

ВЛИЯНИЕ В СОЦИАЛЬНЫХ СЕТЯХ: ВАРИАНТЫ ФОРМАЛИЗАЦИИ¹

Губанов Д. А.²

(ФГБУН Институт проблем управления
им. В.А. Трапезникова РАН, Москва)

Предложено обобщение разработанных ранее вариантов формализации влияния и влиятельности пользователей онлайн-социальных сетей на основе акциональной модели влияния, которая базируется на эмпирических данных (на действиях пользователей социальной сети) и интересах управляющего субъекта (центра). Рассмотрен ряд вопросов, в зависимости от ответа на каждый из которых понятие влияние агента следует формализовать несколько различным образом: оказывают ли влияние все действия агента, либо только начальные действия; интересует ли центр воздействие на действия агента или воздействие на последствия действий агента; считает ли центр, что максимально возможное влияние на любого агента принимает одно и то же значение, либо нет. Рассмотрены различные комбинации ответов на эти вопросы и приведены соответствующие варианты формализации понятия влияния. Исследованы свойства предлагаемых вариантов формализации (в частности, монотонность и аддитивность влияния, соотношение влияния и влиятельности различных агентов). Приведены примеры расчета влияния и влиятельности пользователей онлайн-социальной сети Reddit. Проведено сравнение разных вариантов формализации влиятельности и показаны существенные различия между ними.

Ключевые слова: социальная сеть, влияние, акциональная модель, варианты формализации влияния.

1. Введение

В современных онлайн-социальных сетях отдельные пользователи играют большую роль в установлении повестки дня, в формировании позитивного или негативного отношения к событиям, персонам и т. п. Возможно поэтому вопросам моделирования и анализа влияния участников социальных сетей посвящено большое число исследовательских работ (см., например, обзоры [1, 13]).

¹ Работа выполнена при частичной поддержке РФФИ (проекты 18-29-22042мк, 20-07-00190А).

² Дмитрий Алексеевич Губанов, к.т.н., с.н.с. (dmitry.a.g@gmail.com).

Со второй половины прошлого века в рамках теории социально-сетевых анализа разрабатываются и исследуются различные меры структурной центральности (близость узла, степень узла, посредничество узла и др. – см., например, [10, 20]), которые характеризуют значимость занимаемой узлом позиции в структуре связей статической сети. Такой подход является математически строгим, но не позволяет при оценке влияния пользователя учесть его роль в информационных процессах, протекающих в социальной сети [5]. К тому же обилие мер центральности (см. пример перечня из более чем двухсот мер [9]), являющихся зачастую неочевидными и необозримыми, несмотря на предпринимаемые попытки аксиоматизации [20], затрудняет выбор подходящей для решения конкретной прикладной задачи меры центральности.

Подход, основывающийся на моделировании информационных процессов в социальных сетях на микроуровне, в отличие от социально- сетевого подхода позволяет естественным образом определять влиятельных пользователей: это пользователи сети, опосредованное влияние которых вызывает наибольшее изменение ее состояния. Примерами данного подхода являются модели динамики мнений в социальных сетях [1, 12, 17] и модели распространения информации/поведения [7, 14, 23]. Однако такой подход пока слабо подходит для расчета влияния пользователей онлайн-социальных сетей из-за сложности идентификации параметров выбранной модели и ее валидации (тем не менее работы в этом направлении ведутся, см., например, [6, 15, 16]). Помимо задачи идентификации для моделей распространения информации/влияния возникает NP-трудная задача определения наиболее влиятельного множества пользователей, которая даже с учетом аппроксимационных алгоритмов является сложной для очень больших социальных сетей [11].

