

МОДЕЛИРОВАНИЕ АКТИВНОСТИ ИНТЕРНЕТ-БОТОВ В ПОЛИТИЧЕСКОЙ ОНЛАЙН-КОММУНИКАЦИИ⁸²

Филиппов И.Б., Юрескул Е.А., Стукал Д.К., Ахременко А.С.

Национальный исследовательский университет «Высшая школа экономики»,
Россия. г. Москва, ул. Мясницкая, д. 20,

ibfilippov@gmail.com, eyureskul@hse.ru, dstukal@hse.ru, aakhremenko@hse.ru

В этой работе представлена новая агентно-ориентированная модель политической коммуникации с участием ботов – автоматизированных аккаунтов социальных сетей. Модель предполагает максимально точное – в рамках неизбежных упрощений – воспроизведение алгоритмов социальной сети Twitter. Содержательно модель ориентирована на анализ политической коммуникации.

Ключевые слова: интернет-боты, политическая коммуникация, агентно-ориентированная модель, вычислительная модель, Twitter

Введение

Социальные сети (Social Networking Sites, SNS) уже достаточно продолжительное время находятся в фокусе исследований на стыке политических и компьютерных наук. Кейсы успешной протестной самоорганизации во время «Арабской весны» и Евромайдана показали, что социальные сети могут служить площадкой для мобилизации населения и организации коллективного действия. Помимо этого, социальные сети стали эффективной площадкой политической коммуникации: виртуальное общение стало заменой традиционным платформам для обмена политизированной информацией и формирования политических позиций индивидов. Последнее особенно релевантно для исследователей-политологов, поскольку сформированные в ходе онлайн-дискуссий политические позиции имеют долгосрочные последствия для реальных политических систем: в первую очередь, речь идет о результатах демократических выборов. Также важным аспектом распространения социальных сетей стала поляризация аудитории: доступ к политизированной информации может стать причиной идеологических и эмоциональных расколов [1].

Параллельно с этим возрос исследовательский интерес к ботам в социальных сетях (social bots) – автоматизированным алгоритмам, выполняющим заданные функции за человека. Боты являются неотъемлемой частью экосистемы SNS, и изначально они были направлены на автоматизацию рутинных задач: публикацию сообщений в установленное время и формирование новостной ленты (feed). Одновременно возникли «злонамеренные» (malicious) боты, публикующие исключительно рекламные сообщения (спам), ссылки на недобросовестные сайты (фишинг); или случайным образом добавляющие пользователей в друзья с целью манипуляции популярностью того или иного аккаунта. В последние 10 лет автоматизированные алгоритмы были в значительной степени улучшены: помимо рутинных задач, современные боты могут выполнять и более сложные задачи, в том числе распространение пропаганды, подавление нежелательной информации (suppression of dissent) или манипулирование структурой социальной сети [2]. Доступ к API социальной сети Twitter и появление многочисленных облачных бот-сервисов (Bots as a Service, BaaS) привели к массовому использованию ботов, в том числе для формирования общественного дискурса онлайн. Возникли такие понятия, как astroturfing (создание информационного фона вокруг политиков, имитирующего низовой активизм) и Twitter bombs (массированное распространение выгодной заказчику дискуссии на политическую тематику) [3].

Повышенный интерес к деятельности ботов возник в контексте многочисленных кейсов, в которых автоматизированные алгоритмы использовались для воздействия на общественное мнение в политических целях. К примеру, боты применялись во время президентской кампании во Франции в 2017 году для распространения негативной информации об Э. Макроне (Macronleaks) [4]; а также для манипулирования общественным мнением по поводу президентских выборов в Германии [5], голосованием по поводу Brexit [6] и референдумом о независимости Каталонии [7]. Но наиболее ярким кейсом, привлечшим внимание исследователей, стало использование ботов для манипулирования общественным мнением во время президентской компании Д. Трампа в 2016 году: исследования показали, что целые группы ботов внедрялись в Twitter-сообщества, дискредитировали кандидатов от Демократической партии и релевантные для них новости, одновременно распространяя позитивные сообщения о Трампе [3].

