УДК 519.8

DOI: http://doi.org/10.25728/pu.2021.2.2

ИНФОРМАЦИОННЫЕ СООБЩЕСТВА В СОЦИАЛЬНЫХ СЕТЕВЫХ СТРУКТУРАХ. Ч. 2. Математические сетевые модели формирования сообществ[1]

Д.А. Губанов, И.В. Петров

Аннотация. Дан обзор математических моделей формирования информационных сообществ в условиях неопределенности, в которых индивиды стремятся прийти к истинному представлению относительно интересующего их вопроса. Подробно рассмотрены модели динамики представлений, в которых моделируется изменение представлений одних индивидов под влиянием других индивидов в социальной сети с нетривиальной структурой. Приведено два класса моделей: модели с рациональными (байесовскими) индивидами и модели с наивными (эвристическими) индивидами. Для моделей каждого из классов приведены условия формирования информационных сообществ в социальных сетевых структурах. В моделях с байесовскими агентами для возникновения различных информационных сообществ, как правило, рациональность индивидов ограничивается и вводятся допущения о различной информированности индивидов с учетом структуры сети. В моделях с наивными индивидами применяются различные модификации механизма формирования представлений.

Ключевые слова: социальные сетевые структуры, информационное сообщество, формирование информационных сообществ, формирование представлений, наивные индивиды, рациональные индивиды.

ВВЕДЕНИЕ

Как было отмечено в первой части [1] настоящего обзора, выявление и исследование в социальных сетях информационных сообществ — множеств индивидов, обладающих близкими и устойчивыми представлениями по заданному вопросу, — представляет собой важную задачу во многих предметных областях. Для решения этой задачи необходимо иметь понимание закономерностей динамики представлений в социальной сети. Особенности обработки информации индивидом исследуются в области когнитивистики, психологии и социальной психологии (см., например, работы [2, 3]), а для моделирования динамики представлений в сетях с учетом этих особенностей разрабатываются формальные микроуровневые модели (см., например, работы [4—8]). Модели динамики представлений и формирования информационных сообществ в социальных сетях, имеющие как микроэкономические, так и когнитивные и социально-психологические основания, рассмотрены в первой части [1] обзора. В ней кратко изложена концепция информационного сообщества, а также введена общая концептуальная модель обработки информации и принятия решений индивидом в социальной сети, в рамках которой агенты стремятся устранить неопределенность относительно изучаемого параметра внешней среды, наблюдая как внешние сигналы, так и действия своих соседей в социальной сети. Рассмотрены факторы, влияющие на динамику представлений в сети и на формирование информационных сообществ. Обзор моделей показал, что рациональные агенты в обществе с вырожденной структурой, как правило, достигают истинного представления относительно исследуемого вопроса, и для возникновения раз-

[1] Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 19-17-50225 и № 18-29-22042.

личных информационных сообществ необходимо модифицировать рациональность индивидов и их информированность тем или иным образом (см., например, работы [9—11]). За рамками рассмотрения в первой части обзора осталось изучение влияния на формирование информационных сообществ двух ключевых факторов: структуры социальной сети и агентов с эвристическими правилами обновления представлений. Эти вопросы будут рассмотрены в настоящей части обзора.

Структура второй части обзора такова: в § 1 рассмотрено формирование информационных сообществ в моделях с байесовскими агентами, взаимодействующими в сети; в § 2 рассмотрено формирование информационных сообществ в сети агентов с эвристическими правилами обновлений представлений.

1. ФОРМИРОВАНИЕ ИНФОРМАЦИОННЫХ СООБЩЕСТВ В ОБЩЕСТВЕ С СЕТЕВОЙ СТРУКТУРОЙ В МОДЕЛЯХ С БАЙЕСОВСКИМИ АГЕНТАМИ

В моделях с сетевой структурой задается конечное или счетное число индивидов. Основными элементами сетевых моделей формирования представлений являются структура информированности, множество действий агентов и их функции выигрыша, а также наблюдаемость действий других агентов. Опишем их детально.

Структура информированности агентов. Состояние природы — значение параметра $\theta \in \Theta$ — представляет собой ненаблюдаемую индивидами реализацию случайной величины. Каждый i-й индивид обладает частной информацией — частным сигналом s_i (случайная величина, распределение которой зависит от значения θ). Значение сигнала предоставляет информацию об истинном значении θ. Частные сигналы условно независимы относительно состояния природы θ. Частное представление индивида (private belief) задается изначально и не меняется во времени. Представление агента в некоторый момент времени t будет зависеть от его наблюдений в предыдущие периоды.

Действия агентов и их выигрыши. Каждый агент i может однократно, в заданный момент времени, выполнить действие $x_i \in X$, которое приводит к выигрышу $u(x_i, \theta)$. В момент выбора действия агент руководствуется субъективной вероятностью θ и ожидаемым выигрышем от выполнения действия $U = E_i[u(x_i, \theta)]$, учитывая всю имеющуюся у него информацию. Информативность действия агента для наблюдателей зависит от множества X.

В случае бинарных действий $X = \{0, 1\}$ и пространства состояний $\Theta = \{0, 1\}$ функция выигрыша задается как $u(x, \theta) = \theta - c$, $0 < c < 1$. В ситуации неопределенности выигрыш определяется так:

$$u(x) = (E[\theta] - c)x.$$

Стандартный способ задания процесса выбора с континуальными действиями — предположить, что агент выбирает действие $x \in R^1$, которое максимизирует ожидаемое значение квадратичной функции выигрыша:

$$u(x, \theta) = -E[(x - \theta)^2].$$

Оптимальное действие $x = E[\theta]$, в этом случае ожидаемый агентом выигрыш $U = -\mathrm{Var}(\theta)$.

Публичная информация (*public information*) *и история действий.* Порядок действий агентов (протокол взаимодействия) задан заранее. Агент t ($t \ge 1$) выбирает действие в момент t. История действий в этот момент задается так:

$$h_t = \{x_1, ..., x_{t-1}\}.$$

Агент t знает историю h_t в момент выбора действия. В начале периода t (перед принятием решения) общим знанием агентов является:
- априорное распределение вероятностей состояния природы θ,
- распределение частных сигналов и функции выигрышей всех агентов,
- история предыдущих действий h_t.

При этом выигрыши агентов не наблюдаемы.

С учетом описанных выше элементов *процесс обновления представлений* индивидов состоит в следующем. В момент времени $t \ge 1$ распределение вероятностей состояния природы θ, которое основано исключительно на общедоступной или публичной информации (h_t), называется *публичным или общественным представлением* (public belief — функция распределения состояния природы $F(\theta|h_t)$). Агент t использует публичное представление и частную информацию (сигнал s_t) для формирования своего представления о состоянии природы, имеющего распределение $F(\theta|h_t, s_t)$. Затем он выбирает действие, которое максимизирует зависящий от его представления выигрыш $E[u(x_t, \theta)]$. Оставшиеся агенты знают функцию выигрыша агента t и его модель принятия решений. Наблюдаемое действие x_t воспринимается ими как сообщение о доступной ему информации — частном сигнале s_t. С учетом этого агенты обновляют публичное представление $F(\theta|h_{t+1})$.

Примечание. Социальное научение является *эффективным*, если действие индивида полностью раскрывает его частную информацию. Это возможно, если множество допустимых действий достаточно велико.

Приведем в качестве базовых моделей модель обновления представлений с континуальными действиями и модель обновления представлений с дискретными действиями, в которых индивиды наблюдают действиях *всех* предшественников.

В *модели формирования представлений с континуальными действиями* состоянием природы является реализация в начальный момент времени случайной величины, имеющей нормальное распределение (нормальной случайной величины или нормально распределенного случайного вектора) $N(\bar{\theta}, 1/\rho_\theta)$. Задано счетное число индивидов $i = 1, 2, \ldots$ Каждый индивид i получает индивидуальный сигнал s_i, который равен сумме истинного значения и шума $\epsilon_i \sim N(0,1/\rho_\epsilon)$:

$$s_i = \theta + \epsilon_i.$$

Выигрыш агента $u(x, \theta) = -E[(x - \theta)^2]$. Индивид t выбирает действие $x_t \in R$. Публичной информацией на начало периода t является априорное распределение $N(\bar{\theta}, 1/\rho_\theta)$ и история предыдущих действий $h_t = \{x_1, \ldots, x_{t-1}\}$.

Предположим, что публичное мнение о значении θ в период t задается нормальным распределением $N(\mu_t, 1/\rho_t)$. Тогда это же предположение справедливо на момент $t = 1$ для параметров $\mu_1 = \theta$ и $\rho_1 = \rho_\theta$. Можно показать, что оно истинно для каждого следующего периода. В любой период обновление публичного представления проходит в три этапа.