На практике необходимы показатели влияния, которые, с одной стороны, являются релевантными решаемой задаче и отражают динамику реальных информационных процессов, а с другой – могут быть рассчитаны для больших онлайн-социальных сетей с сотнями миллионов взаимодействующих между собой пользователей. Именно поэтому разрабатывается большое число слабо обоснованных, нестрогих, но вычисли-

тельно эффективных эвристических методов расчета влияния, в которых предпринимаются попытки учесть различные аспекты влияния и особенности конкретной онлайн-овой социальной сети (см., например, обзор [19]): от простейших показателей в виде числа друзей и уровня активности, модифицированных мер структурной центральности (обычно специальных версий PageRank [22]) до показателей, основанных на методах машинного обучения [18]. При наличии соответствующих исходных данных эти методы достаточно легко могут быть применены на практике, однако нельзя считать доказанной обоснованность их использования для оценки влиятельности пользователей.

В работах [2, 3] предложен акциональный подход к расчету влияния, который основан на эмпирических данных (на действиях пользователей социальной сети) и интересах управляющего субъекта. В данной работе обобщаются различные варианты формализации влияния и рассматриваются свойства предлагаемых формализаций влияния и влиятельности агентов социальной сети. В разделе 1 описана акциональная модель, в разделе 2 определены понятия влияния и влиятельности, а также разные случаи формализации влияния и влиятельности, в разделе 3 приведены свойства предлагаемых мер влияния и влиятельности. В разделе 4 рассмотрены примеры расчета случаев функций влияния/влиятельности пользователей реальной социальной сети.

2. Акциональная модель

Кратко опишем согласно [3] формальную модель распространения действий в социальной сети. В этой модели базовым элементом анализа является действие, совершенное агентом (пользователем сети).

Пусть участниками сети являются агенты из множества

$$N = \{1, 2, \dots, n\},$$

которые совершают действия того или иного вида из фиксированного множества

$$K = \{1, 2, \dots, k\}$$

в те или иные моменты времени из интервала T . Видом действия может быть написание поста, написание комментария к посту

и т.д. Обозначим множество действий (написание конкретного поста, комментария и т.д.) через Δ . Будем считать это множество конечным.

Каждое действие $a \in \Delta$ характеризуется тремя параметрами: совершившим его агентом, видом действия и моментом времени t , в который действие было совершено:

$$a(i, j, t), i \in N, j \in K, t \in T.$$

Определим функцию $\alpha(a)$, которая каждому действию $a \in \Delta$ ставит в соответствие совершившего его агента $\alpha \in N$.

Далее, пусть на множестве действий задано бинарное отношение частичного порядка « a является причиной b » (или, что будем далее считать эквивалентным, « b является последствием a »), обозначаемое следующим образом:

$$a \rightarrow b.$$

Пример такого отношения в реальной онлайн-социальной сети: a – создание поста, b – создание комментария к этому посту.

Будем считать, что бинарное отношение удовлетворяет следующим свойствам.

1. Рефлексивность: для любого $a \in \Delta$ справедливо $a \rightarrow a$.
2. Транзитивность: если $a \rightarrow b$ и $b \rightarrow c$, то $a \rightarrow c$.
3. Антисимметричность: если $a \rightarrow b$ и $b \rightarrow a$, то a и b совпадают.

Если $a \rightarrow b$ и $a \neq b$, но при этом не существует такого $c \in \Delta$, что $a \rightarrow c$ и $c \rightarrow b$, то будем говорить, что a является *непосредственной причиной* b (или, что будем далее считать эквивалентным, b является *прямым последствием* a). Это позволяет выделить класс бинарных отношений, в которых у каждого действия существует не более одной непосредственной причины. Будем называть такие бинарные отношения *однозначными*.

Приведем пример неоднозначного бинарного отношения. Пусть a – пост, b – комментарий к этому посту, c – другой пост, при этом комментарий b содержит ссылку на пост c . Тогда, если считать справедливым $a \rightarrow b$ и $c \rightarrow b$, бинарное отношение является неоднозначным.