⁸² Исследование выполнено за счет гранта Российского научного фонда (проект № 20-18-00274), Национальный исследовательский университет «Высшая школа экономики»

Учитывая актуальность применения ботов в политических кампаниях, значительная часть современной исследовательской литературы по данной теме посвящена вопросам распознавания бот-аккаунтов в социальных сетях. Общепринятый подход к классификации разделяет пользователей социальных сетей на «живых» людей (human), полностью автоматизированные алгоритмы (bot) и алгоритмы, напрямую управляемые живым человеком (cyborg) [8]. Исходные попытки распознавания ботов включали в себя лабораторные эксперименты, в которых на небольшой выборке определялись эвристические механизмы, используемые живыми людьми для определения автоматизированных аккаунтов: имя пользователя, количество подписчиков, достоверность публикуемой информации, репутация и «вовлеченность» в ближайшее окружение [9]. Современные принципы классификации пользователей основаны на многоуровневом анализе как открытой информации пользовательского профиля, так и особенностей поведения, характерных для ботов. К примеру, алгоритм BotOrNot выделяет следующие черты аккаунта бота: отсутствие геоданных и данных о пользовательском устройстве; избыточная активность, среди которой преобладают репосты (ретвиты) чужих публикаций; преобладание односторонних подписок над взаимными; аккаунт создан недавно, а имя пользователя состоит из случайного набора символов, обладающего повышенной лингвистической энтропией [3]. Как правило, алгоритмы распознавания ботов используют принципы машинного обучения на большой выборке публикаций (сотни тысяч пользователей и миллионы сообщений), что требует значительных вычислительных ресурсов; в свою очередь, разработчики ботов адаптируются к алгоритмам распознавания, делая поведение машин более похожим на поведение «живых» пользователей.

Разработка алгоритмов распознавания позволила исследователям сформулировать ряд эмпирических предпосылок, лежащих в основе изучения деятельности ботов в социальных сетях. Так, исследования показали, что в ходе политизированных дискуссий боты воспринимаются живыми людьми так же, как рядовые пользователи [9], однако демонстрируют иную динамику публикаций (живые люди проявляют большую активность в ответ на информационные поводы, тогда как боты публикуют сообщения с постоянной частотой [3]); при этом боты, как правило, начинают общение с периферии социальной сети (малого количества друзей и подписок) и постепенно все более и более «встраиваются» в процесс онлайн-взаимодействия. При обмене политизированными мнениями боты напоминают фанатиков (zealots): там, где живые люди могут изменить свое мнение при контакте с опровергающей информацией, боты продолжают продвигать заданную хозяином повестку [10]. Наконец, ботам доступен ограниченный набор стратегий: позитивное или негативное воздействие; воздействие на информацию или на сеть (т.н. BEND framework [11]). Эти соображения легли в основу моделирования поведения ботов в социальных сетях.

В контексте важности изучения ботов в политических дискуссиях ключевой исследовательский вопрос, которым задаются ученые, звучит следующим образом: насколько эффективно применение ботов для манипулирования общественным мнением? Попытки ответить на него сопряжены с целым рядом методологических трудностей. Во-первых, проблематично отследить эффект бот-кампаний на конечное поведение избирателей: пользователи социальных сетей, как правило, не транслируют свои идеологические предпочтения в эксплицитной форме (для оценки идеологии приходится применять сложные математические модели анализа публикаций [12]); дополнительные сложности возникают при необходимости изучения динамики их предпочтений и изоляции эффекта пропаганды, проведенной ботами. При этом эмпирические исследования сетевыми методами показывают, что боты могут влиять на динамику взаимодействий в социальных сетях, а значит, в какой-то мере действительно формируют общественное мнение [7]. Во-вторых, микро-исследования и small-N лабораторные эксперименты не отражают всего масштаба агитационных и дезинформационных кампаний, тогда как проведение масштабных экспериментов на большой выборке настоящих пользователей невозможно по этическим соображениям [13].

Учитывая вышесказанное, наиболее перспективным направлением изучения ботов в социальных сетях представляется агентно-ориентированное моделирование (agent-based modelling, ABM). Парадигма ABM позволяет зафиксировать как индивидуальный уровень поведения пользователей (агентов), а сочетание абстрактных моделей поведения с методиками сетевого анализа – воспроизвести динамику социального взаимодействия в рамках большого сообщества пользователей. Первые разработки в данном направлении опирались на абстрактную модель эпидемии (Susceptible, Infected, Recovered, SIR) [14], однако такие подходы не учитывали специфику именно информационных взаимодействий: как правило, динамика мнения пользователя, увидевшего опровержение ложной информации, отличается от динамики выздоровления. Более совершенные

модели учитывали ограниченную рациональность пользователей, обусловленную невозможностью получить всю информацию о структуре сети [15].

