

© 2022 г. Д.А. ГУБАНОВ, канд. техн. наук (dmitry.a.g@gmail.com)
(Институт проблем управления им. В.А. Трапезникова РАН, Москва)

МЕТОДЫ АНАЛИЗА ИНФОРМАЦИОННОГО ВЛИЯНИЯ В АКТИВНЫХ СЕТЕВЫХ СТРУКТУРАХ¹

Рассмотрен ряд методов и примеров, обосновывающих возможность и целесообразность использования акционального подхода для решения прикладных задач анализа информационных процессов в активных сетевых структурах.

Ключевые слова: активная сетевая структура, информационное влияние, акциональная модель, методы анализа активных сетевых структур.

DOI: 10.31857/S0005231022050075, EDN: ABSYZX

1. Введение

В данной статье рассматриваются построение и применение прикладных методов анализа информационных процессов в активных сетевых структурах, которые основываются на предложенном ранее оригинальном акциональном подходе к расчету информационного влияния [1, 2]. Активная сетевая структура (АСС) — это социальная структура, которая состоит из множества агентов и заданного на нем множества отношений [3]. В АСС отсутствует иерархия, а агенты АСС обладают внутренней активностью и проявляют ее в самостоятельно выбираемых действиях. Частными случаями АСС являются социальные сети, организационные и экономические сети.

К настоящему времени разработан ряд методов анализа сетевых структур (см., например, публикации [4–8]), базирующихся на результатах теории моделирования процессов в сложных сетях, теории социально- сетевого анализа, теории статистического анализа и машинного обучения, в частности: методы анализа сетевой центральности (network centrality), методы выявления сообществ (community detection), методы классификации и кластерного анализа. Акциональный подход позволяет решать эти же прикладные задачи по-новому, он имеет прочное концептуальное основание в виде наблюдаемых минимальных элементов информационного взаимодействия в сети — действий агентов — и учитывает как специфику решаемой практической задачи, так и предпочтения лица принимающего решения.

Структура статьи следующая. В разделе 2 кратко изложена модель информационного влияния в активной сетевой структуре — акциональная модель. В разделе 3 рассмотрена технология анализа информационных процес-

¹ Разделы 3 и 4 выполнены при финансовой поддержке Российского научного фонда (грант № 20-11-20059).

сов в активных сетевых структурах на основе акционального подхода. Раздел 4 содержит примеры, в которых рассматриваются различные прикладные задачи и предлагаются методы их решения: в подразделе 4.1 рассмотрены методы расчета влияния и влиятельности агентов АСС, а также анализа защищенности агентов АСС от информационных воздействий; в подразделе 4.2 приведены методы расчета влияния связей в АСС, а также выявления структур и устойчивых каналов распространения активности; в подразделе 4.3 приведен метод выявления информационных сообществ в АСС.

2. Модель информационного влияния в активной сетевой структуре

Кратко опишем модель информационного влияния, в основном следуя работам [1, 2] (обзоры моделей и методов влияния см. в [9–12]). В основе этой модели лежит действие, совершенное участником сети (агентом). Вводятся обозначения:

- $N = \{1, 2, \dots, n\}$ – множество агентов, входящих в АСС,
- $K = \{1, 2, \dots, k\}$ – фиксированное множество видов действий агентов АСС (в онлайн-социальной сети видом действия, в частности, может быть создание связи подписки или выставление лайка комментарию),
- Δ – конечное множество выполненных агентами действий (например, создание конкретной связи подписки) в заданный период времени T , каждое действие $a \in \Delta$ задается его автором $i \in N$, видом $j \in K$ и моментом совершения $t \in T$.

На множестве Δ задается бинарное отношение частичного порядка $a \rightarrow b$ (действие a является *причиной* действия b), которое обладает свойством рефлексивности (для любого $a \in \Delta$ выполняется $a \rightarrow a$), антисимметричности (если $a \rightarrow b$ и $b \rightarrow a$, то $a = b$) и транзитивности (если $a \rightarrow b$ и $b \rightarrow c$, то $a \rightarrow c$). Если $a \rightarrow b$ и $a \neq b$, но в то же время не существует такого $c \in \Delta$, что $a \rightarrow c$ и $c \rightarrow b$, то $a \downarrow b$ (a является *непосредственной причиной* b). В этом случае бинарные отношения называются *однозначными*. Пример такого отношения в социальных медиа: a – создание учетной записи, b – создание связи подписки на учетную запись.

Для заданного множества $A \subseteq \Delta$ множество всех его последствий определяется следующим образом:

$$\pi(A) = \{b \in \Delta \mid \exists a \in A \ a \rightarrow b\}.$$

Кроме того, выделяется множество *начальных действий* $\Delta^0 \subseteq \Delta$:

$$\Delta^0 = \{a \in \Delta \mid \forall b \in \Delta \ (b \rightarrow a) \Rightarrow (a = b)\}.$$

В рамках акционального подхода проблема расчета влиятельности рассматривается с точки зрения некоторого управляющего субъекта — *центра*.