Если задано множество $A \subseteq \Delta$, то можно определить множество всех действий, являющихся последствиями действий из A :

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

Отметим, что для всех множеств $A \subseteq \Delta$ выполняется включение

$$A \subseteq \pi(A),$$

которое справедливо в силу рефлексивности бинарного отношения.

Среди всех действий Δ выделим множество Δ^0 *начальных действий*, которые не являются последствиями какого-либо другого действия:

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

Заметим, что для однозначных бинарных отношений у каждого действия существует ровно одно начальное действие, являющееся его причиной. Поэтому множества $\pi(A)$ и $\pi(B)$ не пересекаются для любых непересекающихся $A, B \in \Delta^0$.

Рассмотрим теперь проблему расчета влияния с точки зрения некоего управляющего органа (*центра*). Пусть центр определяет (исходя из собственных интересов, планов и т.п.), какие именно действия агентов в социальной сети являются значимыми. Для формализации точки зрения центра введем в рассмотрение *значимость множества действий* – функцию $\Phi(S)$, которая каждому множеству действий $S \subseteq \Delta$ ставит в соответствие неотрицательное вещественное число:

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

Естественно предположить, что если к некоторому множеству действий добавить еще действия, то значимость множества увеличится (по крайней мере, не уменьшится). Поэтому будем считать, что значимость множества действий является монотонно возрастающей функцией:

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

Кроме того, примем естественное предположение о том, что хотя бы какие-то действия обладают положительной значимостью: $\Phi(\Delta) > 0$.

Важный класс функций значимости составляют *аддитивные* функции, для которых выполняется соотношение

$$\Phi(A \cup B) = \Phi(A) + \Phi(B)$$

для любых непересекающихся $A, B \in \Delta$.

Для решения конкретных прикладных задач значимость Φ должна быть корректно определена, также должны существовать эффективные алгоритмы расчета ее значения.

3. Влияние агентов и мета-агентов: варианты формализации

Будем называть мета-агентом любое непустое подмножество пользователей сети. Мета-агентом является как каждый отдельный агент $i \in N$ (одноэлементное подмножество $\{i\}$), так и множество всех агентов N .

Определим влияние и влиятельность мета-агентов на основе акциональной модели. Для каждого мета-агента $I \subseteq N$ определим множество $\delta \subseteq \Delta$ всех совершенных им действий

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

а также множество совершенных им начальных действий

$$\delta_I^0 = \{a \in \Delta^0 \mid \alpha(a) \in I\}.$$

Предварительное неформальное понимание влияния можно сформулировать следующим образом: влияние мета-агента $I \subseteq N$ на мета-агента $J \subseteq N$ велико, если деятельность агентов из множества J в достаточно большой степени обусловлена деятельностью агентов из множества I . Формализовать это понимание можно различным образом в зависимости от решаемой практической задачи.

Далее рассмотрим три вопроса, в зависимости от ответа на каждый из которых (ответы пронумерованы) понятие влияние мета-агента следует формализовать несколько различным образом.

Вопрос А. Оказывают ли влияние все действия (A1), либо только начальные (A2)? Предположим, что некий пользователь реальной онлайн-социальной сети не пишет оригинальных постов (т.е. не привносит в сеть какого-либо нового содержания), однако его репосты обладают большой популярностью.

Если центр выбирает ответ А1, то такой пользователь является влиятельным, если же ответ А2 – не является.

Вопрос В. Интересует ли центр воздействие на действия мета-агента (В1) или воздействие на последствия действий мета-агента (В2)? Приведем гипотетический пример. Пользователь *j* написал несколько постов, и все они получили широкое распространение исключительно благодаря тому, что пользователь *i* сделал их репосты. Если центр выбирает ответ В1, то пользователь *i* не влияет на пользователя *j* (который может даже не знать о его существовании). Однако если центр выбирает ответ В2, то влияние является большим: если бы не пользователь *i*, то о пользователе *j*, возможно, никто бы не узнал.

