см. [Гихман, Скороход, 1965]) решением уравнения

$$\begin{cases} \frac{\partial u}{\partial t} = Au, \\ u(x, t_0) = g(x), \end{cases} \tag{30}$$

где $g \in L_2([0, 1], m)$ — произвольная непрерывная функция. Разложим в ряд Фурье по семейству собственных функций оператора A данные задачи (30) и ее решение будем искать в виде ряда

$$u(x, t) = \sum_{n=0}^{\infty} u_n(t)\psi_n(x).$$

Подставляя указанное представление функции $u(x, t)$ в (30), получим, что функции $u_n(t)$, $n \geq 0$, являются решениями уравнений

$$\frac{du_n}{dt} = \lambda_n u_n, \quad u_n(t_0) = d_n, \quad n \geq 0,$$

где $d_n = \langle g, \psi_n \rangle_m$. Решение данного уравнения можно записать как

$$u_n(t) = d_n e^{\lambda_n(t-t_0)}, \quad n \geq 0, \quad t \geq t_0.$$

Таким образом, решение уравнения (30) для рассматриваемого нами случая можно записать как

$$u(x, t) = \sum_{n=0}^{\infty} e^{\lambda_n(t-t_0)} \psi_n(x) \langle g, \psi_n \rangle_m = \int_0^1 g(y) m(y) \sum_{n=0}^{\infty} \psi_n(x) \psi_n(y) e^{\lambda_n(t-t_0)} dy. \quad (31)$$

Сравнивая правые части (29) и (31), в силу произвольного выбора функции g приходим к следующему утверждению

Предложение 3. При условии, что мощность и тональность информационного воздействия являются фиксированными, т. е. $a_t \equiv a$ и $s_t \equiv s$, плотность переходных вероятностей

$$p(y, t | x, t_0) = \frac{\partial}{\partial y} P(X_t < y | X_{t_0} = x)$$

случайного процесса X , заданного как решение стохастического дифференциального уравнения (11), равна

$$p(y, t | x, t_0) = m(y) \sum_{n=0}^{\infty} \psi_n(x) \psi_n(y) e^{\lambda_n(t-t_0)}. \quad (32)$$

Пользуясь данным выражением для плотности, легко убедиться в том, что

$$\lim_{t \rightarrow \infty} p(y, t | x, t_0) = \pi(y),$$

причем скорость сходимости к плотности стационарного распределения тем выше, чем больше значение величин b и $a + cM$, входящих в выражение (27) для собственных значений λ_n .

Обратимся теперь к вопросу вычисления условных моментов для процесса X_t , $t \geq 0$. А именно, нас будет интересовать задача вычисления величин вида

$$E[X_T^k | X_0 = x_0], \quad k = 1, 2, \dots, \quad (33)$$

— условных моментов процесса X в некоторый момент будущего $T \geq 0$, при условии того, что текущая доля сторонников политика известна и равна $x_0 \in [0, 1]$. Для практических приложений наиболее интересны случаи $k = 1$ и $k = 2$, необходимые для расчета среднего значения и дисперсии величины X_T .

Воспользовавшись решением (31) задачи (30) для случая $g(x) = x^k$, а также тем, что $\langle x^k, \psi_n(x) \rangle_m = 0$ при $n > k$, получим следующий результат.

Предложение 4. *k*-й условный момент случайной величины X_t , $t > 0$, может быть вычислен как

$$E[X_t^k | X_0 = x_0] = \sum_{n=0}^k e^{\lambda n t} \psi_n(x_0) \langle x^k, \psi_n(x) \rangle_m. \quad (34)$$

Пользуясь данным результатом, получим, что условное математическое ожидание для процесса X равняется

$$r(x, t) = E[X_t | X_0 = x_0] = r_\infty + e^{-(a+cM)T} (x_0 - r_\infty), \quad (35)$$

где

$$r_\infty = \frac{a\left(\frac{1+s}{2}\right) + cM_+}{a + cM}.$$

При $t \rightarrow +\infty$ имеем $r(x, t) \rightarrow r_\infty$. Таким образом, долгосрочный прогноз рейтинга политического лидера в рамках рассматриваемой модели допускает эвристическую интерпретацию как долю сил $a\left(\frac{1+s}{2}\right) + cM_+$, направленных на рост числа агентов, одобряющих его деятельность, в общем объеме усилий $a + cM$ (см. (5)). Примечательно, что параметр b , отвечающий за межличностное взаимодействие представителей электората, в оценке ожидаемого значения не фигурирует. При $M = 0$ долгосрочный прогноз рейтинга будет равен $\frac{1+s}{2}$. Таким образом, при отсутствии лидеров мнений, являющихся безусловными сторонниками или противниками данного политика, его рейтинг будет однозначно определяться (при $a > 0$) тональностью новостного фона. И соответственно, при отсутствии информационного воздействия, т. е. $a = 0$, рейтинг одобрения будет определяться долей сторонников политика из числа лидеров мнений $\frac{M_+}{M}$.