Центр задает значимость действий участников сети исходя из своих предпочтений, для формализации которых вводится функция *значимости множества действий*

$$\Phi : 2^{\Delta} \rightarrow [0, +\infty).$$

Считается, что, во-первых, значимость множества действий является монотонно возрастающей функцией

$$(1) \quad \Phi(A) \leq \Phi(B), \quad \text{если} \quad A \subseteq B,$$

а во-вторых, по крайней мере некоторые действия в АСС имеют положительную значимость ($\Phi(\Delta) > 0$). Для решения задач на практике эта функция должна быть корректно определена.

Далее на основе акциональной модели можно определить влияние и влиятельность мета-агентов. *Мета-агентом* называется любое непустое подмножество агентов АСС. Для мета-агента $I \subseteq N$ определяется как множество всех выполненных им действий

$$\delta_I = \{a \in \Delta | \alpha(a) \in I\},$$

так и множество выполненных им начальных действий

$$\delta_I^0 = \{a \in \Delta^0 | \alpha(a) \in I\},$$

где функция $\alpha(a)$ каждому действию $a \in \Delta$ ставит в соответствие его автора $\alpha \in N$.

Если деятельность агентов из множества $J \subseteq N$ в достаточно большой степени вызвана деятельностью агентов из множества $I \subseteq N$, то считается, что информационное влияние мета-агента I на мета-агента J велико. Это понимание формализуется с учетом поставленной практической задачи (возможные варианты формализации приведены в [13]). В частности, функцию влияния можно определить так:

$$\chi(I, J) = \begin{cases} \frac{\Phi(\pi(\delta_I^0) \cap \delta_J)}{\Phi(\delta_J)}, & \Phi(\delta_J) > 0; \\ 0, & \Phi(\delta_J) = 0. \end{cases}$$

При совпадении мета-агента J со всем множеством агентов ($J = N$) функция влияния характеризует влияние мета-агента I на всю АСС. В этом случае влияние называется *влиятельностью* $\varepsilon(I)$.

3. Технология анализа активных сетевых структур

Расчет информационного влияния одних агентов сети на других агентов позволяет ставить и решать задачи мониторинга, анализа, прогноза и информационного управления в АСС [14]. Опишем кратко технологию анализа

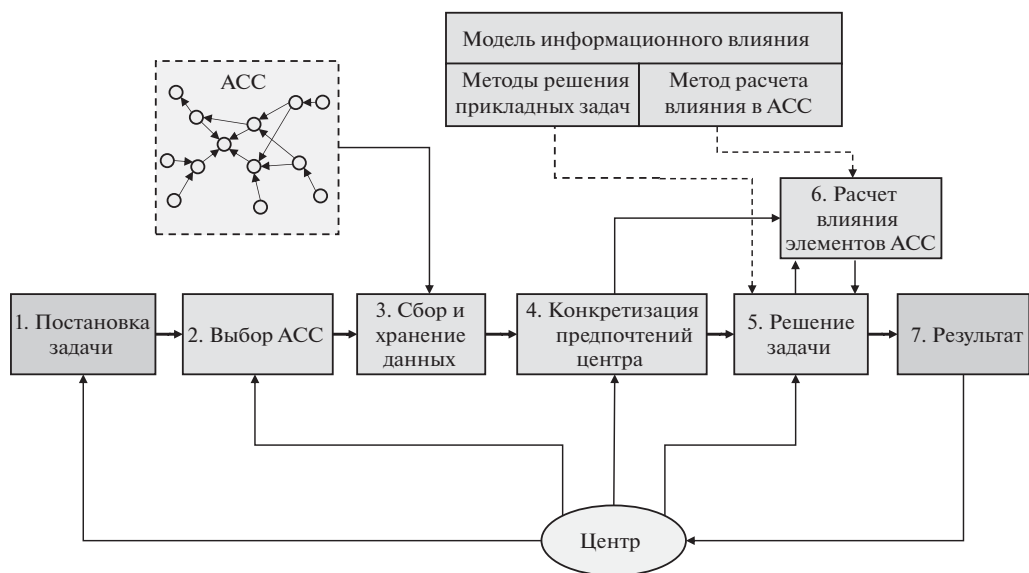


Рис. 1. Технология анализа АСС.

информационных процессов в АСС, которая включает в себя модель информационного влияния (см. рис. 1).

На этапе 1 происходит постановка задачи; на этапе 2 — выбор АСС, на этапе 3 осуществляются сбор данных АСС, структурирование и интеграция данных в хранилище данных; на этапе 4 выполняется конкретизация предпочтений центра (конкретизация функции значимости действий в АСС; выбор варианта формализации функции влияния); на этапе 5 происходит решение прикладной задачи при помощи разработанных на основе модели информационного влияния методов; на этапе 6 выполняется расчет влияния одних элементов АСС на другие; на этапе 7 результаты анализа в подходящем представлении предоставляются центру, который их оценивает и корректирует свои решения.