Вопрос С. Считает ли центр, что максимально возможное влияние на любого мета-агента принимает одно и то же значение, т.е. влияние является нормированной величиной (С1), либо нет (С2)? Ответ С1 в некотором смысле «уравнивает» пользователей с разным уровнем активности в сети. Это актуально в случае, когда центр интересуется не столько количество действий в сети, сколько количество совершивших их пользователей.

Комбинируя различные ответы на вопросы А–С, можно получить разные случаи (см. таблицу 1) и, соответственно, разные варианты формализации понятия влияния.

Таблица 1. Варианты формализации влияния

	В1 (воздействие на действия мета-агента)		В2 (воздействие на последствия действий мета-агента)	
	С1 (нормированное влияние)	С2 (ненормированное влияние)	С1 (нормированное влияние)	С2 (ненормированное влияние)
А1 (все действия)	Случай 1	Случай 2	Случай 3	Случай 4
А2 (начальные действия)	Случай 5	Случай 6	Случай 7	Случай 8

Далее для функции влияния мета-агента I на мета-агента J будем использовать одно и то же обозначение¹ $\chi(I, J)$, каждый раз явно оговаривая, какой именно случай рассматривается.

Рассмотрим каждый из случаев.

Случай 1. Сочетанию $A1, B1, C1$ отвечает следующая функция влияния:

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

Нетрудно видеть, что в этом случае для любого $J \subseteq N$ такого, что $\Phi(\delta_J) > 0$, справедливо соотношение

$$\chi(I, J) \leq \frac{\Phi(\pi(\delta_N) \cap \delta_J)}{\Phi(\delta_J)} = \chi(N, J) = 1.$$

Аналогичное соотношение $\chi(I, J) \leq \chi(N, J) = 1$ справедливо для приведенных ниже случаев 3, 5 и 7. Кроме того, для любого $I \subseteq N$ такого, что $\Phi(\delta_I) > 0$, справедливо $\chi(I, I) = 1$, т.е. каждый агент влияет на самого себя (это справедливо также для случая 3).

Случай 2. Сочетанию $A1, B1, C2$ отвечает следующая функция влияния:

$$\chi(I, J) = \Phi(\pi(\delta_I) \cap \delta_J).$$

Случай 3. Сочетанию $A1, B2, C1$ отвечает следующая функция влияния [2]:

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

Случай 4. Сочетанию $A1, B2, C2$ отвечает следующая функция влияния:

$$\chi(I, J) = \Phi(\pi(\delta_I) \cap \pi(\delta_J)).$$

Случай 5. Сочетанию $A2, B1, C1$ отвечает следующая функция влияния [3]:

¹ Для упрощения записи примем следующее соглашение: если аргументом функции влияния (влиятельности) является одноэлементное множество $\{i\}$ ($i \in N$), то вместо $\{i\}$ в аргументе будем писать i

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

Случай 6. Сочетанию A2, B1, C2 отвечает следующая функция влияния:

$$\chi(I, J) = \Phi(\pi(\delta_I^0) \cap \delta_J).$$

Случай 7. Сочетанию A2, B2, C1 отвечает следующая функция влияния:

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

Случай 8. Сочетанию A2, B2, C2 отвечает следующая функция влияния:

$$\chi(I, J) = \Phi(\pi(\delta_I^0) \cap \pi(\delta_J)).$$

Отметим важный частный случай, когда мета-агент J совпадает со всем множеством агентов (т.е. $J = N$) и функция влияния характеризует влияние мета-агента I на всю сеть, которое назовем *влиятельностью* и обозначим $\varepsilon(I)$:

$$\varepsilon(I) = \chi(I, N).$$

Тогда $\delta_J = \pi(\delta_J) = \Delta$ и ответ на вопрос В неважен. Выпишем значение влиятельности для каждого из восьми перечисленных выше случаев:

$$\text{Случай 1, случай 3: } \varepsilon(I) = \frac{\Phi(\pi(\delta_I))}{\Phi(\Delta)}.$$

$$\text{Случай 2, случай 4: } \varepsilon(I) = \Phi(\pi(\delta_I)).$$

$$\text{Случай 5, случай 7: } \varepsilon(I) = \frac{\Phi(\pi(\delta_I^0))}{\Phi(\Delta)}.$$

$$\text{Случай 6, случай 8: } \varepsilon(I) = \Phi(\pi(\delta_I^0)).$$

В зависимости от поставленной практической задачи влияние и влиятельность могут быть вычислены в соответствии с одним из описанных случаев.

4. Свойства влияния и влиятельности

Сформулируем свойства предлагаемых формализаций влияния и влиятельности.

Начнем с выполняемого для всех случаев свойства.

Утверждение 1. Функция влияния $\chi(I, J)$ является монотонной по первому аргументу, т.е. если $I_1 \subseteq I_2$, то для любого J выполняется неравенство $\chi(I_1, J) \leq \chi(I_2, J)$.

Доказательство. Для случаев 1 и 2 утверждение доказывает следующая цепочка соотношений:

$$\begin{aligned} I_1 \subseteq I_2 &\Rightarrow \delta_{I_1} \subseteq \delta_{I_2} \Rightarrow \pi(\delta_{I_1}) \subseteq \pi(\delta_{I_2}) \Rightarrow \pi(\delta_{I_1}) \cap \delta_J \subseteq \pi(\delta_{I_2}) \cap \delta_J \Rightarrow \\ &\Rightarrow \Phi(\pi(\delta_{I_1}) \cap \delta_J) \leq \Phi(\pi(\delta_{I_2}) \cap \delta_J) \Rightarrow \chi(I_1, J) \leq \chi(I_2, J). \end{aligned}$$

Для прочих случаев доказательство проводится аналогично с учетом замены δ_{I_1} и δ_{I_2} на $\delta_{I_1}^0$ и $\delta_{I_2}^0$ соответственно, а также δ_J на $\pi(\delta_J)$. •

Утверждение 1 означает, что чем «больше» мета-агент, тем больше его влияние, независимо от прочих обстоятельств.

Следующие два утверждения относятся к аддитивности функции влияния. Необходимо отметить, что свойство аддитивности является важным с вычислительной точки зрения, поскольку позволяет рассчитывать влияние мета-агентов (или на мета-агентов) как сумму влияний отдельных агентов (соответственно, на отдельных агентов).

Утверждение 2. Если бинарное отношение является однозначным, а функция значимости – аддитивной, то в случаях 5–8 функция влияния является аддитивной по первому аргументу, т.е. для любых $I_1, I_2, J \subseteq N, I_1 \cap I_2 = \emptyset$, выполняется равенство $\chi(I_1 \cup I_2, J) = \chi(I_1, J) + \chi(I_2, J)$.

Доказательство. Для случая 6 утверждение доказывает следующая цепочка соотношений:

$$\begin{aligned} \chi(I_1 \cup I_2, J) &= \Phi(\pi(\delta_{I_1 \cup I_2}^0) \cap \delta_J) = \Phi((\pi(\delta_{I_1}^0) \cup \pi(\delta_{I_2}^0)) \cap \delta_J) = \\ &= \Phi((\pi(\delta_{I_1}^0) \cap \delta_J) \cup (\pi(\delta_{I_2}^0) \cap \delta_J))^{(*)} = \\ &= \Phi(\pi(\delta_{I_1}^0) \cap \delta_J) + \Phi(\pi(\delta_{I_2}^0) \cap \delta_J) = \chi(I_1, J) + \chi(I_2, J). \end{aligned}$$

Здесь при осуществлении ключевого логического перехода (*) использована аддитивность функции Φ , а также тот факт, что вследствие однозначности бинарного отношения множества $\pi(\delta_{I_1}^0)$ и $\pi(\delta_{I_2}^0)$ (а также $\pi(\delta_{I_1}^0) \cap \delta_J$ и $\pi(\delta_{I_2}^0) \cap \delta_J$ соответственно) не пересекаются.