Рассмотрим некоторые численные примеры, демонстрирующие возможности расчета переходных вероятностей. На графиках рис. 1, 2, 3 отобразим динамику плотности $p(y, t | x, 0)$ переходной вероятности при различных значениях t . Параметры модели для обоих графиков выбраны следующими: $a = 0.30$, $b = 0.20$, $c = 0.10$, $M_+ = 2$, $M_- = 5$. На всех трех графиках выбрано начальное состояние $x = 0.05$. Однако в первом случае (график рис. 1) параметр, моделирующий тональность новостей, имеющих отношение к рассматриваемому политику, выбран равным $s = 0$, на графике рис. 2 $s = 0.90$ и на графике рис. 3 изображена динамика плотности переходной вероятности для случая $s = -0.90$. Над каждым графиком изображено среднее значение величины рейтинга одобрения политика, а также ее стандартное отклонение. Таким образом, становится видным влияние мощности и тональности информационного воздействия на долю сторонников политического лидера: разница между математическим ожиданием рейтинга на момент времени $t = 0.3$ в случае информационного фона с тональностью $s = -0.9$ и аналогичным вариантом, но с тональностью $s = 0.9$, составляет 23.9 процентных пункта. Также видно, как с течением времени t плотность переходной вероятности стремится к своему стационарному значению.

Заключение

В представленной работе мы рассмотрели модель формирования рейтинга одобрения политического лидера под влиянием процессов, протекающих в цифровом публичном пространстве, таких как взаимодействие и обмен мнениями между агентами, наличие лидеров мнений, являющихся сторонниками или противниками данного политика, и появление информационных и новостных сообщений с определенной эмоциональной тональностью. Представленная модель дает возможную интерпретацию такого факта, как наличие положительной взаимосвязи (см. [O'Connor, 2010; Cummings et al., 2010]) между числом и тональностью сообщений, относящихся к политическому лидеру, и результатами социологических опросов по выявлению уровня

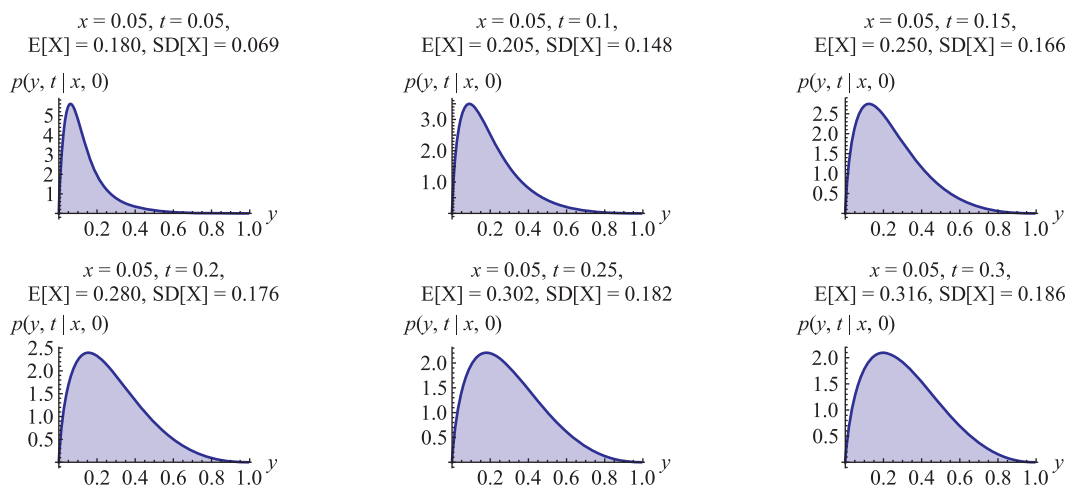


Рис. 1. Плотность переходной вероятности $p(y, t | x, 0)$ при $x = 0.01$ с течением времени t для параметров $a = 0.30, b = 0.20, c = 0.10, M_+ = 2, M_- = 5$. Значение параметра s выбрано равным 0