Приведем наиболее востребованный в прикладных проектах вариант конкретизации предпочтений центра, называемый стандартным в примерах раздела 4 (возможны, конечно, и другие варианты, выбираемые на этапе 4). В этом варианте для центра значимы посты (репосты) и комментарии, в которых упоминается определенный объект (организация, персона и т.п.), а также лайки к таким постам и комментариям. Причем действия должны быть совершены в течение заданного периода времени T . Значимость для центра каждого вида действия определяется известной положительной константой. Рассматриваются следующие виды действий: (1) публикация поста или репоста, (2) выставление лайка посту, (3) публикация комментария к посту или комментарий, (4) выставление лайка комментарий. Таким образом, множество видов действий $K = \{1, 2, 3, 4\}$. Бинарное отношение причинности $a \rightarrow b$ выполнено в следующих случаях: a — пост, b — комментарий к нему; a — пост

или комментарий, b – лайк посту или комментарий; a – пост, b – его репост; a – комментарий, b – комментарий к нему; a и b совпадают. Значимость множества действий определяется как $\Phi(S) = \sum_{a \in S} \Phi(a)$, где $\Phi(a) = 1$, если a – совершенное в течение периода T действие, которое является либо постом с упоминанием заданного объекта, либо комментарием к такому посту, либо лайком такому посту/комментарии, иначе $\Phi(a) = 0$. В качестве подходящего для центра варианта формализации влияния используется приведенный ранее в разделе 2 пример.

Такая технология (и соответственно реализующий ее программный комплекс сбора, структурирования и хранения разнородных данных, а также программный комплекс анализа данных) применяется для исследований онлайн-социальных сетей LiveJournal, Facebook, VKontakte и Reddit.

4. Примеры решения прикладных задач

Приведем примеры решения прикладных задач анализа информационных процессов в АСС на основе предлагаемого подхода:

- расчет влияния и влиятельности агентов АСС, анализ защищенности агентов АСС от информационных воздействий;
- расчет влияния связей между агентами АСС, а также выявление структур и устойчивых каналов распространения активности в АСС;
- выявление информационных сообществ в АСС.

4.1. Расчет влияния и влиятельности агентов, анализ защищенности от информационных воздействий

Вопрос оценки влиятельности участников АСС (в частности, онлайн-социальных сетей) представляет практический интерес, поскольку влиятельные индивиды во многом определяют тематику обсуждений в сети и отношение к обсуждаемым темам. Например, мнения участников сетей можно наблюдать, отслеживая реакцию “лидеров мнений” на актуальные для центра события. Также информация о наиболее влиятельных участниках сети полезна для решения задачи информационного управления — формирования в сети полезной для центра информированности.

Рассмотрим расчет влиятельности пользователей онлайн-социальной сети ВКонтакте [2] (в публикации [1] представлен пример расчета влиятельности на основе акциональной модели для сети Facebook, а в [13] для сети Reddit). Значимыми являются действия за 2015 г. (период T): посты, в которых содержится слово “Назарбаев”, а также их репосты, комментарии и лайки к ним. Предположения о действиях и их взаимосвязи принимаются согласно стандартной конкретизации предпочтений центра (см. раздел 3), однако помимо введенного в его рамках варианта значимости отдельных действий для центра (будем называть его *ненормированным*) рассмотрим и *нормированный* вариант [15]. В этом варианте значимость множества действий каждого агента должна быть равна единице. Считается, что $\Phi(a) = 1/|\delta_{\alpha(a)}|$, если a –

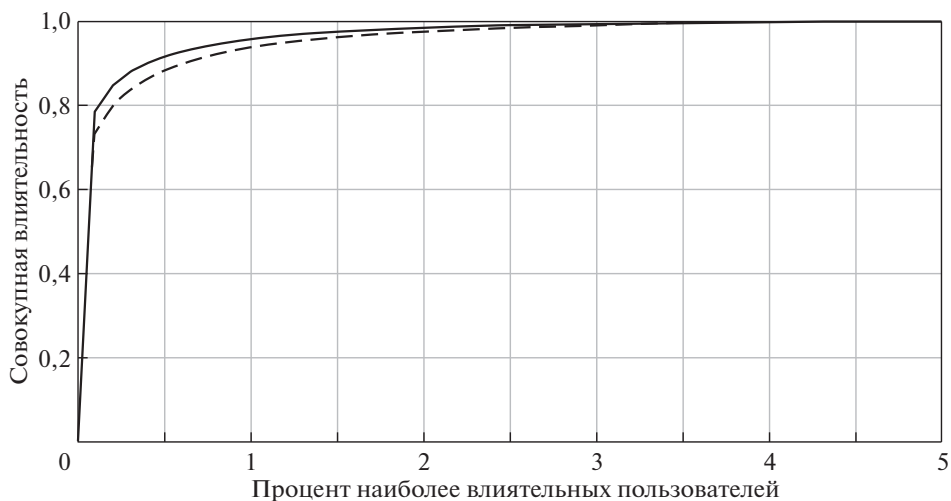


Рис. 2. Зависимость совокупной влиятельности участников сети от их числа.

совершенное в течение периода T действие, которое является либо постом с упоминанием заданного объекта, либо комментарием к такому посту, либо лайком такому посту/комментария, иначе $\Phi(a) = 0$. Следовательно, влияние отдельного агента на расчет влиятельности не зависит от степени его активности. Такая значимость приводит к нормированному способу расчета влиятельности.

Расчет влиятельности пользователей, выполнивших значимые действия, показал, что ненулевую влиятельность имеет всего 38 тысяч участников сети. Совокупная влиятельность всего одного процента наиболее влиятельных участников составляет около 95% влиятельности всей сети, совокупная влиятельность двух процентов наиболее влиятельных составляет 98% влиятельности всей сети (на рис. 2 штриховой линией показано распределение по нормированной влиятельности, сплошной линией показано распределение по ненормированной влиятельности).