Для получения доказательства в случае 8 достаточно заметить δ_J на $\pi(\delta_J)$. Для получения доказательства в случае 5 (случае 7) следует промежуточные члены цепочки равенств разделить на $\Phi(\delta_J)$ (на $\Phi(\pi(\delta_J))$) (если $\Phi(\delta_J) = 0$ ($\Phi(\pi(\delta_J)) = 0$), то утверждение для случая 5 (случая 7) является тривиальным). •

Ясно, что влияние мета-агента также в данных случаях является аддитивной функцией: для любых непересекающихся множеств $I_1, I_2 \subseteq N$, выполняется равенство $\varepsilon(I_1 \cup I_2) = \varepsilon(I_1) + \varepsilon(I_2)$.

Утверждение 3. Если функция значимости является аддитивной, то в случаях 2 и 6 функция влияния является аддитивной по второму аргументу, т.е. для любых $I, J_1, J_2 \subseteq N, J_1 \cap J_2 = \emptyset$, выполняется равенство $\chi(I, J_1 \cup J_2) = \chi(I, J_1) + \chi(I, J_2)$.

Доказательство. Для случая 2 утверждение доказывает следующая цепочка соотношений:

$$\begin{aligned} \chi(I, J_1 \cup J_2) &= \Phi\left(\pi(\delta_I) \cap \delta_{J_1 \cup J_2}\right) = \Phi\left(\pi(\delta_I) \cap (\delta_{J_1} \cup \delta_{J_2})\right) = \\ &= \Phi\left(\left(\pi(\delta_I) \cap \delta_{J_1}\right) \cup \left(\pi(\delta_I) \cap \delta_{J_2}\right)\right) = \\ &= \Phi\left(\pi(\delta_I) \cap \delta_{J_1}\right) + \Phi\left(\pi(\delta_I) \cap \delta_{J_2}\right) = \\ &= \chi(I, J_1) + \chi(I, J_2). \end{aligned}$$

Для получения доказательства в случае 6 достаточно заметить δ_I на δ_I^0 . •

Утверждение 4. Если бинарное отношение является однозначным, а функция значимости – аддитивной, то в случае 6 функция влияния является аддитивной по обоим аргументам, т.е. для любых $I_1, I_2, J_1, J_2 \subseteq N, I_1 \cap I_2 = J_1 \cap J_2 = \emptyset$, выполняется равенство

$$\chi(I_1 \cup I_2, J_1 \cup J_2) = \chi(I_1, J_1) + \chi(I_2, J_1) + \chi(I_1, J_2) + \chi(I_2, J_2).$$

Доказательство. Утверждение 2 (аддитивность по первому аргументу) справедливо для случаев 5–8, утверждение 3 (аддитивность по второму аргументу) – для случаев 2 и 6. Таким образом, для случая 6 справедливы оба утверждения, что и означает аддитивность по обоим аргументам. •

Утверждение 5. Если множество последствий действий мета-агента I содержится в множестве последствий действий мета-агента J , то влиятельность мета-агента I не превосходит влиятельности мета-агента J :

$$\pi(\delta_I) \subseteq \pi(\delta_J) \Rightarrow \varepsilon(I) \leq \varepsilon(J) \quad (\text{случаи 1–4}),$$

$$\pi(\delta_I^0) \subseteq \pi(\delta_J^0) \Rightarrow \varepsilon(I) \leq \varepsilon(J) \quad (\text{случаи 5–8}).$$

Доказательство.