Современные разработки в данной области учитывают особенности работы сайтов социальных сетей и специфику сетевой структуры. Так, в модели [11] пользователи делятся на «невежд», «ботов» и «контрпропагандистов». Ботам доступно всего две стратегии (информационных маневра): поддержка информационного сообщения (backing) или установление связей между не связанными ранее группами пользователей (bridging). Авторы показали, что обе стратегии предполагают одну и ту же схему поведения (внедрение в существующие дискуссии и распространение информации), однако эффективность информационных маневров напрямую зависит от того, насколько плотно бот включен в социальную сеть: боты, находящиеся на периферии сети, не могут значимо повлиять на распространение информации.

В работе [10] авторы с помощью абстрактной агентно-ориентированной модели тестируют теорию «спирали молчания». Согласно данной теории, живые пользователи постоянно оценивают собственную идеологическую позицию и ее соответствие ближайшему социальному окружению. Поскольку каждый человек неизбежно стремится соответствовать идеалам своей социальной группы, транслируемые в новостной ленте сообщения влияют на вероятность того, что пользователь будет высказывать собственную позицию: в случае, если он не видит широкой поддержки своих взглядов, он предпочтет промолчать. Боты в модели не подвержены этому эффекту, однако могут менять соотношение транслируемых сообщений в пользу той или иной идеологии, снижая таким образом вероятность распространения противоположных мнений живыми пользователями. Результаты численных экспериментов показали, что важную роль в эффективности бот-кампаний играет структура социальной сети: чем более плотно связаны между собой пользователи в локальной подсети, тем меньше возможностей у бота запустить «спираль молчания»; в свою очередь, боты с большим числом подписчиков обладают в глазах пользователей большим авторитетом, что повышает их эффективность. Аналогичная модель представлена в работе [13], использующей уже не абстрактную, а сетевую структуру: авторы вводят понятие «привлекательности» пользователя-бота, основанное на числе его подписчиков; и «климата мнений» – соотношения идеологических позиций ближайших соседей живого пользователя, которое влияет на вероятность трансляции того или иного сообщения.

В настоящей работе мы опираемся на современный опыт агентно-ориентированного моделирования социальной сети Twitter и проводим вычислительный эксперимент, в котором формулируем список факторов (как на уровне индивидуального пользователя, так и на уровне сетевой структуры), влияющих на эффективность информационных маневров, проводимых ботами.

1 Модель

Предлагаемая модель направлена на контролируемое воспроизведение в упрощённом виде политически значимых взаимодействий участников коммуникации в социальной сети Twitter. Ниже мы опишем основные предпосылки, из которых мы исходили при разработке дизайна обсуждаемой модели и покажем, как они были реализованы. После этого будет описано функционирование модели в целом и отмечено, какую роль в ней на данный момент играют боты.

В реальных социальных медиа взаимодействие пользователей всегда опосредованно интерфейсом системы и её внутренними алгоритмами. Пользователь взаимодействует с набором экранов, часть из которых выглядит для всех пользователей идентично, другая часть – персонализирована для каждого пользователя. Классическим примером персонализированного экрана в Twitter является «фид» (feed), то есть лента пользователя, куда складываются наиболее актуальные твиты, написанные авторами, на которых данный пользователь подписан. Этот экран генерируется внутренними алгоритмами сети заново каждый раз, когда пользователь заходит в сеть, а состав и порядок сообщений, туда попавших, зависит от того, на кого пользователь подписан и что он делал в предыдущие заходы.

Взаимодействие пользователя с другими участниками коммуникации в социальных сетях никогда не бывает непосредственным не только потому, что между пользователями всегда стоит интерфейс сети, который жёстко определяет механики возможных взаимодействий, но и потому, что коммуникация между пользователями как правило возникает как реакция на стимул (твит), предоставленный одним из коммуницирующих без изначального указания адресата. Типичный сценарий выглядит так: сначала один пользователь пишет твит, адресованный неограниченному кругу читателей, а потом уже завязывается публичное обсуждение этого твита с некоторыми из прочитавших. Это обстоятельство коренным образом отличает коммуникацию в социальных сетях от коммуникации в мессенджерах (хотя услуги мессенджера могут предоставляться тем же сайтом,

который предоставляет и услуги социальной сети). Таким образом, нужно отметить, что объектом, на который направлена коммуникация, является публикация, а не её автор. Единственной механикой взаимодействия между пользователями непосредственно является механика подписки/отписки.

Исходя из сказанного выше, взаимодействия в модели, которую мы предлагаем, состоят главным образом из взаимодействия пользователей с экранами, которые им генерирует внутренний алгоритм социальной сети. Между собой пользователи взаимодействуют напрямую только через подписку или отписку. Такая мотивация диктует необходимость иметь в модели объекты разных типов: во-первых, самих пользователей, во-вторых, их экраны, в-третьих – их публикации, и, наконец, саму алгоритмическую систему.