Результаты оценки влиятельности показывают, что предложенный метод позволяет с учетом предпочтений центра эффективно находить относительно небольшое число участников сети, оказавших наибольшее влияние на остальную сеть.

Рассмотрим теперь метод оценки информационной защищенности в АСС. Следует отметить, что информация самым существенным образом влияет на жизнь современного общества. Некоторые его ключевые группы являются особенно уязвимыми к целенаправленным информационным воздействиям. При этом высокая защищенность от информационных воздействий может быть как позитивным (например, защищенность от экстремистских идеологий), так и негативным фактором (например, защищенность от пропаганды вакцинации). Поэтому оценка того, насколько те или иные группы населения “защищены” от информационного влияния заданного множества источников воздействий, представляет практический интерес.

Метод информационной защищенности позволяет рассчитать степень защищенности одного мета-агента сети от информационных воздействий другого мета-агента. Мета-агент T (цель) *информационно защищен* от мета-агента S (источник), если действия S не влияют (прямо или косвенно) на действия T .

Будем придерживаться стандартной конкретизации предпочтений центра. Функция влияния мета-агента S на мета-агента T определяется так:

$$\chi(S, T) = \begin{cases} \frac{\Phi(\pi(\delta_S) \cap \delta_T)}{\Phi(\delta_T)}, & \Phi(\delta_T) > 0, \\ 0, & \Phi(\delta_T) = 0 \end{cases}$$

(в [13] это вариант формализации 1).

Тогда информационная защищенность мета-пользователя T от мета-пользователя S определяется следующим образом:

$$\xi(S, T) = 1 - \chi(S, T).$$

Приведем простейший модельный пример расчета. Пусть множество агентов $N = \{1, 2, 3\}$, источник $S = \{1\}$, цель $T = \{2\}$, множество конкретных действий одного вида $\Delta = \{a_1, a_2, a_3, a_4\}$ ($\alpha(a_1) = 1$, $\alpha(a_2) = 2$, $\alpha(a_3) = 3$, $\alpha(a_4) = 2$), причинно-следственные связи между действиями $a_1 \downarrow a_2$, $a_3 \downarrow a_4$. Тогда информационная защищенность цели равна 0,5.

4.2. Расчет влияния связей между агентами, выявление структур и устойчивых каналов распространения активности

Во многих случаях представляет практический интерес оценка значимости не только отдельных участников, но и связей между участниками АСС. Для распространения информации может оказаться важной регулярная реакция одного участника сети на сообщения другого участника. Вызывают интерес также такие связи между участниками (или их подмножествами), которые значимы не только для самих этих участников, но и играют существенную роль в распространении информации в сети в целом. Эти связи образуют каналы распространения информации. На практике анализ значимых связей и каналов позволяет оценить информационный ландшафт АСС в заданной предметной области. Информация о таких связях полезна и для решения задачи информационного управления, например формирования в АСС полезной для центра информированности путем точечного воздействия на связи для ускорения или замедления распространения нужной информации в сети.

Формализуем используемые понятия при помощи акциональной модели [16].

Сетевая значимость непосредственной связи от агента i к агенту j определяется как нормированная значимость последствий действий, выполнен-

ных агентом j под непосредственным влиянием агента i :

$$\xi_{ij} = \frac{\Phi(\pi(\delta_{ij}))}{\Phi(\Delta)},$$

где $\delta_{ij} = \{b \in \delta_j | \exists a \in \delta_i a \downarrow b\}$ – совокупность действий агента j , выполненных под непосредственным влиянием агента i . Сетевая значимость непосредственной связи показывает, насколько большую роль с точки зрения центра играет наличие этой связи для распространения информации в сети.

Индивидуальную значимость для агента j его непосредственной связи с агентом i можно определить как долю значимости последствий действий агента j , которые обусловлены непосредственной связью с агентом i , по отношению к значимости всех действий агента j :

$$\eta_{ij} = \frac{\Phi(\pi(\delta_{ij}))}{\Phi(\pi(\delta_j))}.$$

Выделяются три различных случая связи между агентами.

Случай 1. Малое ξ_{ij} – связь не представляет интереса для рассмотрения.

Случай 2. Большое ξ_{ij} , малое η_{ij} – связь имеет значение для формирования информационной картины в сети, а агент j является влиятельным безотносительно к связи с агентом i .

Случай 3. Большое ξ_{ij} , большое η_{ij} – связь имеет значение для формирования информационной картины в сети, а влияние агента j в большой степени обусловлено его информационной связью с агентом i .

Каналом распространения считается такая последовательность агентов $i_1, i_2, \dots, i_m$, в которой: (1) агенты не повторяются (уникальны), (2) сетевая значимость связей между соседями является довольно большой, т.е. для некоторого параметра $h > 0$ выполняется $\xi_{12} \geq h, \xi_{23} \geq h, \dots, \xi_{m-1,m} \geq h$. *Интенсивностью канала* $i_1, i_2, \dots, i_m$ называется величина $H = \min\{\xi_{12}, \xi_{23}, \dots, \xi_{m-1,m}\}$. Каналы распространения с высокой интенсивностью можно находить, увеличивая значение параметра h и выявляя соответствующие связи между агентами.