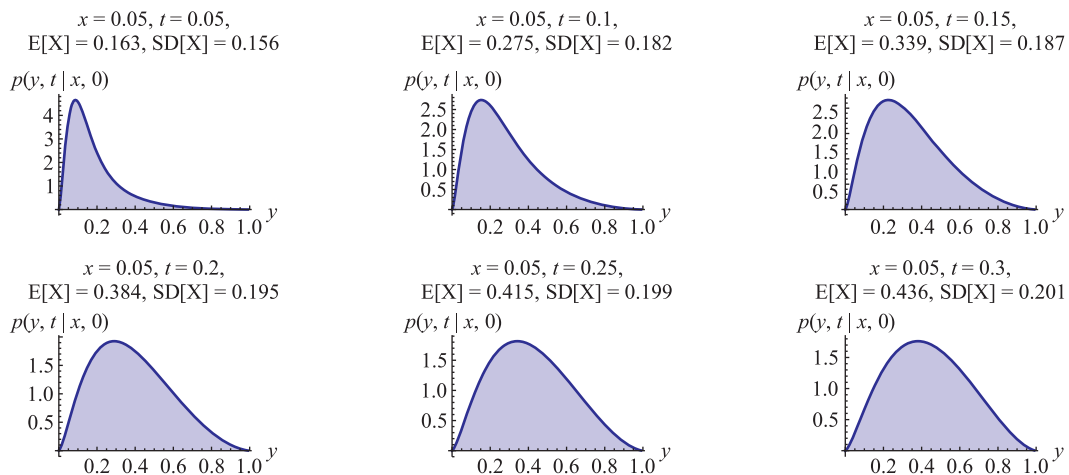


Рис. 2. Плотность переходной вероятности $p(y, t | x, 0)$ при $x = 0.01$ с течением времени t для параметров $a = 0.30, b = 0.20, c = 0.10, M_+ = 2, M_- = 5$. Значение параметра s выбрано равным 0.90

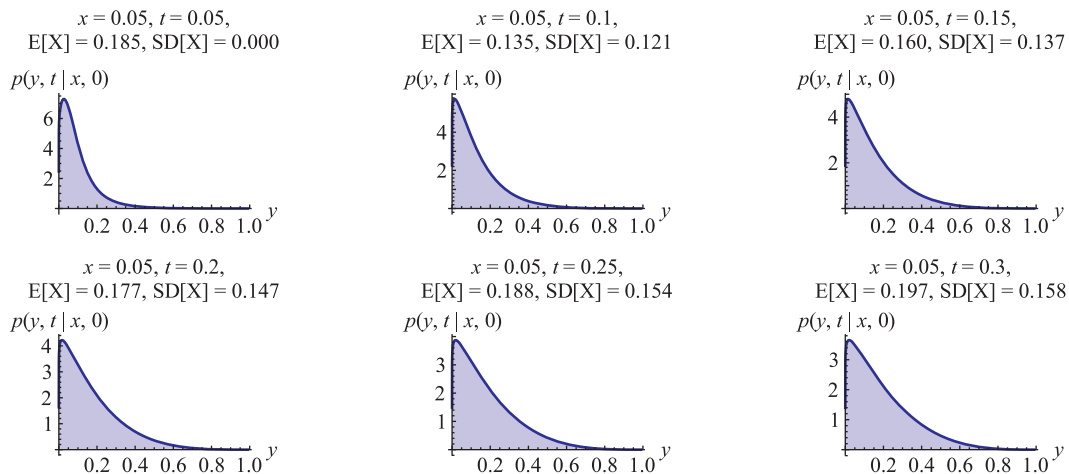


Рис. 3. Плотность переходной вероятности $p(y, t | x, 0)$ при $x = 0.01$ с течением времени t для параметров $a = 0.30, b = 0.20, c = 0.10, M_+ = 2, M_- = 5$. Значение параметра s выбрано равным -0.90

одобрения его деятельности. Различные сочетания и значения параметров a и c , отражающих степень влияния информационного воздействия и позиции лидеров мнений на рейтинги политика, могут объяснять, почему указанная взаимосвязь может как усиливаться, так и ослабевать. Математической основой данной модели является так называемый случайный процесс Якоби. При условии, что параметр, отвечающий за межличностное взаимодействие представителей электората b , не превышает величины $\frac{1}{2}(a+cM)$, единственным стационарным распределением данного процесса будет бета-распределение, параметры которого определяются параметрами модели.