Все твиты и экраны пользователей имеют одинаковый набор алгоритмов и характеристик (хотя они и отличаются по наполнению), в то время как пользователи делятся на три категории: (а) лидеры, (б) обычные пользователи, (в) боты. Отличие лидеров от обычных пользователей состоит в том, что они более активны, могут работать с большим количеством постов, занимают центральное место в топологии сети пользователей и – самое главное в рамках данной работы – могут быть «хозяевами» ботов. Боты отличаются от пользователей тем, что они полностью копируют характеристики лидера-хозяина за исключением иерархических, обладают расширенным набором стратегий взаимодействия с сетью и составляют крайне плотную подсеть – все боты одного хозяина связаны между собой отношениями подписки.

Помимо различных объектов, в модели также присутствует три разных эпистемологических уровня. Первый уровень – это уровень политических ценностей, на котором могут взаимодействовать только сами пользователи. Каждый пользователь обладает экзогенно заданной идеологической позицией, которую он может передавать твитам своего авторства, а другие пользователи могут идеологические позиции этих твитов считывать. Алгоритмы сети считывать политическую позицию пользователей и сообщений не в состоянии. Второй уровень – уровень данных, отображаемых в интерфейсе. Знание о том, где топологически и иерархически находится аккаунт пользователя и характеристики успешности продвижения поста доступны для считывания как алгоритмам сети, так и другим пользователям. Наконец, на третьем уровне находятся недоступные пользователям данные, которыми оперируют только алгоритмы сети. Это статистические данные о пользователях и состоянии сети, метаданные публикаций и информация о всех действиях пользователей.

Предлагаемая модель работает в следующем порядке. Сначала экзогенно задается набор управляющих параметров модели (таблица 1).

Таблица 1. Управляющие параметры модели

Параметр	Содержание
M	Число исходящих рёбер у пользователей при инициализации
Alpha	Параметр распределения активности
numleaders	Число лидеров
max_passiveness	Параметр максимальной активности
alpha2	Параметр распределения размеров прочитываемого фида
lifespan	Время актуальности твита
leader_boost	Дополнительная вероятность написать твит для лидера
action_probability	Вероятность подписки при ретвите, вероятность написать собственный твит, 2 x вероятность отписки
totalP	Общая популяция сети (включая ботов)
P	Число не-ботов в сети
B	Доля ботов
Bots	Число ботов
tolerance	Максимально допустимое расхождение позиций
Steps	Число ходов в модели
botshareFirst	Список долей ботов по лидерам
leader_positions	Список позиций лидеров

Затем генерируется сеть пользователей. Для этого мы используем направленную версию алгоритма Барабаши-Альберт [16] с заданным числом исходящих рёбер (здесь это число равно 4).

Алгоритм Барабаши-Альберт генерирует граф с лидерской топологией, то есть распределение степеней вершин, что в нашей модели соответствует распределению числа подписчиков, подчиняется степенному распределению, наблюдаемому в реальных социальных сетях. На этом этапе все пользователи равноправны, а ботов в модели ещё нет.

Так как модель имеет лидерскую топологию, в ней можно выделить несколько самых популярных вершин, каждой из которых присваивается тип лидера. На рисунке 1 приведён пример сети из 50 вершин, получающейся в результате генерации описанным образом, прямоугольниками отмечены вершины, которые могут претендовать на лидерский класс. Лидерский класс присваивается нескольким самым популярным вершинам, пока число лидеров не станет равно параметру numleaders.