Каналы распространения информации в сети могут образовывать связанные подграфы. *Связной структурой каналов с интенсивностью H* называется максимальное по включению множество агентов сети, в котором каждая пара агентов включена в некоторый канал с интенсивностью, большей или равной H .

Представляют особый интерес такие каналы распространения информации, связи в которых имеют большую индивидуальную значимость. Совместное выполнение двух неравенств $\xi_{ij} \geq h, \eta_{ij} \geq h'$ для достаточно больших порогов h и h' означает, что агент j фактически является своеобразным ретранслятором действий агента i , вместе с тем являясь достаточно популярным для прочих агентов.

Приведем пример нахождения связных структур каналов в ВКонтakte [16], для которого применяется стандартная конкретизация предпочтений

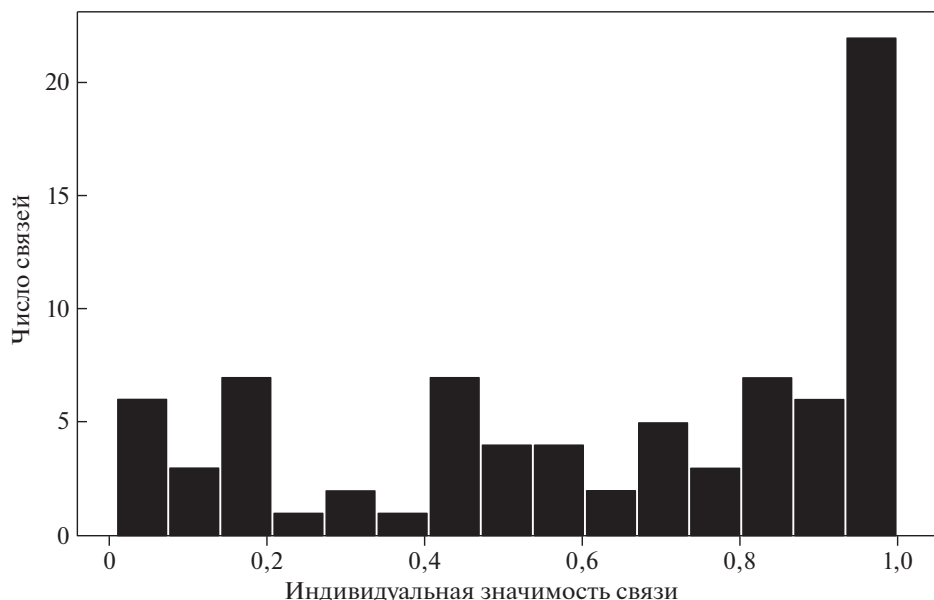


Рис. 3. Гистограмма связей по их индивидуальной значимости.

центра по объекту “Назарбаев”. Значимость каждого из действий для центра одна и та же.

Рассчитаем сетевую значимость связей и рассмотрим связи с большими $\xi_{ij} \geq h = 250^2$, т.е. интересуют указанные выше случаи 2 и 3. Оказывается, значимость 99,995% связей сети меньше данного h . На рис. 3 показана гистограмма распределения найденных связей по их индивидуальной значимости.

Как можно видеть, среди связей с большой сетевой значимостью высока доля связей с большой индивидуальной значимостью (превалирует случай 3 “большое ξ_{ij} и большое η_{ij} ”). Таким образом, для связей, являющихся значимыми для формирования информационной картины в сети, популярность агентов, находящихся на конце связей, во многих случаях обусловлена их информационной связью с агентами в начале связи (они являются ретрансляторами).

Для интенсивности $h = 250$ выявлены связные структуры каналов ВКонтакте (рис. 4,а), а затем связные структуры каналов, в которых связи дополнительно имеют большую индивидуальную значимость ($h' = 0,8$) (рис. 4,б).

Центральными узлами таких структур являются открытые страницы ВКонтакте (паблики) и группы, освещающие различные стороны жизни общества. Публикуемая ими информация ретранслируется как пабликами, так и активными пользователями сети. Даже для больших значений интенсивности h можно выделить устойчивые длинные цепочки пабликов, специализирующиеся на определенной тематике. Эти цепочки, однако, исчезают, если

² Здесь и далее значения h следует делить на $2,2 \cdot 10^6$.