Также исследован характер границ диапазона возможных значений рейтинга одобрения политика. Показано, что граница 0 может быть границей выхода (достигнув ее, невозможно более достигнуть никакого другого состояния) при абсолютно негативном ($s = -1$) потоке новостей (либо полном их отсутствии ($a = 0$)) вместе с отсутствием «группы поддержки» из числа лидеров мнений ($M_+ = 0$). Данная ситуация соответствует варианту завершения политической карьеры в связи с полной утратой какого-либо одобрения у представителей электората. При выполнении неравенства $\min\{M_-, M_+\} > \frac{b}{c}$, то есть при наличии достаточно большого числа лидеров мнений, как сторонников, так и противников данного политика, границы $l = 0$, $r = 1$ будут границами входа и перестанут быть достижимыми из внутренних состояний интервала (l, r) , вне зависимости от тональности информационных сообщений. Таким образом, при выполнении указанного условия достижение консенсуса в мнениях, относительно данного политического лидера, будет невозможным. Далее, в работе представлены формулы расчета значений переходных вероятностей и на условном примере продемонстрировано то, насколько тональность новостных сообщений способна повлиять на распределение возможных значений рейтинга одобрения политического лидера.

Применение предложенной модели на практике позволит дать численную оценку того, как изменение тональности и числа сообщений, относящихся к данному политику, способно повлиять на рейтинги его одобрения. Использование аппарата теории случайных процессов открывает возможности по проведению симуляций различных вариантов развития событий с использованием методов численного решения уравнения (11). Однако заметим, что практическое применение полученных результатов ставит отдельные и масштабные задачи разработки эффективного и релевантного для исследуемой области метода анализа текстов, отбора исследуемых цифровых ресурсов и регулярного мониторинга сообщений на них. Также подчеркнем, что в текущем варианте представленной модели мы рассматривали зависимости рейтинга одобрения политика от параметров информационного фона, считая последний заданным. В действительности имеет место и обратная зависимость — зависимость таких характеристик, как тональность сообщений от показателей рейтинга. Исследование данных взаимосвязей может служить развитием представленной статьи.

Список литературы (References)

- Бородин А. Н. Случайные процессы. — СПб.: Лань, 2013.
Borodin Sluchaynye protsessi [Random processes]. — St. Petersburg: Lan', 2013 (in Russian).
- Гихман И. И., Скороход А. В. Введение в теорию случайных процессов. — М.: Наука, 1965.
Gikhman I. I., Skorokhod A. V. Vvedeniye v teoriyu sluchaynykh protsessov [Introduction to the theory of random processes]. — Moscow: Nauka, 1965 (in Russian).
- Михайлов А. П., Маревцева Н. А. Модели информационной борьбы // Математическое моделирование. — 2011. — Т. 23, № 10. — С. 19–32.
Mikhailov A. P., Marevtseva N. A. Models of information warfare // *Mathematical Models and Computer Simulations*. — 2012. — Vol. 4, No. 3. — P. 251–259. (Original Russian paper: *Mikhailov A. P., Marevtseva N. A. Modeli informatsionnoi bor'by* // *Matematicheskoe modelirovaniye*. — 2011. — Vol. 23, No. 10. — P. 19–32.)