- Расчет представления агента t. Он производится при помощи обновления публичного представления $N(\mu_t, 1/\rho_t)$ по правилу Байеса на основе частной информации $s_t = \theta + \epsilon$. Это представление является распределением $N(\tilde{\mu}_t, 1/\tilde{\rho}_t)$ с параметрами

$$\tilde{\rho}_t = \rho_t + \rho_\epsilon,$$

$$\tilde{\mu}_t = \alpha_t s_t + (1 - \alpha_t)\mu_t, \text{ где } \alpha_t = \rho_\epsilon/\tilde{\rho}_t.$$

- Решение агента t (действие x_t). Агент желает максимизировать выигрыш $-E[(x - \theta)^2]$, поэтому он выбирает действие, равное ожидаемому значению θ: $x_t = \tilde{\mu}_t$, т. е. $x_t = \alpha_t s_t + (1 - \alpha_t)\mu_t$.
- Социальное научение. Агенты сети наблюдают действие x_t и обновляют публичное представление относительно значения θ в течение следующего периода. Поскольку правило принятия решений агента t, а также величины α_t и μ_t известны всем агентам, то *наблюдаемое действие* x_t позволяет *полностью раскрыть частный сигнал* s_t. Публичная информация в конце периода t идентична информации агента t: $\mu_{t+1} = \tilde{\mu}_t$ и $\rho_{t+1} = \tilde{\rho}_t$. Следовательно, в периоде $t + 1$ представление по-прежнему нормально распределено $N(\mu_{t+1}, 1/\rho_{t+1})$, и процесс научения может быть продолжен далее. Отметим, что здесь история действий $h_t = \{x_1, \ldots, x_{t-1}\}$ информационно эквивалентна последовательности сигналов $(s_1, \ldots, s_{t-1})$.

Точность публичного убеждения увеличивается по закону $\rho_t = \rho_\theta + (t - 1)\rho_\epsilon$, т. е. дисперсия сойдется к нулю. При этом значимость частных сигналов стремится к нулю и агенты, соответственно, имитируют действия друг друга. В случае «зашумленности» наблюдений действий других агентов скорость социального научения падает [12]. Существуют модификации базовой модели социального научения [13], в которых агент в свой ход может *заплатить за нужную точность p* частного сигнала и затем выполнить действие. Тогда при минимальных предположениях о функции затрат $c(p)$ можно доказать, что агенты перестанут «покупать» сигнал после некоторого момента времени и социальное научение прекратится.

В *модели формирования представлений с дискретными действиями* состояние природы $\theta \in \Theta = \{0, 1\}$ устанавливается случайно в начальный момент времени, $\mu_1 = P(\theta = 1)$. Конечное число N или счетное число агентов индексируется целым числом t. Каждый агент получает частный симметричный сигнал $q > 1/2$: $P(s_t = \theta|\theta) = q$. Агент t выбирает действие $x_t \in \{0, 1\}$ в период t (и только в этот период). Выигрыш агента определяется состоянием природы:

$$u(x, \theta) = \begin{cases} 0, & x = 0, \\ \theta - c, & x = 1, \end{cases}$$

где $0 < c < 1$ (в классической BHW-модели (сокр. от фамилий ее авторов — *S. Bikhchandani, D. Hirshleifer, I. Welch* [14]) в качестве состояния природы θ выступает значение выигрыша агента от действия 1 (accept), а параметр c — это затраты от действия 1). Поскольку $x \in \{0, 1\}$, то выигрыш можно записать как $u(x, \theta) = (\theta - c)x$. В ситуации неопределенности агент рассчитывает выигрыш как ожидаемое значение $u(x, \theta)$ с учетом имеющейся у него информации.

Как и раньше, информация, которой располагает агент t, — это частный сигнал и история действий предыдущих агентов h_t. Публичное представление в начале периода t — вероятность состояния 1 с учетом общедоступной истории h_t:

$$\mu_t = P(\theta = 1|h_t).$$

Показано [14], что в таких моделях может быстро возникнуть *информационный каскад*: агенты в последовательности игнорируют свои частные сигналы и действуют так же, как и их предшественники, тем самым не предоставляя своим последователям новой информации, т. е. общество неэффективно агрегирует доступную информацию и может прийти к неверным убеждениям. Возможна модификация модели [15], в которой агенты приобретают информацию, если она может изменить их действие. Если информация может помочь отказаться от текущего консенсуса и является достаточно «недорогой», то агенты придут к правильным представлениям и действиям.

Таким образом, в моделях обновления представлений индивидов с канонической структурой социальной сети (в которой агент наблюдает действия всех предшественников) общество в результате взаимодействий будет представлять собой одно информационное сообщество с истинным или ложным представлением об интересующем его вопросе (см. выше условия истинности или ложности). Рассмотрим теперь модели обновления представлений байесовских агентов, в которых учитывается более сложная топология социальной сети.

Формирование информационных сообществ в сетях с нетривиальной структурой

Рассмотрим для начала показательную модель секвенциального социального научения с нетривиальной структурой. В работе [16] задается счетное множество агентов (индивидов), индексируемых $n \in \mathbb{N}$. Агенты последовательно и однократно принимают решения. Выигрыш агента n зависит от его действия и исходного состояния природы θ. Для простоты предполагается, что и состояние природы, и действия агентов являются бинарными, т. е. для агента n действие $x_n \in \{0, 1\}$, а состояние мира $\theta \in \{0, 1\}$. Выигрыш агента n

$$u_n(x_n, \theta) = \begin{cases} 1, & x_n = \theta, \\ 0, & x_n \neq \theta. \end{cases}$$

Также предполагается, что значения состояния мира равновероятны, т. е. $P(\theta = 0) = P(\theta = 1) = 1/2$. Состояние θ агентам неизвестно. Каждый агент формирует убеждение о состоянии мира, наблюдая частный сигнал $s_n \in \bar{S}$ ($\bar{S}$ — метрическое пространство) и действия остальных агентов. Сигналы условно независимо порождаются согласно вероятностной мере F_θ. Пара (F_0, F_1) называется *сигнальной структурой*. Предполагается, что меры F_0 и F_1 абсолютно непрерывны относительно друг друга, т. е. невозможен сигнал, который бы полностью раскрыл состояние мира.

Каждый агент n наблюдает действия только своих соседей в социальной сети из множества $B(n) \subseteq \{1, 2, ..., n - 1\}$. Каждая сетевая окрестность $B(n)$ порождается согласно некоторому распределению вероятностей Q_n, заданному на множестве всех подмножеств $\{1, 2, ..., n - 1\}$. Каждое распределение Q_n в последовательности $\{Q_n\}_{n \in N}$ не зависит от других распределений и от реализации частных сигналов. Последовательность $\{Q_n\}_{n \in N}$ формирует топологию социальной сети, которая, в отличие от реализованной окрестности $B(n)$ и индивидуального сигнала s_n, является общим знанием. Каноническая в литературе топология, когда каждый агент наблюдает все предыдущие действия, реализуется, если для любого $n \in \mathbb{N}$ вероятность окрестности $\{1, 2, ..., n - 1\}$ равна 1 в распределении Q_n. Возможны, конечно, и другие варианты, например, реализация модели случайного графа.

Информационное множество агента n задается как $I_n = \{s_n, B(n), x_k \text{ для всех } k \in B(n)\}$.

Обозначим множество всех возможных информационных множеств агента как $\boldsymbol{I}_n$. Стратегией агента σ_n: $\boldsymbol{I}_n \to \{0, 1\}$ является отображение из множества возможных информационных множеств во множество действий. Профиль стратегий — последовательность стратегий $\sigma = \{\sigma_n\}_{n \in N}$. Для заданного профиля σ последовательность действий в сети представляет собой случайный процесс $\{x_n\}_{n \in N}$, обозначим порожденную таким процессом меру как P_σ. Профиль стратегий σ^* является *совершенным байесовским равновесием* в чистых стратегиях в игре социального научения, если для любого $n \in N$ стратегия σ_n^* максимизирует ожидаемый выигрыш агента n при заданных стратегиях σ_{-n}^*.

Для заданного профиля σ ожидаемый выигрыш агента n от действия $x_n = \sigma_n(I_n)$ представляет собой $P_\sigma(x_n = \theta|I_n)$. Следовательно, для любого равновесия σ^*

$$\sigma_n^*(I_n) \in \operatorname*{Argmax}_{y \in \{0, 1\}} P_{(y, \sigma_{-n}^*)}(y = \theta|I_n).$$

Утверждается, что в такой игре социального научения существует совершенное байесовское равновесие в чистых стратегиях (Σ^* — множество таких равновесий).

Представляет интерес вопрос: приводит ли равновесное поведение к асимптотическому научению? Формально *асимптотическое научение* возникает в равновесии σ, если значение x_n сходится по вероятности к θ:

$$\lim_{n\to\infty} P_\sigma(x_n = \theta) = 1.$$

Действия агентов могут быть охарактеризованы как функция суммы двух апостериорных представлений — частного убеждения агента и *социального убеждения* (social belief). Утверждается, что в равновесии $\sigma \in \Sigma^*$ решение агента n, $x_n = \sigma_n(I_n)$, представляет собой

$$x_n = \begin{cases} 1, p_n + q_n > 1, \\ 0, p_n + q_n < 1, \end{cases}$$

иначе $x_n \in \{0, 1\}$. Здесь $p_n = P_\sigma(\theta = 1|s_n)$ — *частное убеждение*, $q_n = P_\sigma(\theta = 1|B(n), x_k, k \in B(n))$ — *социальное убеждение*.