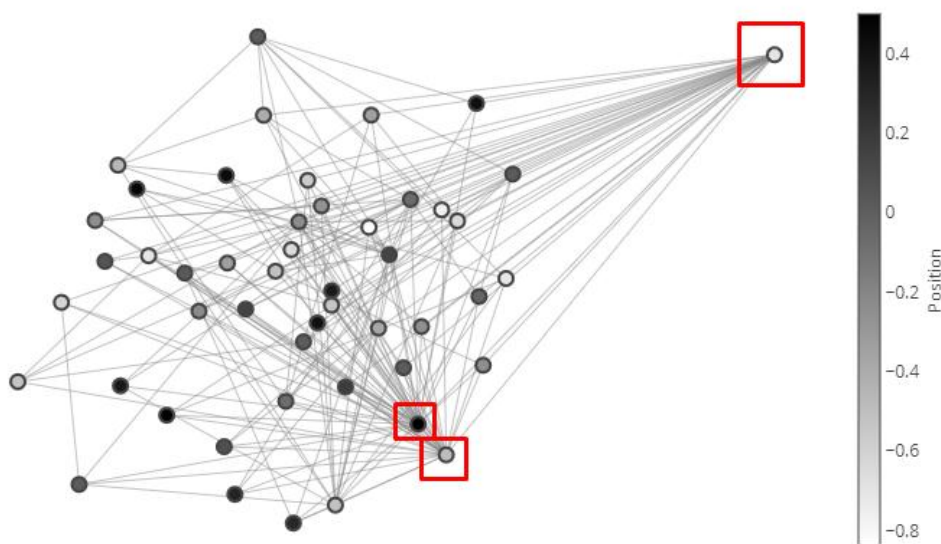


Рис. 1. Пример сети с 50 вершинами; прямоугольниками отмечены претенденты на статус лидера

После генерации исходной сети лидерам присваиваются значения идеологической позиции из `leader_positions`, списка позиций для лидеров. Позиции присваиваются контролируемо, самый популярный лидер получает первую позицию из списка, второй по популярности – вторую, и так далее. После того, как лидерам выданы все позиции, начинается процесс генерации позиции для пользователей. Пока в графе остаются пользователи без позиции, выполняется следующее:

- 1) Позиции генерируются из Beta-распределения с параметрами (2, 2), после чего они удваиваются, и из них вычитается 1. Таким образом, полученная позиция находится в пределах от -1 до 1.
- 2) Позиция присваивается случайной вершине из списка вершин, которые ещё не имеют позиции. Если позиция вершины отличается от средней позиции соседей данной вершины более, чем на `tolerance`, подбирается другая вершина.
- 3) Если сгенерированная позиция не может быть присвоена ни одной вершине, генерируется другая позиция.

После генерации позиций, все пользователи получают из степенного распределения характеристики активности и работоспособности – число постов, которые могут быть прочитаны за одну сессию (один ход активности). Лидеры получают к этим характеристикам бонус.

После того, как сеть с «реальными» пользователям инициализирована, в неё добавляются боты. Каждый лидер получает по соответствующей доле ботов от популяции ботов `bots`, каждый бот копирует идеологическую позицию хозяина, его активность и работоспособность. После раздачи ботов лидерам, все боты каждого хозяина связываются между собой. В результате сеть начинает выглядеть так, как на рисунке 2, на котором изображена та же сеть, что на рисунке 1 после добавления одному из лидеров 6 ботов. Группа ботов и их лидер изображены в прямоугольнике.

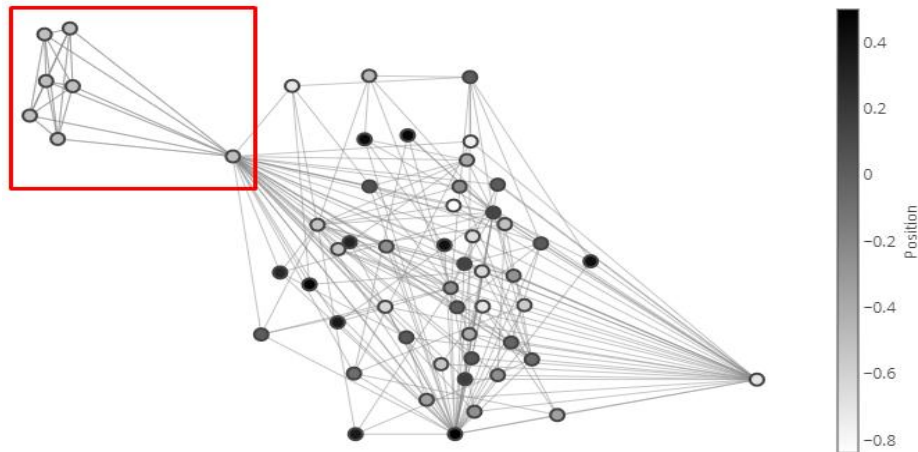


Рис. 2. Присоединение ботов к сети

После этого сеть полностью инициализирована, всем вершинам в сети добавляются пустые фиды, создаются глобальные пустые справочники твитов и ретвитов и работа модель запускается.