Случаи 1–4:

$$\pi(\delta_I) \subseteq \pi(\delta_J) \xRightarrow{\text{в силу (1)}} \Phi(\pi(\delta_I)) \leq \Phi(\pi(\delta_J)) \xLeftrightarrow{\text{по определению}} \varepsilon(I) \leq \varepsilon(J).$$

Случаи 5–8:

$$\pi(\delta_I^0) \subseteq \pi(\delta_J^0) \xRightarrow{\text{в силу (1)}} \Phi(\pi(\delta_I^0)) \leq \Phi(\pi(\delta_J^0)) \xLeftrightarrow{\text{по определению}} \varepsilon(I) \leq \varepsilon(J). \bullet$$

Утверждение 6. Если влияние i -го агента на любого из агентов не меньше влияния на него j -го агента, то в случае 3 влиятельность i -го агента не меньше влиятельности j -го агента:

$$\forall k \in N \quad \chi(i, k) \geq \chi(j, k) \Rightarrow \varepsilon(i) \geq \varepsilon(j).$$

Если функция значимости является аддитивной, то утверждение является верным и для случаев 2 и 6.

Доказательство. Докажем утверждение для случая 3. Поскольку левая часть соотношения из утверждения 6 выполняется для всех k , она справедлива также для $k = i$ и $k = j$:

$$(*) \quad \chi(i, i) \geq \chi(j, i), \quad \chi(i, j) \geq \chi(j, j).$$

Если $\varepsilon(j) = 0$, то неравенство $\varepsilon(i) \geq \varepsilon(j)$ очевидно. Пусть $\varepsilon(j) > 0$, при этом $\chi(j, j) = 1$. Тогда из (*) с учетом $0 \leq \chi(i, j) \leq 1$ вытекает, что $\chi(i, j) = 1$. Следовательно, $\varepsilon(i) > 0$. Поэтому выполняется

$$\frac{\chi(i, j)}{\chi(j, i)} \geq 1 \Rightarrow \frac{\Phi(\pi(\delta_i) \cap \pi(\delta_j)) / \Phi(\pi(\delta_j))}{\Phi(\pi(\delta_i) \cap \pi(\delta_j)) / \Phi(\pi(\delta_i))} \geq 1 \Rightarrow \frac{\varepsilon(i)}{\varepsilon(j)} \geq 1.$$

Если функция значимости является аддитивной, то согласно утверждению 3 функция влияния в случаях 2 и 6 является аддитивной по второму аргументу, т.е.

$$\varepsilon(i) = \sum_{k \in N} \chi(i, k), \quad \varepsilon(j) = \sum_{k \in N} \chi(j, k).$$

Тогда $\varepsilon(i) \geq \varepsilon(j)$, поскольку по условию утверждения 6 $\forall k \in N \chi(i, k) \geq \chi(j, k)$. •

Утверждение 7. Если влияние мета-агента I на мета-агента J больше влияния мета-агента J на мета-агента I , то в случае 3 влияние мета-агента I больше влияния мета-агента J :

$$\chi(I, J) > \chi(J, I) \Rightarrow \varepsilon(I) > \varepsilon(J).$$

Доказательство. Если $\chi(I, J) > \chi(J, I)$, то $\chi(I, J) > 0$. Тогда $\Phi(\pi(\delta_J) \cap \pi(\delta_I)) > 0$ и соответственно $\Phi(\pi(\delta_I)) > 0$, следовательно, $\chi(J, I) > 0$. Поэтому, исходя из определения влияния, в случае 3 получаем:

$$\chi(I, J)\varepsilon(J) = \frac{\Phi(\pi(\delta_I) \cap \pi(\delta_J))}{\Phi(\Delta)} = \chi(J, I)\varepsilon(I) \Rightarrow \frac{\chi(I, J)}{\chi(J, I)} = \frac{\varepsilon(I)}{\varepsilon(J)}. \bullet$$