Частное убеждение агента n не зависит от профиля стратегий. Пользуясь правилом Байеса, его можно записать так:

$$p_n = \left(1 + \frac{dF_0}{dF_1}(s_n)\right)^{-1}.$$

Определяется носитель индивидуальных убеждений — интервал $[\underline{\beta}, \bar{\beta}]$, где $\underline{\beta} = \inf\{r \in [0, 1]|P(p_1 \le r) > 0\}$ и $\bar{\beta} = \sup\{r \in [0, 1]|P(p_1 \le r) < 1\}$. Сигнальная структура имеет *ограниченные частные убеждения*, если $\underline{\beta} > 0$ и $\bar{\beta} < 1$, и *неограниченные*, если $\underline{\beta} = 1 - \bar{\beta} = 1$. Если частные убеждения являются неограниченными, то агенты могут получить сколь угодно сильный сигнал в пользу того или иного состояния.

Для изложения дальнейших результатов, связанных с асимптотическим социальным научением, приведем некоторые свойства сетевых топологий и сигнальных структур. Топология сети обладает свойством *расширения наблюдений* (expanding observations), если для всех $K \in N$

$$\lim_{n\to\infty} Q_n(\max_{b \in B(n)} b < K) = 0.$$

Справедлива теорема о том, что если топология сети $\{Q_n\}_{n \in N}$ не обладает свойством расширения наблюдений, то *не существует равновесия σ, в котором достигается асимптотическое научение*. Если топология сети не обладает свойством расширения наблюдений, то это эквивалентно тому, что существует конечное множество агентов, действия которых с положительной вероятностью будет наблюдать бесконечное число агентов, что, в свою очередь, не позволит им агрегировать рассредоточенную в сети информацию (такое конечное множество агентов называется *чрезмерно влиятельным* [16]).

Если же топология сети $\{Q_n\}_{n \in N}$ обладает свойством расширения наблюдений и из сигнальной структуры (F_0, F_1) следует неограниченность частных убеждений, то справедлива *теорема о возникновении асимптотического научения в каждом равновесии* $\sigma \in \Sigma^*$. Теорема, в частности, гарантирует научение в случае наличия агентов, которые влиятельны (в том смысле, что их действия видны всему обществу), но не чрезмерно, поскольку они не являются единственными источниками информации в сети. В качестве примера можно рассмотреть сеть, в которой действия первых K агентов наблюдаются всеми другими агентами, но при этом каждый агент также наблюдает за своим непосредственным соседом, т. е. $B(n) = \{1, 2, ..., K, n - 1\}$. Такая топология сети обладает свойством расширения наблюдений и, следовательно, ведет к научению в случае неограниченных частных убеждений. Этот вывод противоречит результатам в небайесовских моделях научения (см. работы [17, 18]), в которых новые убеждения агентов представляют собой средневзвешенную сумму частных убеждений и убеждений агентов, за которыми они наблюдают: если первые K агентов являются влиятельными в том смысле, что их действия видят остальные агенты, то асимптотического научения не будет.

Теперь пусть сигнальная структура (F_0, F_1) такова, что частные убеждения являются ограниченными, а топология сети $\{Q_n\}_{n \in N}$ удовлетворяет одному из условий:

— $B(n) = \{1, ..., n - 1\}$ для всех n (см. также статью [19]),

— $|B(n)| \le 1$ для всех n,

— существует такая константа M, что $|B(n)| \le M$ для всех n и с вероятностью 1

$$\lim_{n\to\infty} \max_{b \in B(n)} b = \infty,$$

тогда в любом равновесии $\sigma \in \Sigma^*$ асимптотическое научение достигнуто не будет. Из этого, в частности, следует отсутствие асимптотического научения в такой сети, в которой каждый агент n равномерно случайно и независимо выбирает $M \ge 1$ соседей из множества $\{1, ..., n - 1\}$.

В качестве отступления заметим, что в рассматриваемой модели агенты выполняют действие однократно, после чего получают выигрыш. Но в некоторых ситуациях *действие может быть отложено*, агенты могут обмениваться сообщениями о том, что они знают, не неся существенных затрат (однако, теряя время) и получая, в свою очередь, дополнительную информацию. Приведем пример функции выигрыша агента [20]:

$$u_i(x_i, \theta) = \\ = \begin{cases} \delta^\tau \pi, \text{если } x_{i,\tau} = \theta \text{ и } x_{i,t} = \text{wait для } t < \tau, \\ x, \quad \text{иначе,} \end{cases}$$

где $x_i = [x_{i,t}]_{t=0,1...}$ — последовательность действий i-го агента ($x_i \in \{\text{wait}, 0, 1\}$), $\pi > 0$ — выигрыш агента, $\delta \in (0, 1)$ — коэффициент дисконтирования. На качественном уровне аналогом такой двухэтапной модели в случае агентов с эвристическими правилами обновления представлений является модель [21], в которой агенты сначала формируют свои мнения, а затем одновременно выполняют действие в соответствии с функциями выигрыша.

Условия достижения научения в социальной сети [16] являются довольно мягкими. Типичным итогом работы байесовских моделей социального научения является консенсус, достигаемый в долгосрочной перспективе. Чтобы получить информационные сообщества с различными представлениями, необходимо ослабить требование рациональности индивидов социальной сети.

В частности, распространена концепция *квазибайесовского обновления агентов* [22—24], которая заключается в том, что каждый агент в сети считает, что действия других агентов вызваны исключительно их частными сигналами (такая концепция связана с *когнитивными ограничениями* — ограниченной глубиной вывода агента). В работе [24] исследуется секвенциальное (последовательное) социальное научение в социальной сети (сети наблюдений). Предполагается, что мир (природа) может находиться в одном из двух равновероятных состояний $w \in \{0, 1\}$. Имеется счетное множество агентов, индексируемых $i = 1, 2, 3, ...$, которые действуют однократно и по очереди (последовательно). Агент i на своем шаге получает частный сигнал о состоянии мира $s_i \sim N(1, \sigma^2)$, если $w = 1$, или $s_i \sim N(-1, \sigma^2)$, если $w = 0$, и, кроме того, наблюдает действия своих предшественников в ориентированной сети наблюдений $N_i \subseteq \{1, 2, ..., i-1\}$. На основе этой информации агент формирует свое убеждение относительно состояния мира w и выбирает действие $a_i \in [0, 1]$, которое максимизирует его функцию полезности $u_i(a_i, w) = -(a_i - w)^2$, точнее, ожидаемую полезность $\mathbb{E}[-(a_i - w)^2]$ (таким образом, выбранное им действие соответствует его убеждению относительно вероятности события $\{w = 1\}$). Агенты в модели являются байесовскими, однако авторами указанной работы принимается довольно *сильное допущение о наивности участников сети*: i-й агент ошибочно полагает, что действие его предшественника j в сети наблюдений обусловлено исключительно получаемым им частным сигналом (у него нет предшественников), т. е. он считает, что $a_j = P[w = 1|s_j]$, и недооценивает коррелированность действий его предшественников. Интересно, что оптимальное действие агента можно вывести по правилу, сходному с правилом обновления в модели формирования мнений ДеГроота. Доказано, что принятое допущение приводит к тому, что в более плотных сетях наблюдений [24] общество (агенты сети) в долгосрочной перспективе чаще (по сравнению с разреженными сетями) приходит к ошибочной оценке истинного состояния мира (*ошибочному научению*), т. е. либо $\lim_{n\to\infty} a_n = 0$, если $w = 1$, либо $\lim_{n\to\infty} a_n = 1$, если $w = 0$. Такой эффект объясняется тем, что в разреженных сетях «ранние» агенты не сильно влияют друг на друга, следовательно, достигнутый консенсус «включает в себя» больше независимых источников информации и скорее всего будет правильным. Показано также, что в случае континуальных действий агенты почти наверное придут к консенсусу, т. е. разногласия агентов исчезнут. Однако, если действия агентов считать бинарными (для любого i-го агента $a_i \in \{0, 1\}$), а множество агентов разбить на две группы с четными и нечетными номерами так, что в соответствии со стохастической блочной моделью вероятность наличия связи наблюдения от i-го агента к j-му равна q_s, если они находятся в одной группе, и равна q_d, если они находятся в разных группах ($q_s > q_d > 0$), то существует положительная вероятность того, что все нечетные агенты выберут действие 0, в то время как все четные агенты выберут действие 1. Таким образом, в связной сети *сформируются информационные сообщества с противоположными представлениями о состоянии природы.*

В работе [25] агенты являются *локально байесовскими*, т. е. они обрабатывают информацию как байесовские агенты, но при этом каждый из них считает, что его эго-сеть (локальная сеть) и есть вся исходная глобальная сеть агентов, которая является неориентированной и связной (при этом, в отличие от рассмотренных выше моделей, агент не считает, что его соседи руководствуются только частными сигналами). В каждый момент времени агенты формируют свои убеждения относительно состояния мира на основании полученного на предыдущем шаге частного сигнала и сообщений своих соседей (точнее, полной истории сообщений соседей, поскольку агенты обладают совершенной памятью), а затем обмениваются своими убеждениями. Предлагается простое правило обновления убеждений, согласно которому агенты приписывают неожиданные изменения в убеждениях соседей получаемым ими новым частным сигналам (агенты считают, что за пределами их локальной сети не существует других агентов). Показывается, что в некоторых сетях убеждения агентов колеблются, не стабилизируясь.

Ограничения на рациональность агентов также накладываются в моделях с повторяющимися действиями агентов в сети, в которых агенты многократно пересматривают свои убеждения и действия (повторяющееся байесовское обновление): модели повторяющихся действий с локально-оптимальными агентами, с эвристическим включением информации от соседей и рациональными ожиданиями агентов.