В начале каждого хода создаётся копия справочника ретвитов и словаря ретвитов. На информацию из копий при принятии решений будут опираться алгоритмы сети и пользователи, в то время как исходная сеть будет подвергаться изменениям. Это необходимо для того, чтобы все вершины принимали свои решения «одномоментно» и порядок моделирования действий вершин не оказывал влияния на результат. После того, как копии сделаны, в той версии сети, в которую пользователи будут вносить изменения, проводится чистка фидов от твитов, со времени создания которых прошло больше чем *lifespan* ходов. После этого для каждой вершины выполняется следующий порядок действий:

1. Определяется, будет ли вершина активна на данном ходу. Если нет, следующие пункты не выполняются;
2. Из всего потока поступивших пользователю постов формируется актуальный фид, который будет полностью прочитан пользователем;
3. Пользователь прорабатывает фид, при выполнении определённых условий отписывается от/ подписывается на некоторых авторах прочитанных твитов, выбирает твит для возможного ретвита;
4. Если пользователь – бот, а не один из твитов в фиде не был выбран для ретвита, то он случайным образом выбирает для ретвита один из исторических твитов своего хозяина. При выборе все твиты равноправны.
5. С заданной вероятностью (*action_probability* для обычного пользователя и *action_probability + leader_boost* для лидера) пользователь пишет свой твит.
6. Результатом выполнения всех этих действий становится список твитов, который отправится в фиды ко всем подписчикам пользователя.

После того, как для всех вершин выполнена данная последовательность действий, полученные списки твитов применяются для модификации фидов вершин, а в модели начинается на следующий ход.

Определение активности вершины опирается на две характеристики каждой вершины. Первая характеристика – *clock*, это внутренние «часы бездействия» пользователя. Помимо этого, у каждой вершины есть экзогенно заданный параметр *passiveness*. Если пользователь – лидер, то он получает дополнительный бонус от 0 до 1 к этой вероятности. Вершина будет активна с вероятностью, определяемой по следующей формуле:

$$P(\text{activity}) = \begin{cases} \frac{1}{(\text{passiveness} - \text{clock} + 1)} & | \text{leader} = 0 \\ \frac{1}{(\text{passiveness} - \text{clock} + 1)} + \text{leader_boost} & | \text{leader} = 1 \end{cases} \quad (1)$$

Если вершина проходит проверку активности, тогда clock обнуляется. В противном случае к текущему значению прибавляется 1.

Хотя сам фид формируется из твитов, отправленных в него вершинами, на которые подписан данный пользователь, в условиях, когда внимание пользователя ограничено, критически важной является ранжирование твитов. Алгоритм сети пытается сделать это, опираясь на свои представления о полезности данного твита для пользователя. Алгоритм «не знает», какова идеологическая позиция пользователя, поэтому ему остаётся опираться только на данные об успешности автора и данного твита. Твиты в фиде упорядочиваются по убыванию по внутренней оценке алгоритма score. Оценка score полезности конкретного твита для пользователя определяется по следующему правилу:

где n – число вхождений твита в фид через разные источники, $indegree$ – число подписчиков автора твита, $retweets$ – число ретвитов к моменту просмотра, $subscribed_followings$ – число подписок данного пользователя, подписанных на автора твита. Все эти переменные с префиксом $max_$ - максимальные значения данных переменных во всём фиде.

Проработка фида пользователем состоит из двух элементов. Во-первых, пользователь оценивает каждый твит на основании его совпадения с собственной позицией, числа подписчиков автора и отношения числа ретвитов не-ботами к просмотрам (этот член призван смоделировать некоторую меру качества твита в смысле его способности нравиться читателю). Пользовательская оценка качества твита $value$ формируется по следующей формуле:

$$value = \left(1 - \sqrt{(\text{position}_{\text{reader}} - \text{position}_{\text{tweet}})^2}\right) \times (\ln(\text{indegree} + 1) + 1) \times \left(\frac{\text{retweets}_{\text{real_users}}}{\text{views}_{\text{real_users}}} + 1\right) \quad (2)$$

Помимо оценки полезности, пользователь также проверяет, не превосходит ли модуль разницы между $\text{position}_{\text{reader}}$ и $\text{position}_{\text{tweet}}$ максимально допустимый порог $tolerance$. Если это происходит, с вероятностью $\text{action_probability}/2$ происходит отписка от автора твита или того, за счёт чьего ретвита данный пост попал к пользователю в фид. После выбора самого полезного твита с вероятностью $\text{action_probability}$ происходит ретвит, и с такой же вероятностью в случае ретвита происходит подписка на автора твита, если до этого её не было.