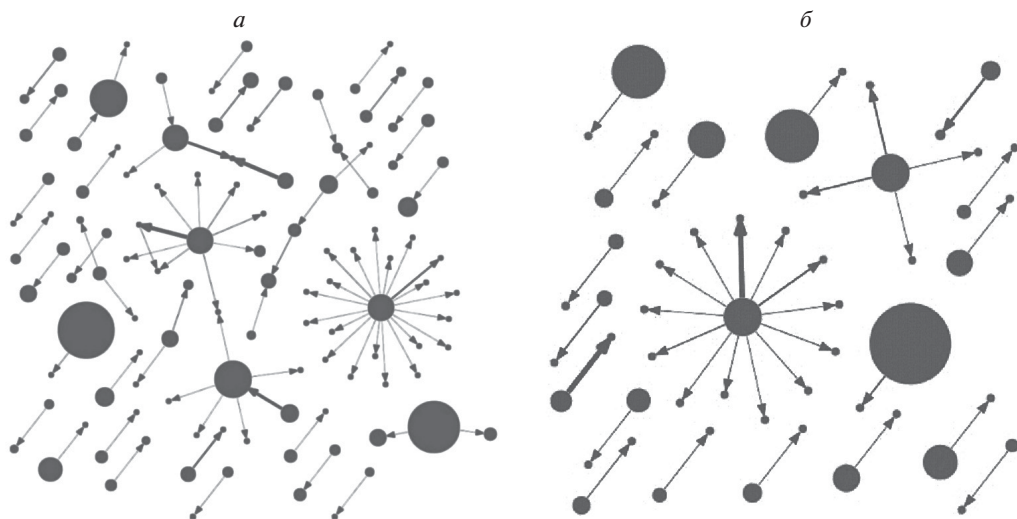


Рис. 4. Связные структуры каналов: $a - h = 250$, $b - h = 250$, $h' = 0,8$. Размер вершины нелинейно соответствует влиятельности участника сети, а толщина ребра соответствует сетевой значимости ребра.

потребовать, чтобы в них все последующие узлы являлись ретрансляторами предыдущих узлов.

На рис. 4 показано, что найденные связные структуры каналов в целом имеют простые конфигурации: чаще всего это цепь из двух вершин. Более нетривиальные структуры можно найти, ослабляя ограничения h и h' . Поиск и содержательный анализ таких структур является перспективным направлением дальнейших прикладных исследований.

4.3. Выявление информационных сообществ

Одной из прикладных задач, представляющих научный и практический интерес, является задача выявления информационных сообществ в АСС [17]. Информационные сообщества состоят из индивидов со сходными представлениями и, как правило, слабо пересекаются по источникам информации, оказывающим на них влияние. Задача выявления информационных сообществ представляется довольно важной, поскольку можно предположить, что участники таких сообществ обладают сходными представлениями, а следовательно, и в ряде случаев и сходным поведением. Таким образом, типизация пользователей позволяет планировать и осуществлять управляющие информационные воздействия [3, 9, 18].

Для решения задачи выявления информационных сообществ предлагается метод, который основывается на выявлении наиболее влиятельных агентов АСС согласно акциональной модели. Здесь важным считается влияние, порождаемое оригинальными действиями (т.е. важно введение в сеть новой информации). Поскольку в онлайн-овых социальных сетях мала доля поль-

зователей с заметной влиятельностью (см. выводы подраздела 4.1), то для характеристики участника сети можно использовать влияния на него со стороны наиболее влиятельных участников (источников информации).

Опишем схему предлагаемого метода.

Шаг 1. Расчет влиятельности участников сети в соответствии с акциональной моделью и отбор наиболее влиятельных участников.

Шаг 2. Формирование стохастических векторов влияний на всех участников сети со стороны влиятельных индивидов (в результате данного шага каждый участник задается стохастическим вектором, характеризующим влияние на него).

Шаг 3. Введение расстояния между двумя участниками (метрики).

Шаг 4. Разбиение участников сети на заданное число сообществ.

На первом и втором шагах в соответствии со стандартной конкретизацией предпочтений центра рассчитывается влиятельность индивидов — источников оригинальных действий и влияние выбранных источников оригинальных действий на действия индивидов сети. Рассчитанные значения влияния представляются в виде стохастической по строкам матрицы $P = (p_{ij})$ размерностью $n \cdot l$, где l — число источников.

Для дальнейшей формализации задачи нахождения сообществ требуется для начала уточнить, что понимается под расстоянием между двумя стохастическими векторами влияния $p_1 = (p_{11}, \dots, p_{1l})$ и $p_2 = (p_{21}, \dots, p_{2l})$. Будем пользоваться следующим расстоянием, которое представляется более подходящим, чем евклидово расстояние [19, 20]:

$$(2) \quad d(p_1, p_2) = 1 - \sum_{j=1}^l \min(p_{1j}, p_{2j}) = \frac{1}{2} \sum_{j=1}^l |p_{1j} - p_{2j}|.$$

Справедливость второго равенства в (2) для стохастических векторов легко доказать [20].

На заключительном шаге ставится задача кластеризации индивидов сети. Если задано желаемое число кластеров k , то необходимо найти такие центры кластеров $a^{(r)} \in \mathfrak{R}^l$, $r = 1, \dots, k$, при которых достигается минимум целевой функции

$$\begin{aligned} F(a^{(1)}, \dots, a^{(k)}) &= \sum_{i=1}^n \min_{r=1, \dots, k} d(P_{i*}, a^{(r)}) = \\ &= \sum_{i=1}^n \min_{r=1, \dots, k} \frac{\mathbf{1}^T |P_{i*} - a^{(r)}|}{2} \end{aligned}$$

при следующих ограничениях $\mathbf{1}^T a^{(r)} = 1$, $a^{(r)} \succeq 0$, $r = 1, \dots, k$.

Минимизация вогнутой функции на выпуклом множестве является вычислительно сложной задачей. На практике для поиска глобального минимума

применяют эвристические алгоритмы поиска локального минимума. Предлагается итерационный алгоритм кластеризации, в котором выделяется два шага.

1. Назначение i -го пользователя кластеру $u_i = \underset{j}{\operatorname{argmin}} d(P_{i*}, a^{(j)});$
2. Поиск центра j -го кластера $a^{(j)} = \underset{a: \mathbf{1}^T a=1, a \geq 0}{\operatorname{argmin}} \sum_{i: u_i=j} d(P_{i*}, a).$

Поиск центроида с учетом заданных ограничений требует отдельного рассмотрения. В [20] для этой задачи приведен алгоритм получения точного решения.