Модели повторяющихся действий с локально-оптимальными агентами. В этих моделях агенты в каждый момент времени выбирают наилучший ответ, основываясь на своих текущих убеждениях (которые формируются рационально), не учитывая влияние своих действий на остальных агентов и возможность получения в будущем дополнительной информации. Если действия агентов континуальны и априорные представления агентов одинаковы, то в любой связной сети достигается консенсус как в случае дискретных состояний природы [23], так и в случае нормально распределенных [26]. В случае дискретных действий агентов ($x_n(t) \in \{0, 1\}$) в связной сети также возможно итоговое достижение консенсуса (см. работы [27, 28]). В модели [28] каждый агент выполняет локально-оптимальное действие в каждый момент времени с учетом своих текущих представлений.

Модели повторяющихся действий с эвристическим включением информации от соседей. Яркий пример такого подхода — модель [29, 30], которая отчасти напоминает модель ДеГроота. Агенты имеют априорные представления относительно состояния природы $\theta \in \{0, 1\}$. В начале каждого периода каждый агент получает частный сигнал и наблюдает убеждения своих соседей. Агент n в момент времени t имеет представление $p_n(t) = P(\theta = 1)$, он обновляет свое представление сначала по правилу Байеса с учетом полученного сигнала $s_n(t)$

$$p'_n(t) \equiv P(\theta = 1|s_n(t)) =$$
$$= \frac{P(s_n(t)|\theta = 1)p_n(t)}{P(s_n(t)|\theta = 1)p_n(t) + P(s_n(t)|\theta = 0)(1 - p_n(t))},$$

а затем усредняет его с полученными убеждениями соседей по правилу ДеГроота

$$p_n(t + 1) = a_{nn}p'_n(t) + \sum_m a_{nm}p_m(t),$$

где матрица A задает веса соседей. Если получаемые агентами сигналы неинформативны, то фактически убеждения агентов формируются по правилу ДеГроота (см. § 2). Если сигналы информативны, граф сети является сильно связным, а каждый агент «доверяет» себе, то представления агентов почти наверное сойдутся к правильной оценке состояния природы.

Модели повторяющихся действий с рациональными ожиданиями агентов. В работе [31] множество состояний природы $\Theta = \{0, 1\}$. В начальный момент времени каждый агент n получает информативный сигнал s_n. В каждый момент времени агент n наблюдает действие каждого своего соседа $m \in B(n)$, затем выбирает действие $x_n(t)$ и получает выигрыш

$$u(x_n(t), h_n(t), s_n) = P(\theta = x_n(t)|h_n(t), s_n),$$

где $h_n(t)$ — история действий соседей к началу момента t. Агенты дисконтируют свои будущие выигрыши с коэффициентом $\lambda \in (0, 1)$ и играют в повторяющуюся игру с неполной информацией. Если граф сети является L-локально связным и существует ограничение сверху d на количество наблюдаемых соседей, то все агенты в бесконечной сети (большой сети) почти наверное (с высокой вероятностью) придут к истинной оценке состояния природы. Граф G является *L-локально связным*, если для каждого ребра (n, m) существует путь из m в n длиной не более L. Свойство L-связности и существование ограничения d можно интерпретировать как отсутствие в сети чрезмерно влиятельных агентов.

2. ФОРМИРОВАНИЕ ИНФОРМАЦИОННЫХ СООБЩЕСТВ В СОЦИАЛЬНОЙ СЕТИ В МОДЕЛЯХ С ЭВРИСТИЧЕСКИМИ ПРАВИЛАМИ ОБНОВЛЕНИЯ ПРЕДСТАВЛЕНИЙ АГЕНТОВ

В байесовских моделях изначально не учитываются психологические компоненты личности индивида. В то же время из психологии и социальной психологии известно, что индивиды обладают когнитивными ограничениями и подвержены действию различных социо-психологических факторов (среди которых предрасположенность к своей точке зрения, социальное влияние одних индивидов на других, конформизм и т. п.). Для объяснения возникающих эффектов разрабатываются различные теории и модели. Со второй половины XX века разрабатываются и совершенствуются математические модели, в которых применяются простые эмпирические правила обновления представлений агентов и которые демонстрируют наблюдаемые на практике эффекты. Основополагающие работы в области *моделирования динамики мнений* исследуют и моделируют в первую очередь феномен согласования мнений агентов (консенсус), когда взаимодействие между членами сети приводит к постепенному уменьшению различий между мнениями участников взаимодействий. Этот феномен объясняется в социальной психологии целым рядом причин, в том числе конформизмом, результатом принятия доказательств (убеждением), неполной информированностью, неуверенностью в собственных решениях и т. п.

В классических формальных моделях динамики мнений (см. работы [4, 7, 32—35] рассматривается последовательное усреднение непрерывных (*континуальных*) мнений агентов в дискретном времени. Существуют различные вариации такого рода моделей, в которых усреднение происходит в непрерывном времени [36, 37].

Приведем немного модифицированный пример классической модели ДеГроота, в которой рассматривается динамика формирования консенсуса мнений в сетевой структуре. В этой структуре узлы из множества $N = \{1, ..., n\}$ на каждом шаге формируют свое текущее мнение по такому правилу: мнение агента является взвешенной суммой мнений всех остальных агентов сети, а также своего мнения на предыдущем шаге:

$$x_i^{(t+1)} = \sum_{j \in N} a_{ij} x_j^{(t)}, \quad t \geq 0,$$

где $x_i^{(0)}$ — мнение i-го агента в некий начальный момент времени. Параметр $a_{ij} \in [0, 1]$ отражает степень влияния j-го агента на i-го агента $\left(\sum_j a_{ij} = 1\right)$.

В матричном виде динамику мнений можно записать так:

$$x^{(t+1)} = Ax^{(t)},$$

где A — стохастическая по строкам матрица влияний.

В качестве отступления отметим, что для модели ДеГроота можно привести микроэкономические основания и обозначить связь с байесовскими моделями. В частности, если начальные представления индивидов зашумлены, то правило обновления ДеГроота является оптимальным на первом шаге [17]: новым мнением индивида является взвешенная сумма мнений его соседей, причем весом мнения соседа является точность его информации. В последующие моменты времени индивид должен скорректировать веса своих соседей с учетом того, что поступающая информация может повторяться. Сделать это непросто, поэтому правило ДеГроота с постоянными весами можно рассматривать как поведенческую эвристику. Другое микроэкономическое основание — представление агентов как игроков, стремящихся найти решение простой координационной игры [38], в этой игре динамика локально оптимального наилучшего ответа (совпадающая с динамикой по ДеГрооту) ведет к равновесию по Нэшу.

Динамика мнений по ДеГрооту приводит, в частности, к тому, что в социальной сети, имеющей сильную связность, достигается консенсус. Мнения агентов начинают совпадать, поскольку каждый агент оказывает прямое или косвенное влияние на любого другого агента сети и различия во мнениях агентов нивелируются.

Структура сети взаимодействий накладывает ограничения на возможность достижения консенсуса. Очевидно, например, что в несвязной сети консенсус может быть достигнут лишь в особых случаях. Различия во мнениях агентов могут наблюдаться и в сильно связных сетях, если, например, агенты имеют в какой-то степени «нечувствительные» к влиянию предубеждения [39]. В такого рода моделях мнение агента на каждом шаге формируется как взвешенная сумма мнений на предыдущем шаге и своего начального мнения:

$$x^{(t+1)} = \Lambda A x^{(t)} + [I_n - \Lambda]x^0,$$

где $\Lambda = I_n - \mathrm{diag}(A)$.

Содержательно начальные мнения агентов можно интерпретировать как индивидуальные предпочтения или укоренившиеся убеждения, влияние которых сохраняется в процессе обмена мнениями.

Сходную с рассмотренной динамику мнений можно получить при помощи модели формирования мнений со сложными узлами [40], в которой каждый узел состоит из двух взаимодействующих между собой агентов — внешнего и внутреннего. Информационный обмен узла с другими узлами сети осуществляет внешний агент, внутренний агент (который интерпретируется как доверенное лицо внешнего — друг или консультант) взаимодействует только с соответствующим внешним.

Многомерным обобщением модели с «нечувствительными» агентами является модель [41], в которой рассматривается не один, а сразу несколько взаимосвязанных вопросов (m различных тем), по каждому из которых у каждого из агентов имеется свое мнение. Мнение i-го агента ($i \in N$) по m различным темам задается вектором $x_i^{(t)} = (x_i^{(t)}(1), ..., x_i^{(t)}(m))$. Динамика мнений i-го агента в момент t задается так:

$$x_i^{(t)} = \lambda_{ii} \sum_{j \in N} a_{ij} y_j^{(t-1)} + (1 - \lambda_{ii}) x_i^{(0)},$$

$$y_j^{(t-1)} = C x_j^{(t-1)},$$

где C — матрица взаимовлияния обсуждаемых тем, а $y_j^{(t-1)}$ — выпуклые комбинации j-го агента по нескольким темам. В матричном виде динамику можно представить так:

$$x^{(t)} = [(\Lambda A) \otimes C] x^{(t-1)} + [(I_n - \Lambda) \otimes I_m] x^{(0)},$$

где $\otimes$ — произведение Кронекера, $\Lambda = I_n$ или $\Lambda = I_n - \mathrm{diag} A$ (в зависимости от модели).