В такой конфигурации боты могут влиять на коммуникацию через повышение его внутренней оценки за счёт влияния на оценку score твитов их хозяина, так как боты увеличивают и число подписчиков своего хозяина, и число ретвитов его постов. Помимо этого, более редким, но тоже важным механизмом влияния ботов является непосредственное участие ботов в коммуникации: за счёт ретвитов их постов хозяином боты могут стать видимыми для остальных пользователей и быть интегрированными в граф коммуникации обычных пользователей, или, по крайней мере, подписаться на тех, кого ретвитил их хозяин, и прорабатывать их твиты в том числе. То, что за счёт действий хозяина боты начинают инкорпорироваться в общую сеть, иллюстрирует рисунок 3. На этом рисунке представлена сеть пользователей, изображённая на рисунке 2, 90 ходов спустя. Боты, которые изначально существовали отдельной группой, получают связи с другими вершинами, и получают возможность работать с твитами не только хозяина, но и других вершин.

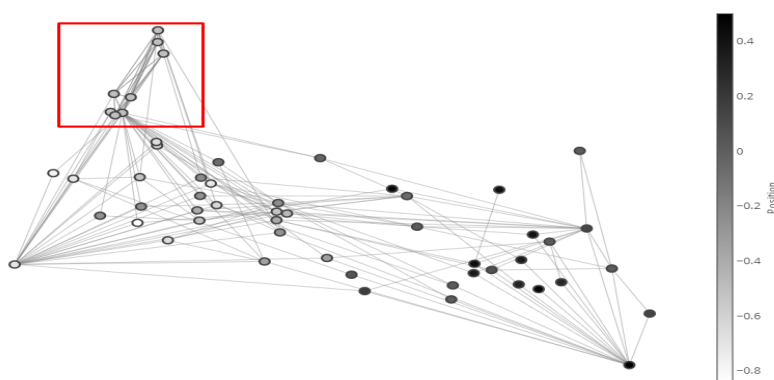


Рис. 3. Сеть пользователей и ботов (в прямоугольнике) через 90 ходов после инициализации модели

2 Численный эксперимент

Для исследования заложенных в модели возможностей ботов влиять на коммуникацию в социальной сети мы провели серию вычислительных экспериментов, в рамках которых исследовались особенности коммуникационной сети при разной доле ботов, созданных лидерами. Доля ботов менялась от 0.01 до 0.46 с шагом 0.05. Измерялись два основных показателя включенности ботов в коммуникацию – средняя доля ботов среди подписок и прочтений обычных пользователей. С учетом описанных выше каналов влияния ботов на коммуникацию, гипотеза состояла в положительной связи между долей ботов в сети и двумя измеряемыми признаками.

Анализ экспериментальных данных позволил подтвердить сформулированную гипотезу. Коэффициент корреляции Пирсона равен 0.84 (p -value = 0.002) для доли ботов в прочтениях и 0.89 (p -value < 0.001) для доли ботов в подписках. Визуализация результатов экспериментов (см. Рис. 4) позволила также выявить нелинейный вид этой зависимости, характеризующийся затуханием эффекта, т.е. уменьшением предельного эффекта доли ботов по мере увеличения этой доли. Изучению факторов, обуславливающих затухание эффекта, будет посвящено дальнейшее исследование данной вычислительной модели.