Таким образом, продемонстрирована полная схема метода, реализацию которого можно использовать для выявления информационных сообществ — подмножеств пользователей, характеризующихся общими источниками информации.

Приведем пример выявления информационных сообществ для следующей имитационной модели социальной сети. Пусть в сети обсуждается конечное множество вопросов K . Предпочтения каждого пользователя сети $i \in N$ по этим вопросам задаются стохастическим вектором $\theta_i \sim \operatorname{Dir}(\alpha)$ (вероятностное распределение Дирихле), где вектор $\alpha \in R_{>0}^{|K|}$ определяет априорную значимость вопросов. Каждый вопрос $k \in K$ освещается теми или иными источниками информации из множества V с вероятностями, заданными стохастическим вектором $\phi_k \sim \operatorname{Dir}(\beta)$, где вектор $\beta \in R_{>0}^{|V|}$ определяет априорную значимость источников информации. Пусть пользователь сети i совершает множество действий δ_i за рассматриваемый период T . Каждое действие $a \in \delta_i$ с вероятностью $\theta_{i,k}$ посвящено вопросу k . Если действие пользователя связано с вопросом k , то с вероятностью $\phi_{k,v}$ оно связано с освещением вопроса в источнике $v \in V$. Используя такую имитационную модель, можно рассчитать влияние источников информации на пользователей сети (используется стандартная конкретизация предпочтений центра). Для расчета влияния абсолютное число действий, выполненных пользователем, не является существенным, поэтому предположим, что каждый из пользователей выполнил одно и то же число действий m . Для такой модели (параметры $m = 1000$, $|K| = 5$, $|N| = 1000$, $|V| = 50$, все компоненты вектора β равны 0,02, все компоненты вектора α равны 0,05) проведена серия вычислительных экспериментов, в каждом из которых выполнен поиск сообществ. Среднее значение качества кластеризации NMI (нормированная взаимная информация) при использовании стандартного метода K-means составило 0,91 (максимум качества, 1,0, достигается при совпадении кластеризации с эталонным разбиением), при использовании модифицированного метода – 0,98 (при этом используются одни и те же рассчитанные значения влияния). На рис. 5 представлена диаграмма рассеяния участников сети для одного запуска имитационной модели (размерность векторов влияний на участников сети снижена при помощи метода главных компонент), точки одного и того же сообщества обозначены одинаковыми символами.

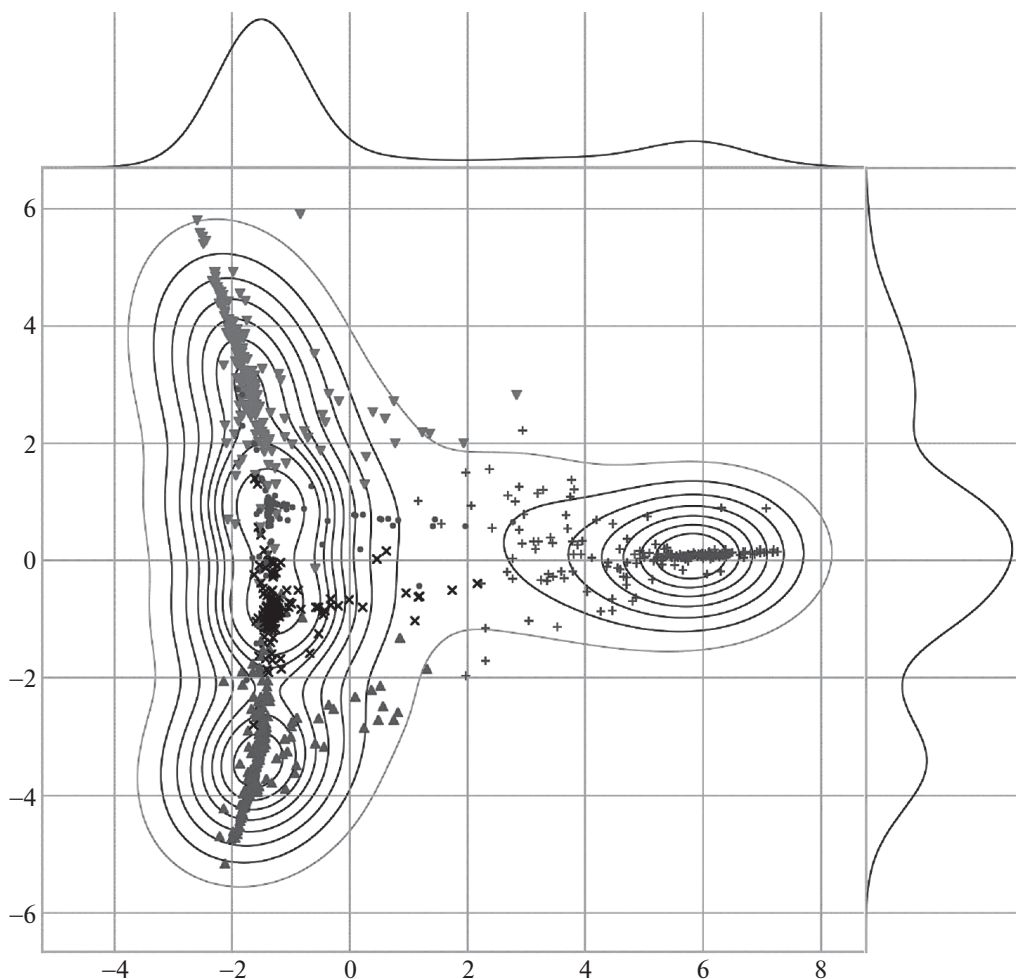


Рис. 5. Диаграмма рассеяния участников сети.