В целом отметим, что в рассмотренных выше моделях, несмотря на учет дополнительных факторов (наличие предубеждений и наличие взаимовлияющих тем), сохраняющих определенное рассогласование мнений, взаимовлияние между агентами приводит со временем к уменьшению разногласий. В частности, предположение об усреднении подразумевает, что мнения никогда не выйдут за пределы диапазона начальных мнений.

Для моделей динамики формирования мнений получены многочисленные теоретические результаты, как правило, связанные с достижением консенсуса в сети. Для исследования такого рода моделей применяется аппарат теории стохастических матриц, теории однородных и неоднородных марковских цепей. Известно, что динамику мнений можно моделировать при помощи цепей Маркова. В однородной марковской цепи достижение консенсуса определяется сходимостью степеней стохастической матрицы. В работах [4, 42] приведены некоторые достаточные условия сходимости степеней стохастической матрицы. Для класса стохастических матриц, не гарантирующих консенсус, в работе [42] приведены необходимые условия достижения консенсуса, а в работе [43] найдены минимальные изменения начальных представлений агентов, приводящие к консенсусу. Кроме того, известны более частные результаты. Например, в работе [38] для случайных сетей, являющихся по сути стохастично блочными, определена зависимость скорости сходимости представлений агентов (обновляемых по правилу простого усреднения) от значения гомофилии.

Выше шла речь о долгосрочном консенсусе, на практике часто интересно знать возможность достижения разногласий за конечное время. Как влияет структура сети на *среднесрочные разногласия и формирование «среднесрочных» информационных сообществ*? Если начальные представления агентов одинаковы, то консенсус достигается сразу же и не зависит от структуры сети. В общем случае консенсус зависит от начальных представлений и структуры сети. Можно рассматривать наихудший случай [44]: анализировать скорость сходимости представлений (количество шагов, необходимое для того, чтобы различия в представлениях были гарантированно малы) для любых начальных представлений. Для простоты рассматриваются неразложимые и примитивные матрицы влияния A. Для типичной такой матрицы справедливо (почти всюду)

$$A^t = \sum_{l=1}^{n} \lambda_l^t P_l,$$

для нее выполняются свойства:

— $\lambda_1 = 1, \lambda_2, \lambda_3, ..., \lambda_n$ — различные собственные значения A, упорядоченные по невозрастанию модуля,

— матрица P_l — проекционный оператор, который соответствует нетривиальному одномерному собственному подпространству, отвечающему собственному значению λ_l,

— $P_1 = A^\infty$ и $P_1 x^{(0)} = x^{(\infty)}$,

— $P_l \mathbf{1} = \mathbf{0}$ для всех $l > 0$.

Матрица P_1 соответствует матрице результирующих влияний A^∞ и определяет устойчивое состояние системы (результирующие представления агентов), остальные матрицы $P_{l>1}$ отражают отклонение от этой результирующей матрицы в момент времени t. То, как быстро матрица P_1 будет доминировать, зависит от величины λ_2: чем меньше значение, тем быстрее возникнет устойчивое состояние в сети [45]. Оказывается, что для представлений агентов в момент t выполняется

$$\frac{1}{2}|\lambda_2|^t - (n-2)|\lambda_3|^t \le \sup_{x^{(0)} \in [0,1]^n} \|x^{(t)} - x^{(\infty)}\|_\infty \le$$
$$\le (n-1)|\lambda_2|^t.$$

Таким образом, величина $|\lambda_2|^t$ определяет максимальное отклонение представления агента в сети от результирующего представления. А матрица P_2 главным образом определяет отклонение представлений от результирующего представления (консенсуса) и соответствует *метастабильному состоянию сети, когда большая часть разногласий исчезает, но сохраняется некоторая устойчивая часть разногласий.*

Поскольку матрица $P_2 = \sigma\rho^{\mathrm{T}}$, где ρ^{T} — левый собственный вектор матрицы A, соответствующий собственному значению λ_2, а σ — правый собственный вектор матрицы A, соответствующий собственному значению λ_2, то при достаточно больших значениях t отклонение $x^{(t)} - x^{(\infty)}$ в сущности равно $\lambda_2^t\,\sigma(\rho^{\mathrm{T}}x^{(0)})$. Таким образом, если собственное значение λ_2 — положительное вещественное число, то отклонение i-го агента от консенсуса пропорционально компоненте σ_i безотносительно начальных представлений $x^{(0)}$. Упорядочение среднесрочных представлений агентов определяется одним зависящим от сетевой структуры числом. В работе [17] приводится интересная интерпретация: мнения индивидов по множеству различных вопросов могут быть хорошо аппроксимированы линией, позиция индивида на этой линии (в лево-правом спектре) и определяет его позицию по всем вопросам (например, мнения многих людей по широкому кругу принципиально не связанных вопросов могут характеризоваться мерой их консервативности/либеральности).

Возникает вопрос о том, *как структура сети (матрица влияний) влияет на сохранение различных информационных сообществ в сети*. В частности, можно оценить влияние расщепления (разбиения) сети на группы, например, при помощи показателя «узкого места» графа или константы Чигера [45]:

$$\Phi_*(A) = \min_{S:S\subseteq N,\ \sum_{i\in S}\pi_i \le \frac{1}{2}} \frac{\sum_{i\in S,\, j\notin S}\pi_i A_{ij}}{\sum_{i\in S}\pi_i},$$

где π — левый собственный вектор матрицы A, соответствующий собственному значению λ_1 (вектор влияний агентов). Эта константа мала, если существует некоторое множество агентов, имеющее не более половины влияния в обществе, которое подвергается относительно небольшому внешнему влиянию. Если матрица влияния соответствует обратимой марковской цепи, то константа ограничивает значения λ_2 [45]: $\Phi_*^2(A) \le 1 - \lambda_2 \le 2\Phi_*(A)$, т. е. чем меньше константа, тем больше значение λ_2, и это влияет на убывание разногласий (тем дольше они сохраняются в сети).

Выше предполагается, что матрица влияния (доверия) в модели динамики представлений не изменяется во времени. Существуют исследования, в которых матрица влияний меняется со временем. Модель динамики представлений агентов была обобщена в работе [46], где матрица влияния меняется на каждом шаге и итеративный процесс задается произведением матриц:

$$x^{(t)} = A^{(t)}x^{(t-1)}.$$

Решение задачи согласования мнений в такой постановке сводится к исследованию сходимости неоднородных цепей Маркова. Отметим также, что рассмотренный класс моделей формирования мнений имеет тесную связь с исследованиями, посвященными консенсусу в многоагентных системах, которых на настоящий момент огромное число (см. обзоры [47, 48]). Полученные в этой области теоретические результаты могут быть перенесены и на область социальных сетей (в случае, конечно, простых правил обновления представлений, которые не слишком хорошо учитывают специфику предметной области).

Вопросам формирования консенсуса в работах, являющихся развитием классической модели ДеГроота, уделяется много внимания, вопросам о возможности возникновения в этом случае социального научения (формированию общества с истинными представлениями) — существенно меньше.

Формирование истинных представлений в модели ДеГроота. Если представления агентов в сети отражают их оценки значения параметра внешней среды $\mu \in [0, 1]$, то агенты могут прийти как к истинному консенсусу, так и ложному. Поэтому возникает вопрос о возможности *социального научения* в сети с агентами, представления которых обновляются по правилу ДеГроота. В работе [18] начальные представления (частные сигналы) $x_i^{(0)}$ являются независимыми случайными величинами, которые имеют математическое ожидание μ, положительную дисперсию и значения которых лежат на отрезке [0, 1]. Известно, что при достаточно слабых условиях представления агентов сходятся к одному и тому же итоговому мнению. Для того, чтобы оценить сходимость к истинному значению результирующей оценки в обществе по мере его роста, рассматривается последовательность матриц влияния $(A(n))_{n=1}^{\infty}$, индексируемая числом агентов в каждой сети n. Считается, что каждая сеть (соответствующая матрица влияния) является *сходящейся*, т. е. относительно нее любые начальные представления агентов сходятся к некоторому пределу. Последовательность $(A(n))_{n=1}^{\infty}$ называется *мудрой*, если $\operatorname{p}\lim\limits_{n\to\infty} \max\limits_{i\le n} |x_i^{(\infty)}(n) - \mu| = 0$.

Одно из условий «мудрости» общества связано с влиянием агентов. Без ограничения общности агенты упорядочиваются по убыванию влияния так, что $s_i(n) \ge s_{i+1}(n) \ge 0$ для любого i и n, где $s_i(n)$ — вес, с которым начальное мнение i-го агента входит в итоговое мнение сети n. Утверждается, что последовательность сходящихся стохастических матриц $(A(n))_{n=1}^{\infty}$ является мудрой, если и только если соответствующие векторы влияний таковы, что $s_1(n) \to 0$, т. е. влияние наиболее влиятельного агента в обществе стремится к нулю по мере роста общества.