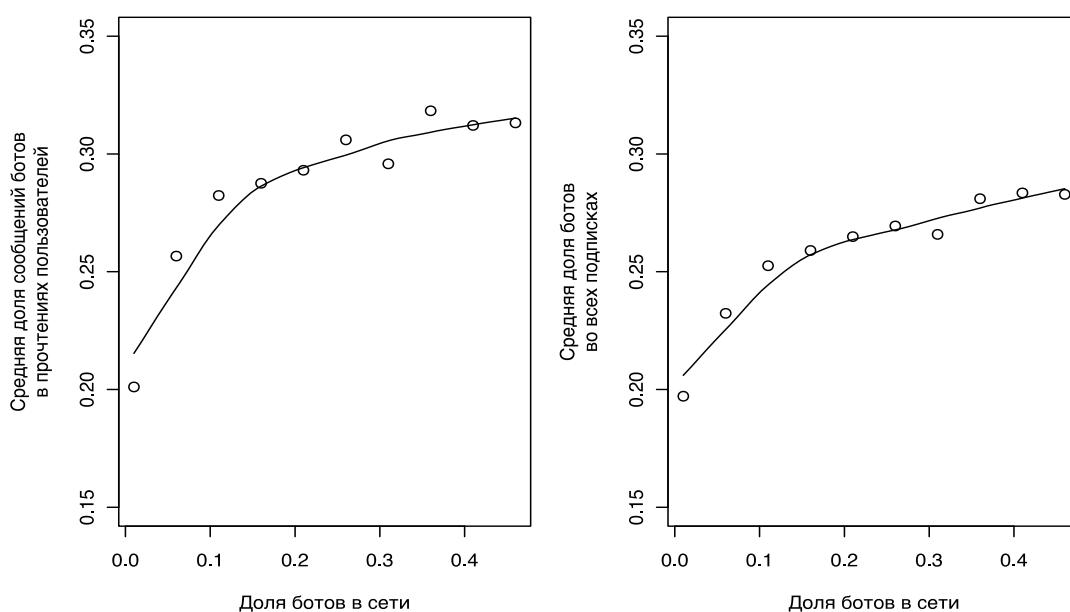


Рис. 4. Эффекты доли ботов в сети

Заключение

В этой работе была представлена новая агентно-ориентированная модель политической коммуникации с участием ботов – автоматизированных аккаунтов социальных сетей. Модель предполагает максимально точное – в рамках неизбежных упрощений – воспроизведение алгоритмов социальной сети Twitter. Содержательно модель ориентирована на анализ политической коммуникации; именно этим обусловлена особая роль, которую в модели играют лидеры – вершины с максимальным числом связей (степенями вершин). Кроме того, сам по себе интервал $[-1, 1]$, на котором определены политические позиции, предполагает возможности исследования процессов идеологической поляризации.

Среди наиболее перспективных направлений дальнейших исследований отметим исследование эффективности различных стратегий воздействия ботов на политическую коммуникацию. С этой целью нами запланирована масштабная серия вычислительных экспериментов над описанной выше моделью.

Литература

1. *Iyengar S., Sood G., Lelkes Y.* Affect, not ideology: a social identity perspective on polarization // *Public Opinion Quarterly*. Vol. 76. 2012. – P. 405-431.
2. *Beskow D.M., Carley K.M.* It's all in a name: detecting and labelling bots by their name // *Computational and Mathematical Organizational Theory*. Vol. 25. 2019. – P. 24–35.
3. *Bessi A., Ferrara E.* Social bots distort the 2016 U.S. Presidential election online discussion // *First Monday*. Vol. 21. 2016, №11.
4. *Ferrara E.* Disinformation and social bot operations in the run up to the 2017 French presidential Election // *First Monday*. Vol. 22. 2017, №8.
5. *Neudert L.-M., Kollanyi B., Howard Ph.* Junk News and Bots during the German Parliamentary Election: What are German Voters Sharing over Twitter? // *Data Memo*. Oxford, UK: Project on Computational Propaganda. 2017, №7.
6. *Howard Ph., Kollanyi B.* Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum // *arXiv:1606.06356 [Physics]*. 2016, June 20.
7. *Stella M., Ferrara E., De Domenico M.* Bots increase exposure to negative and inflammatory content in online social systems // *Proceedings of the National Academy of Sciences*. Vol. 115. 2018, №49. – P. 12435-12440.
8. *Chu Z., Gianvecchio S., Wang H., Jajodia S.* Detecting Automation of Twitter Accounts: Are You a Human, Bot, or Cyborg? // *IEEE Transactions on Dependable and Secure Computing*. Vol. 9. 2012, №6. – P. 811-824.
9. *Edwards Ch., Edwards A., Spence P., Shelton A.* Is that a bot running the social media feed? Testing the differences in perceptions of communication quality for a human agent and a bot agent on Twitter // *Computers in Human Behavior*. Vol. 33. 2014. – P. 372-376.
10. *Cheng Ch., Luo Y., Changbin Y.* Dynamic mechanism of social bots interfering with public opinion in network // *Physica A: Statistical Mechanics and its Applications*. Vol. 551. 2020.
11. *Beskow D., Carley K.* Agent Based Simulation of Bot Disinformation Maneuvers in Twitter // *Proceedings of the 2019 Winter Simulation Conference*. 2019. – P. 750-761.
12. *Barberá P., Rivero G.* Understanding the Political Representativeness of Twitter Users // *Social Science Computer Review*. Vol. 33. 2015. №6. – P. 712-729.
13. *Ross B., Pilz L., Cabrera B., Brachten F., Neubaum G., Stieglitz S.* Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks // *European Journal of Information Systems*. Vol. 28. 2019, №4. – P. 394-412.
14. *Daley D. J., Kendall D. G.* Stochastic Rumours // *Journal of Applied Mathematics*. Vol. 1. 1965, №1. – P. 42-55.
15. *Carley K. M., Martin M. K., Hirshman B. R.* The Etiology of Social Change // *Topics in Cognitive Science*. Vol. 1. 2009, №4. – P. 621–650.
16. *Barabási A.-L., Albert R., Jeong H.* Scale-Free Characteristics of Random Networks: The Topology of the World-Wide Web // *Physica A: Statistical Mechanics and its Applications*. Vol. 281. 2000. P. 69-77.