Конечно, возможны и иные сценарии использования акциональной модели для решения задач выявления информационных сообществ в АСС. В частности, в онлайн-овых социальных сетях для расчетов можно использовать следующие виды действий: создание учетной записи пользователя и подписка на пользователя.

5. Заключение

Настоящая статья посвящена демонстрации возможности применения акционального подхода для решения прикладных задач в активных сетевых структурах: кратко изложена акциональная модель, рассмотрена технология анализа информационных процессов в АСС на основе акциональной модели, в соответствии с которой приведен ряд примеров решения прикладных задач. Предложенный подход, несмотря на свою относительную простоту

ту, демонстрирует многообразие возможных постановок практических задач. Конкретная задача анализа информационных процессов в АСС должна формулироваться с учетом как специфики решаемой прикладной задачи, так и предпочтений лица, принимающего решения, — и акциональный подход для этого является достаточно гибким.

СПИСОК ЛИТЕРАТУРЫ

1. Губанов Д.А., Чхартишвили А.Г. Акциональная модель влияния пользователей социальной сети // Проблемы управления. 2014. № 4. С. 20–25.
2. Губанов Д.А., Чхартишвили А.Г. Влиятельность пользователей и метапользователей социальной сети // Проблемы управления. 2016. № 6. С. 12–17.
3. Breer V.V., Novikov D.A., Rogatkin A.D. Mob Control: Models of Threshold Collective Behavior. M.: Springer, 2017.
4. Aggarwal C.C. Social Network Data Analytics. N.Y. – Heidelberg: Springer, 2011.
5. Fouss F., Saerens M., Shimbo M. Algorithms and Models for Network Data and Link Analysis. Cambridge University Press, 2016.
6. Kozitsin I.V., Chkhartishvili A.G., Marchenko A.M., et al. Modeling Political Preferences of Russian Users Exemplified by the Social Network VKontakte. Math Models Comput Simul. 2020. V. 12. P. 185–194.
<https://doi.org/10.1134/S2070048220020088>.
7. Temporal Network Theory / Под ред. P. Holme, J. Saramdki. Cham: Springer International Publishing, 2019.
8. Wasserman S., Faust K. Social Network Analysis: Methods and Applications. 1-е изд. Cambridge University Press, 1994.
9. 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.
10. Flache A., et al. Models of Social Influence: Towards the Next Frontiers // J. of Artificial Societies and Social Simulation. 2017. No. 4(20).
11. Proskurnikov A.V., Tempo R. A Tutorial on Modeling and Analysis of Dynamic Social Networks. Part I // Annual Reviews in Control. 2017. (43). P. 65–79.
12. Proskurnikov A.V., Tempo R. A Tutorial on Modeling and Analysis of Dynamic Social Networks. Part II // Annual Reviews in Control. 2018. (45). P. 166–190.
13. Gubanov D.A. A Study of Formalizations of User Influence in Actional Model // Proc. of the 13th Int. Conf. “Management of Large-Scale System Development” (MLSD). Moscow, Russia: IEEE, 2020.
14. Губанов Д.А., Чхартишвили А.Г. Концептуальный подход к анализу онлайн-новых социальных сетей // Управление большими системами. 2013. Вып. 45. С. 222–236.
15. Chkhartishvili A.G., Gubanov D.A. Analysis of User Influence Types in Online Social Networks: An Example of VKontakte // Proc. of the 11th IEEE Int. Conf. on Application of Information and Communication Technologies (AICT2017, Moscow). M.: IEEE, 2017. V. 1.
16. Chkhartishvili A., Gubanov D. On Approaches to Identifying Information Spread Channels in Online Social Networks // Proc. of the 12th Int. Conf. “Management of Large-Scale System Development” (MLSD). Moscow, Russia: IEEE, 2019.

17. *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.
18. *Novikov D.A., Chkhartishvili A.G.* Mathematical Models of Informational and Strategic Reflexion: a Survey // Advances in Systems Science and Applications. 2014. V. 14. No. 3. P. 254–278.
19. *Gubanov D., Petrov I.* Multidimensional Model of Opinion Polarization in Social Networks // 2019 Twelfth Int. Conf. “Management of large-scale system development” (MLSD). IEEE, 2019. P. 1–4.
20. *Бызов Л.Г., Губанов Д.А., Козицин И.В., Чхартишвили А.Г.* Идеальный политик для социальной сети: подход к анализу идеологических предпочтений пользователей // Проблемы управления. 2020.
DOI: <http://doi.org/10.25728/pu.2020.4.x>

Статья представлена к публикации членом редколлегии А.А. Галяевым.

Поступила в редакцию 12.07.2021

После доработки 25.08.2021

Принята к публикации 26.01.2022