Препятствием к мудрости также являются *значимые* (prominent) группы, которые получают непропорциональную долю внимания общества и сбивают его с правильного пути. Для фиксированной сети с числом агентов n *группой* B называется подмножество множества агентов $\{1, ..., n\}$. Группа B является *t-значимой* относительно A (или значимой в пределах t шагов), если любой агент $i \notin B$ подвергается косвенному влиянию группы, т. е. $A_{i,B}^{t} > 0$. Минимальный вес такого влияния называется *p-шаговой значимостью* B: $\pi_B(A; t) = \min_{i \notin B} A_{i,B}^{t}$. *Семейством* называется последовательность групп (B_n), где $B_n \subset \{1, ..., n\}$ для каждого n. Семейство (B_n) *равномерно значимо* относительно $(A(n))_{n=1}^{\infty}$, если существует такая константа $a > 0$, что для каждого n существует такое t, что $\pi_{B_n}(A(n); t) \ge a$. Семейство (B_n) является *конечным*, если оно в итоге перестает расти, т. е. если существует такое q, что $\sup_n |B_n| \le q$. Утверждается, что если существует конечное равномерно значимое семейство по отношению к $(A(n))$, то последовательность не является мудрой. Таким образом, большие общества с «небольшой» группой, которая влияет на каждого в сети, не способны достичь правильных представлений относительно параметра внешней среды.

В рассмотренных выше моделях динамики мнений моделируются феномен консенсуса, когда различия во мнениях взаимодействующих агентов уменьшаются со временем, и сопутствующий феномен социального научения. При этом, как показано, возможен эффект среднесрочных разногласий, обусловленный структурой сети. Во многих случаях в социальных сетях наблюдаются социально-психологические феномены (см. например, статью [49]), связанные с долгосрочными разногласиями и стабильными информационными сообществами: сохранение различий, групповая поляризация (в групповой дискуссии усиливается любая изначально доминирующая точка зрения), поляризация мнений (когда происходит усиление разногласий между двумя оппозиционными группами) и т. п. Классические модели динамики мнений в долгосрочной перспективе плохо объясняют сохранение различий во мнениях или даже усиление радикальных мнений в сильно связных сетевых структурах. Для их моделирования разрабатываются новые формальные математические модели (см. работы [5, 6, 37, 50—53]). В частности, формирование множеств агентов с разными представлениями моделируется в моделях ограниченного доверия (bounded confidence) [5, 6, 51], в которых только достаточно близкие агенты могут оказывать влияние друг на друга (это правило взаимодействия обычно мотивируется явлениями гомофилии и социальной идентификации).

В модели [54] мнения агентов задаются вектором $x = (x_1, ..., x_n) \in R^n$. Агент i воспринимает мнения остальных агентов, только если они достаточно близки к его собственному мнению, т. е. множество влияющих на него агентов $I(i, x) = \{1 \le j \le n \mid |x_i - x_j| \le \epsilon_i\}$, где $\epsilon_i > 0$ — уровень уверенности (обычно $\epsilon_i = \epsilon$). Тогда динамика мнений такова:

$$x_i(t+1) = |I(i, x(t))|^{-1} \sum_{j \in I(i, x(t))} x_j(t).$$

Это динамика мнений по ДеГрооту с зависящей от мнений агентов матрицей доверия: $a_{ij}(x) = 1/|I(i, x)|$, если $j \in I(i, x)$, иначе $a_{ij}(x) = 0$. Приведены необходимые и достаточные условия достижения консенсуса. В частности, показано, что для достижения консенсуса необходимо, чтобы в любой момент времени вектор мнений был ϵ-профилем (вектор является ϵ-профилем, если в упорядочении его компонент по возрастанию значений расстояние между двумя соседними элементами не больше ϵ). В противном случае консенсус невозможен, *сеть доверия расщепляется на компоненты связности, в ней формируются группы индивидов с одинаковыми убеждениями* (*информационные сообщества*). В любом случае агенты придут к равновесию за конечное число шагов [55].

В работе [56] рассмотрены «осторожные» агенты, доверие которых к получаемым сообщениям зависит от содержания сообщений (высказанных в них мнений). Для того чтобы отразить эту зависимость, вводится функция доверия $G(x, u)$, где x — мнение агента, u — получаемое агентом сообщение. Относительно свойств функции доверия вводятся различные предположения, комбинации которых приводят к различным случаям формализации функции доверия, соответствующим: агенту, который независимо от содержания доверяет получаемым сообщениям; агенту-консерватору; агенту-новатору; агенту — умеренному консерватору; и агенту — умеренному новатору. Приведем пример функции доверия агента — умеренного консерватора:

$$G(x, u) = \beta[1 - (1 - \exp(-\gamma|x - u|))\exp(-\gamma|x - u|)],$$

где β и γ — константы. Содержательно агент выделяет и воспринимает информацию, совпадающую с его мнением, до тех пор, пока различие во мнениях не станет достаточно велико, но при очень больших отклонениях вероятность того, что агент заметит такую информацию, растет. В общем случае управляемая динамика мнений агентов в социальной сети описывается выражением:

$$x_i^k = a_{ii} x_i^{k-1} + \beta G_i(x_i^{k-1}, u^{k-1}) u^{k-1} + \\ + \sum_{j \in N_i} a_{ij} G_i(x_i^{k-1}, x_j^{k-1}) x_j^{k-1}, \quad k = 1, 2, ...,$$

где u — уже управление со стороны внешнего субъекта (например, СМИ), а индивидуальные функции доверия $\{G_i(\cdot)\}_{i \in N}$ таковы, что выполняется условие нормировки. В рамках такой модели можно условно считать, что матрица A отражает доверие агентов *источникам информации*, а функции доверия отражают доверие агентов *содержанию информации*. В частном случае однородной и регулярной сети рассматривается задача синтеза оптимального информационного управления, которая формулируется как задача поиска последовательности управлений, максимизирующей критерий эффективности, и может быть решена известными методами. Эти соображения о зависимости степени доверия агента к сообщениям других агентов от их содержания привели к разработке модели [8], в которой моделируются два взаимосвязанных процесса: процесс распространения действий в сети и процесс формирования мнений агентов. При помощи численного моделирования показано, что в сети в результате обмена представлениями формируются различные информационные сообщества.

Выше рассматривались модели динамики представлений «эвристических» индивидов в социальной сети, в которых представления являются континуальными. В целом такого рода модели — в которых мнения меняются постепенно — представляются наиболее естественными. Это подтверждается работами в области социальной психологии и поведенческой экономики, в частности, в работе [57] отмечается, что общество (на примере индийских деревень) делится на агентов двух типов: байесовских и действующих по правилу ДеГроота. Тем не менее, существует большое количество работ, в которых рассматриваемые мнения имеют порядковые или даже номинальные шкалы и т. д. К моделям *с дискретными представлениями/мнениями агентов* относят модели «голосования» (voter model) [58], модели большинства и пороговые модели [59]. Многие из этих моделей, учитывающих сетевую структуру взаимодействий, известны также как модели распространения активности/информации в сети (см., например, статью [60]). Приведем примеры моделей голосования с дискретными представлениями индивидов, которые демонстрируют эффект формирования разногласий в сети.

В работе [61] рассматривается множество из N агентов на $L \times L$ регулярной решетке ($L^2 = N$). Агент $i \in \{1, 2, ..., N\}$ выбирает действие $a^{(i)} \in A = \{a_1, ..., a_g, ..., a_G\}$, руководствуясь своим мнением, описываемым правилом $r^{(i)} \in R = \{r_1, ..., r_k, ..., r_K\}$. Особенность модели состоит в возможности задания множественных отношений между мнениями-правилами и действиями: правила бывают эксклюзивные (обязательно одно действие, остальные запрещены) и инклюзивные (разрешают с равной вероятностью выполнение одного из нескольких действий). Агенты знают множества A и R и матрицу отношений S (размерности $K \times G$), также они могут наблюдать действия соседей в сети, но не их мнения. В начальный момент времени каждый i-й агент случайно (с равной вероятностью) выбирает правило $r^{(i)} \in R$ и действует согласно этому правилу. В каждый следующий момент времени τ случайно выбранный агент i обновляет свои представления о правиле, которому следует его сосед $j \in M_i$, основываясь на наблюдаемом им действии $a^{(j)}(\tau)$:

$$P^{(i)}(r^{(j)}(\tau) = r_k | a^{(j)}(\tau)) = \frac{P^{(i)}(a^{(j)}(\tau) | r_k)}{\sum_{k=1}^{K} P^{(i)}(a^{(j)}(\tau) | r_k)}.$$

Затем (в этот же момент времени) i-й агент принимает правило выбора действий в соответствии с вероятностью его использования соседями в сети.

Показано, что *для существования различных информационных сообществ в сети достаточно возможности применения агентами инклюзивных правил.* В случае применения агентами только эксклюзивных правил модель сводится к классической модели голосования.

Более содержательно богатая модель формирования представлений, основывающаяся на модели голосования, предлагается в работе [62]. В ней отмечается, что в традиционных экономических моделях социального научения редко возможны долгосрочные разногласия, хотя в реальности разногласия возникают чаще, чем консенсус. Во многих экономических моделях социального научения подразумевается, что каждая новая порция информации, получаемая агентом, является истинной: агент наблюдает элемент разбиения пространства состояний, и этот элемент разбиения содержит истинное состояние. Это предположение (в ряде случаев неявное) приводит агентов к консенсусу. В работе [62] предлагается концепция обработки информации агентами, в которой агенты могут получить ложную информацию, которая приводит к тому, что разногласия в моделях на основе этой концепции являются обычным исходом в сети. Вводится два класса аксиом для используемых агентами правил обновления убеждений: аксиомы готовности к научению (правило обновления позволяет агенту научиться и прийти к истинной оценке состояния мира) и аксиомы неманипулируемости (правило обновления приводит агента к одному и тому же убеждению при получении им одной и той же информации независимо от ее представления). В зависимости от выбранных комбинаций аксиом возможны те или иные правила обновления. В работе детально рассматриваются два из них:

— агенты, придерживающиеся правила «неуступчивого» обновления, никогда не отказываются от своих убеждений;

— агенты придерживаются правила неуступчивого обновления, но могут полностью поменять свои убеждения, если новая информация полностью разрешает всю неопределенность относительно состояния мира.