- Михайлов А. П. и др.* Развитие модели распространения информации в социуме // Математическое моделирование. — 2014. — Т. 26, № 3. — С. 65–74.
Mikhailov A. P. et al. Development of a model of information dissemination in society // Mathematical Models and Computer Simulations. — 2014. — Vol. 26, No. 3. — P. 65–74 (in Russian).
- Abramowitz M., Stegun I. (ed.)* Handbook of Mathematical Functions, With Formulas, Graphs, and Mathematical Tables. — New York, USA: Dover Publications, Inc., 1964.
- Alexa Internet — The top 500 sites on the web [Электронный ресурс]: <https://www.alexa.com/topsites/> (дата обращения: 15.07.2019).
Alexa Internet — The top 500 sites on the web [Electronic resource]: <https://www.alexa.com/topsites/> (accessed 15.07.2019).
- Alfarano S., Lux T., Wagner F.* Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach // Journal of Economic Dynamics and Control. — 2008. — Vol. 32, No. 1. — P. 101–136.
- Barberá P., Rivero G.* Understanding the political representativeness of Twitter users // Social Science Computer Review. — 2015. — Vol. 33, No. 6. — P. 712–729.
- Bhattacharya R. N., Waymire E. C.* Stochastic processes with applications. — Society for Industrial and Applied Mathematics, 2009.
- Bollen J., Mao H., Pepe A.* Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena // Fifth International AAAI Conference on Weblogs and Social Media. — 2011.
- Chen C. et al.* Content-based influence modeling for opinion behavior prediction // Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. — 2016. — P. 2207–2216.
- Carro A., Toral R., San Miguel M.* Markets, herding and response to external information // PloS one. — 2015. — Vol. 10, No. 7.
- Cummings D., Oh H., Wang N.* Who needs polls? Gauging public opinion from Twitter data // Unpublished manuscript. — 2010.
- De A. et al.* Learning a linear influence model from transient opinion dynamics // Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management. — ACM, 2014. — P. 401–410.
- Delbaen F., Shirakawa H.* An interest rate model with upper and lower bounds // Asia-Pacific Financial Markets. — 2002. — Vol. 9, No. 3. — P. 191–209.
- Elliott J.* Eigenfunction expansions associated with singular differential operators // Transactions of the American Mathematical Society. — 1955. — Vol. 78, No. 2. — P. 406–425.
- Föllmer H., Schweizer M.* A microeconomic approach to diffusion models for stock prices // Mathematical Finance. — 1993. — Vol. 3, No. 1. — P. 1–23.
- Gardiner C. W. et al.* Handbook of stochastic methods. — Springer, 1985.
- Gayo-Avello D.* I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper — A Balanced Survey on Election Prediction using Twitter Data // arXiv preprint arXiv:1204.6441. — 2012.
- Jedrzejewski A., Sznajd-Weron K.* Impact of memory on opinion dynamics // Physica A: Statistical Mechanics and its Applications. — 2018. — Vol. 505. — P. 306–315.
- Karlin S., Taylor H. E.* A second course in stochastic processes. — Elsevier, 1993.
- Kirman A.* Ants, rationality, and recruitment // The Quarterly Journal of Economics. — 1993. — Vol. 108, No. 1. — P. 137–156.

- Kononovicius A., Gontis V.* Three-state herding model of the financial markets // EPL (Europhysics Letters). — 2013. — Vol. 101, No. 2. — P. 28001.
- Kononovicius A., Gontis V.* Control of the socio-economic systems using herding interactions // Physica A: Statistical Mechanics and its Applications. — 2014. — Vol. 405. — P. 80–84.
- McKean H.P.* Elementary solutions for certain parabolic partial differential equations // Transactions of the American Mathematical Society. — 1956. — Vol. 82, No. 2. — P. 519–548.
- Mälntylä M. V., Graziotin D., Kuutila M.* The evolution of sentiment analysis — A review of research topics, venues, and top cited papers // Computer Science Review. — 2018. — Vol. 27. — P. 16–32.
- Nakov P. et al.* SemEval-2016 task 4: Sentiment analysis in Twitter // Proceedings of the 10th international workshop on semantic evaluation (semeval-2016). — 2016. — P. 1–18.
- Øksendal B.* Stochastic differential equations. — Springer, Berlin, Heidelberg, 2003.
- O'Connor B. et al.* From tweets to polls: Linking text sentiment to public opinion time series // Fourth International AAAI Conference on Weblogs and Social Media. — 2010.
- Peralta A.F. et al.* Analytical and numerical study of the non-linear noisy voter model on complex networks // Chaos: An Interdisciplinary Journal of Nonlinear Science. — 2018. — Vol. 28, No. 7.
- SimilarWeb — Top Websites Ranking [Электронный ресурс]: <https://www.similarweb.com/top-websites/> (дата обращения: 15.07.2019).
- SimilarWeb — Top Websites Ranking [Electronic resource]: <https://www.similarweb.com/top-websites/> (accessed 15.07.2019).
- Stroock D. W., Varadhan S. S.* Multidimensional diffusion processes. — Springer, 2007.
- Van Kampen N. G.* Stochastic processes in physics and chemistry. Vol. 1. — Elsevier, 1992.