Протокол взаимодействия агентов следующий: в каждый дискретный момент времени выбирается случайно одно из ребер графа социальной сети ij; агент i сообщает выбираемое случайно одно из высказываний (входящих в его убеждение B_j) P агенту j; затем j-й агент обновляет свое убеждение согласно правилу $U_j(B_j, P)$, где высказывание P — это подмножество значений состояния мира Ω, т. е. $P \subseteq \Omega$. Агенты сети достигают консенсуса, если в некоторый момент существует такое высказывание P^*, что для каждого i-го агента $P(B_i) = P^*$, где $P(B_i)$ — пересечение высказываний из убеждений i-го агента. Оказывается, что для достижения консенсуса в сети нужно принять довольно сильные допущения относительно истинности начальных убеждений и числа неуступчивых агентов. Показано, что *в больших сетях формирование различных информационных сообществ не только возможно, но и неизбежно.*

ЗАКЛЮЧЕНИЕ

Во второй части обзора рассмотрены вопросы формирования информационных сообществ в обществах с нетривиальной сетевой структурой. Индивиды — члены социума взаимодействуют между собой в рамках этой структуры, и индивид (агент), наблюдая действия соседей в сети, может получить дополнительную информацию относительно интересующего его вопроса.

Рациональные агенты в таких сетях приходят в долгосрочной перспективе к согласию (истинному или ложному — зависит от условий, накладываемых на топологию и/или начальные представления). Для получения различных информационных сообществ необходимо ослабить требование рациональности агентов. В первом классе рассматриваемых моделей агенты являются условно байесовскими, например, они обладают «наивными» представлениями о составе и структуре сети или наивно учитывают сигналы соседей. Во втором классе моделей агенты являются «простыми», формирующими свои представления на основе эвристических правил. В случае индивидов, безусловно доверяющих действиям своих соседей в сети, согласие в социальной сети является обычным исходом взаимодействий — в нем формируется одно информационное сообщество, однако при этом можно указать условия формирования различных среднесрочных (метастабильных) информационных сообществ. Показано, что отличающиеся друг от друга устойчивые информационные сообщества могут быть получены в том случае, когда на индивидов помимо социального влияния действуют и другие социально-психологические факторы (гомофилия, склонность к подтверждению своей точки зрения и т. д.), а также когда в сети наряду с истинной распространяется и ложная информация.

В третьей части обзора будут рассмотрены эмпирические исследования, связанные с вопросами существования информационных сообществ в реальных социальных сетях и их особенностями.

ЛИТЕРАТУРА

1. *Губанов Д.А., Петров И.В.* Информационные сообщества в социальных сетевых структурах. Ч. 1. От основного понятия к математическим моделям формирования // Проблемы управления. — 2021. — № 1. — С. 15—23. [*Gubanov, D.A., Petrov, I.V.* Information Communities in Social Network Structures. Part 1. From Concept to Mathematical Models // Control Sciences. — 2021. — No. 1. — P. 15—23 (In Russian)].
2. *Kadushin, C.* Understanding Social Networks: Theories, Concepts, and Findings. — New York: Oxford University Press, 2012. — 252 p.
3. *Myers, D.G.* Social Psychology. — N.-Y.: McGraw-Hill, 1999. — 760 p.
4. *DeGroot, M.H.* Reaching a Consensus // Journal of American Statistical Assotiation. — 1974. — No. 69. — P. 118—121.
5. *Deffuant, G., Neau, D., Amblard, F., Weisbuch, G.* Mixing Beliefs among Interacting Agents // Advances in Complex Systems. — 2000. — Vol. 03. — P. 87—98.
6. *Hegselmann, R., Krause, U.* Opinion Dynamics under the Influence of Radical Groups, Charismatic Leaders, and Other Constant Signals: A Simple Unifying Model // Networks and Heterogeneous Media. — 2015. — Vol. 10, no. 3. — P. 477—509.

7. *Губанов Д.А., Новиков Д.А., Чхартишвили А.Г.* Социальные сети: модели информационного влияния, управления и противоборства. 3-е изд., перераб. и дополн. — М.: МЦНМО, 2018. — 224 с. [*Gubanov, D.A., Novikov, D.A., Chkhartishvili, A.G.* Social Networks: Models of Information Influence, Control and Confrontation. — Cham, Switzerland: Springer International Publishing, 2019. — 158 p.]
8. *Губанов Д.А., Петров И.В., Чхартишвили А.Г.* Многомерная модель динамики мнений в социальных сетях: индексы поляризации // Проблемы управления. — 2020. — № 3. — С. 26—33 [*Gubanov, D.A., Petrov, I.V., Chkhartishvili, A.G.* Multidimensional Model of Opinion Dynamics in Social Networks: Polarization Indices // Control Sciences. — 2020. — No. 3. — P. 26—33 (In Russian)].
9. *Rabin, M., Schrag, J.L.* First Impressions Matter: A Model of Confirmatory Bias // The Quarterly Journal of Economics. — 1999. — No. 1 (114). — P. 37—82.
10. *Jern, A., Chang, K.K., Kemp, C.* Belief Polarization is not Always Irrational // Psychological Review. — 2014. — No. 2 (121). — P. 206—224.
11. *Acemoglu, D., Ozdaglar, A.* Opinion Dynamics and Learning in Social Networks // Dynamic Games and Applications. — 2011. — No. 1 (1). — P. 3—49.
12. *Vives, X.* How Fast Do Rational Agents Learn? // Review of Economic Studies. — 1993. — Vol. 60. — P. 329—347.
13. *Burguet, R., Vives, X.* Social Learning and Costly Information Acquisition // Economic Theory. — 2000. — No. 1 (15). — P. 185—205.
14. *Bikhchandani, S., Hirshleifer, D., Welch, I.* A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades // The Journal of Political Economy. — 1992. — Vol. 100, no. 5. — P. 992—1026.
15. *Ali, S.N.* Herding with Costly Information // Journal of Economic Theory. — 2018. — Vol. 175. — P. 713—729.
16. *Acemoglu, D., Dahleh, M., Lobel, I., Ozdaglar, A.* Bayesian Learning in Social Networks // The Review of Economic Studies. — 2011. — No. 4 (78). — P. 1201—1236.
17. *DeMarzo, P.M., Vayanos, D., Zwiebel, J.* Persuasion Bias, Social Influence, and Unidimensional Opinions // The Quarterly Journal of Economics. — 2003. — Vol. 118. — P. 909—968.
18. *Golub, B., Jackson, M.O.* Naïve Learning in Social Networks and the Wisdom of Crowds // American Economic Journal: Microeconomics. — 2010. — No. 1 (2). — P. 112—149.
19. *Smith, L., Sorensen, P.* Pathological Outcomes of Observational Learning // Econometrica. — 2000. — Vol. 68, no. 2. — P. 371—398.
20. *Acemoglu, D., Bimpikis, K., Ozdaglar, A.* Dynamics of Information Exchange in Endogenous Social Networks: Endogenous Social Networks // Theoretical Economics. — 2014. — No. 1 (9). — С. 41—97.
21. *Федянин Д.Н., Чхартишвили А.Г.* Об одной модели информационного управления в социальных сетях // Управление большими системами. — 2010. — Вып. 31. — С. 265—275. [*Fedyanin, D.N., Chkhartishvili, A.G.* On a Problem of Information Control in Social Networks // Large-Scale Systems Control. — 2010. — Iss. 31. — P. 265—275. (In Russian)]
22. *Eyster, E., Rabin, M.* Naïve Herding in Rich-Information Settings // American Economic Journal: Microeconomics. — 2010. — No. 4 (2). — P. 221—243.
23. *Mueller-Frank, M.* A General Framework for Rational Learning in Social Networks: Framework for Rational Learning // Theoretical Economics. — 2013. — No. 1 (8). — P. 1—40.
24. *Dasaratha, K., He, K.* Network Structure and Naive Sequential Learning // arXiv:1703.02105 [cs, econ, q-fin]. — 2019. — 40 p.
25. *Li, W., Tan, X.* Locally Bayesian Learning in Networks // Theoretical Economics. — 2020. — No. 1 (15). — P. 239—278.
26. *Mossel, E., Olsman, N., Tamuz, O.* Efficient Bayesian Learning in Social Networks with Gaussian Estimators // 2016 54th Annual Allerton Conference on Communication, Control, and Computing. — Allerton, 2016. — P. 425—432.
27. *Bala, V., Goyal, S.* Learning from Neighbours // The Review of Economic Studies. — 1998. — Vol. 65, no. 3. — P. 595—621.
28. *Gale, D., Kariv, S.* Bayesian Learning in Social Networks // Games and Economic Behavior. — 2003. — No. 2 (45). — P. 329—346.
29. *Jadbabaie, A., Molavi, P., Sandroni, A., Tahbaz-Salehi, A.* Non-Bayesian Social Learning // Games and Economic Behavior. — 2012. — No. 1 (76). — P. 210—225.
30. *Jadbabaie, A., Molavi, P., Tahbaz-Salehi, A.* Information Heterogeneity and the Speed of Learning in Social Networks * / A. Jadbabaie, P. Molavi, A. Tahbaz-Salehi. — 2013. — 68 p.
31. *Mossel, E., Sly, A., Tamuz, O.* Strategic Learning and the Topology of Social Networks: Strategic Learning // Econometrica. — 2015. — No. 5 (83). — P. 1755—1794.
32. *Робертс Ф.С.* Дискретные математические модели с приложениями к социальным, биологическим и экологическим задачам. — М.: Наука, 1986. [*Roberts, F.S.* Discrete Mathematical Models with Applications to Social, Biological, and Environmental Problems. — London: Pearson, 1976. — 576 p.]
33. *French, J.R.* A Formal Theory of Social Power // The Psychological Review. — 1956. — No. 63. — P. 181—194.
34. *Harary, F.* A Criterion for Unanimity in French's Theory of Social Power / Studies in Social Power. — Michigan: Institute of Sociological Research. — 1959. — P. 168—182.
35. *Jackson, M.* Social and Economic Networks. — Princeton: Princeton University Press, 2008. — 648 p.
36. *Abelson, R.P.* Mathematical Models of the Distribution of Attitudes under Controversy // Contributions to Mathematical Psychology. — 1964. — P. 141—160.
37. *Altafini, C.* Consensus Problems on Networks with Antagonistic Interactions // IEEE Transactions on Automatic Control. — 2013. — Vol. 58, no. 4. — P. 935—946.
38. *Golub, B., Jackson, M.O.* How Homophily Affects the Speed of Learning and Best-Response Dynamics // The Quarterly Journal of Economics. — 2012. — Vol. 127. — P. 1287—1338.
39. *Friedkin, N.E., Johnsen, E.C.* Social Influence Network Theory: A Sociological Examination of Small Group Dynamics. — Cambridge: Cambridge University Press, 2011.
40. *Федянин Д.Н., Чхартишвили А.Г.* Консенсус в социальной сети со сложными узлами // Управление большими системами. — 2016. — № 64. — С. 137—150. [*Fedyanin, D.N., Chkhartishvili, A.G.* Consensus in Social Networks with Complex Structure of Social Actors // Large-Scale Systems Control. — 2016. — No. 64. — S. 137—150 (In Russian)]
41. *Parsegov, S.E., Proskurnikov, A.V., Tempo, R., Friedkin, N.E.* Novel Multidimensional Models of Opinion Dynamics in Social Networks // IEEE Transactions on Automatic Control. — 2017. — Vol. 62, no. 5. — P. 2270—2285.
42. *Berger, R.L.* A Necessary and Sufficient Conditions for Reaching a Consensus Using De Groot's Method // Journal of American Statistical Assotiation. — 1981. — Vol. 76. — P. 415—419.
43. *Agaev, R.P., Chebotarev, P.Y.* The Projection Method for Reaching Consensus and the Regularized Power Limit of a Stochastic Matrix // Automation and Remote Control. — 2011. — Vol. 72, iss. 12. — P. 2458—2476.
44. *Golub, B., Sadler, E.* Learning in social networks. — 2017. http://dx.doi.org/10.2139/ssrn.2919146.
45. *Levin, D.A., Peres, Y., Wilmer, E.L.* Markov Chains and Mixing Times. — Providence, RI: American Mathematical Society, 2009.
46. *Chatterjee, S., Seneta, E.* Toward Consensus: Some Convergence Theorems on Repeated Averaging // Journal of Applied Probability. — 1977. — Vol. 14. — P. 159—164.
47. *Агаев Р.П., Чеботарев П.Ю.* Сходимость и устойчивость в задачах согласования характеристик (обзор базовых результатов) // Управление большими системами. — 2010. — № 30-1. — С. 470—505. [*Agaev, R.P., Chebotarev, P.Yu.* Convergence and Stability in Consensus and Coordination Prob-

lems (A Survey of Basic Results) // Large-Scale Systems Control. — 2010. — No. 30-1. — P. 470—505. (In Russian)]
48. *Cao, Y., Yu, W., Ren, W., Chen, G.* An Overview of Recent Progress in the Study of Distributed Multi-Agent Coordination // IEEE Transactions on Industrial Informatics. — Vol. 9, iss. 1. — 2013. — P. 427—438.
49. *Moscovici, S., Zavalloni, M.* The Group as a Polarizer of Attitudes // Journal of Personality and Social Psychology. — 1969. — Vol. 12, no. 2. — P. 125—135.
50. *Del Vicario, M., Scala, A., Caldarelli, G.*, et al. Modeling Confirmation Bias and Polarization // Scientific Reports. — 2017. — Vol. 7 art. — No. 40391. — 9 p.
51. *Hegselman, R., Krause, U.* Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation // Journal of Artificial Societies and Social Simulation. — 2002. — Vol. 5, no. 3.
52. *Jager, W., Amblard, F.* Uniformity, Bipolarization and Pluriformity Captured as Generic Stylized Behavior with an Agent-Based Simulation Model of Attitude Change // Computational & Mathematical Organization Theory. — 2005. — Vol. 10, no. 4. — P. 295—303.
53. *Urbig, D., Lorenz, J., Herzberg, H.* Opinion Dynamics: The Effect of the Number of Peers Met at Once // Journal of Artificial Societies and Social Simulation. — 2008. — Vol. 11, no. 2. — 4 p.
54. *Krause, U.* A Discrete Nonlinear and Non-Autonomous Model of Consensus Formation // Communications in Difference Equations. — 2000. — P. 227—236.
55. *Dittmer, J.* Consensus Formation under Bounded Confidence // Nonlinear Analysis. — 2001. — 47 (7). — P. 4615—4621.
56. *Губанов Д.А., Новиков Д.А.* Модели унифицированного информационного управления в однородных социальных сетях // Управление большими системами. — 2010. — № 30.1. — С. 722—742. [*Gubanov, D.A., Novikov, D.A.* Models of Unified Information Control in Homogeneous Social Networks // Large-Scale Systems Control. — 2010. — No. 30.1. — P. 722—742 (In Russian)]
57. *Chandrasekhar, A.G., Larreguy, H., Xandri, J.P.* Testing Models of Social Learning on Networks: Evidence From Two Experiments // Econometrica. — 2020. — No. 1 (88). — P. 1—32.
58. *Holley, R.A., Liggett, T.M.* Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model // The Annals of Probability. — 1975. — Vol. 3, no. 4. — P. 643—663.
59. *Бреер В.В., Новиков Д.А., Рогаткин А.Д.* Управление толпой: математические модели порогового коллективного поведения. — М.: ЛЕНАНД. — 2016. — 168 с. [*Breer, V.V., Novikov, D.A., Rogatkin, A.D.* Upravlenie tolpoi: matematicheskie modeli porogovogo kollektivnogo povedeniya. — M.: LENAND. — 2016. — 168 s. (In Russian)]
60. *Kempe, D., Kleinberg, J., Tardos, E.* Maximizing the Spread of Influence through a Social Network // Theory of Computing. — 2015. — Vol. 11, no. 4. — P. 105—147.
61. *Tang, T., Chorus, C.* Learning Opinions by Observing Actions: Simulation of Opinion Dynamics Using an Action-Opinion Inference Model // Journal of Artificial Societies and Social Simulation. — 2019. — No. 3 (22). — 20. p.
62. *Sadler, E.* False Information and Disagreement in Social Networks // SSRN Electronic Journal. — 2017. — 28 p.

Статья представлена к публикации членом редколлегии чл.-корр. РАН Д.А. Новиковым.

Поступила в редакцию 26.08.2020, после доработки 09.11.2020.
Принята к публикации 24.11.2020.

Губанов Дмитрий Алексеевич — канд. техн. наук,
✉ dmitry.a.g@gmail.com,

Петров Илья Владимирович — аспирант,
✉ zyxzy@protonmail.ch,

Институт проблем управления им. В.А. Трапезникова РАН, г. Москва.

INFORMATION COMMUNITIES IN SOCIAL NETWORKS. Part II: Networked Models of Formation

D.A. Gubanov[1] and I.V. Petrov[2]

V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

[1]✉ dmitry.a.g@gmail.com, [2]✉ zyxzy@protonmail.ch

Abstract. This survey deals with mathematical models for the formation of information communities under uncertainty. The models of opinion dynamics are considered in detail. Within these models, individuals change their opinions under the influence of other individuals in a social network of a nontrivial structure. Two classes of such models are presented: the models with rational (Bayesian) individuals and the models with naive (heuristic) individuals. For each of the classes, conditions for the formation of information communities in social networks are described. For various information communities to emerge in a society with rational agents, the rationality of individuals is often limited, and some assumptions on different awareness of individuals are introduced considering the network structure. For a society with naive individuals, different modifications of the opinion dynamics mechanism are often adopted.

Keywords: social networks, information community, formation of information communities, analysis of information communities, belief formation.

Funding. This work was supported by the Russian Foundation for Basic Research, projects nos. 19-17-50225 and 18-29-22042.