Утверждение 8. Если в результате совершения новых действий в сети ($\Delta' \supseteq \Delta$) влияние мета-агента I на мета-агента J увеличится, то в случаях 2, 4, 6 и 8 увеличится и влияние мета-агента I :

$$\chi_{\Delta'}(I, J) > \chi_{\Delta}(I, J) \Rightarrow \varepsilon_{\Delta'}(I) > \varepsilon_{\Delta}(I).$$

Доказательство. Докажем утверждение для случая 2. В силу того, что исходное множество действий расширилось, а функция значимости является монотонно возрастающей, справедливы следующие соотношения:

$$\Phi(\pi_{\Delta'}(\delta'_I) \cap \delta'_J) > \Phi(\pi_{\Delta}(\delta_I) \cap \delta_J) \Rightarrow \pi_{\Delta'}(\delta'_I) \cap \delta'_J \supset \pi_{\Delta}(\delta_I) \cap \delta_J.$$

Поскольку $\pi_{\Delta}(\delta_I) \cap \delta_J = \pi_{\Delta}(\delta_I) \cap \delta'_J$, то

$$\pi_{\Delta'}(\delta'_I) \cap \delta'_J \supset \pi_{\Delta}(\delta_I) \cap \delta'_J \Rightarrow \pi_{\Delta'}(\delta'_I) \supset \pi_{\Delta}(\delta_I) \Rightarrow \varepsilon_{\Delta'}(I) > \varepsilon_{\Delta}(I).$$

Аналогичные рассуждения применимы для случаев 4, 6 и 8. •

¹ Функция значимости в данном случае является продолжением исходной функции на множество Δ' . Также следует отметить, что нижний индекс Δ для той или иной функции означает, что функция никак не учитывает новые действия из Δ' (в том числе новые последствия действий из множества Δ).

В заключение данного раздела отметим, что некоторые из отмеченных свойств функции влияния для случаев 3 и 5 приведены в статьях [2, 3].

5. Расчеты влияния и влиятельности на примере сети Reddit

В данном разделе рассматриваются примеры расчета влияния и влиятельности пользователей онлайн-социальной сети Reddit (reddit.com). Для анализа использованы посты и комментарии участников форума politics за 2010 год (224 тысячи постов и 2,2 миллиона комментариев) [21]. Пример одного из деревьев обсуждений приведен на рис. 1.

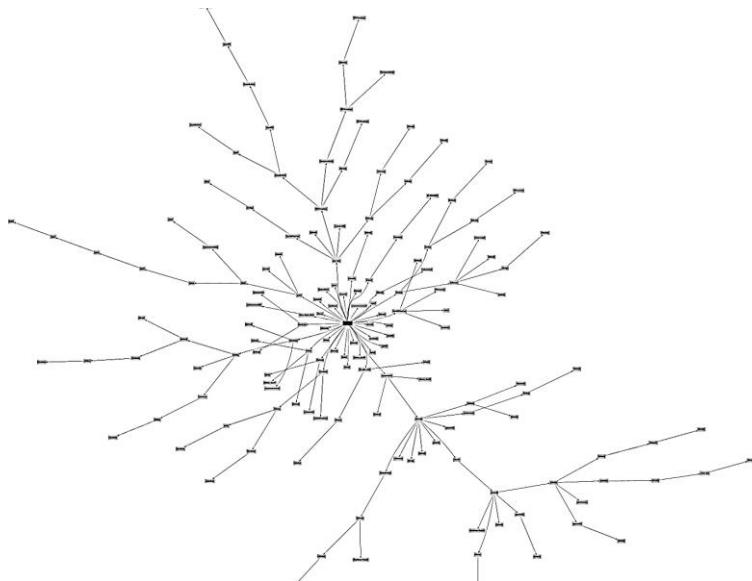


Рис. 1. Фрагмент дерева обсуждений в сети Reddit (центральной вершиной является пост, все остальные вершины являются комментариями к посту или комментариями к комментариям)

В такой сети бинарное отношение между действиями является однозначным. Функцию значимости будем считать адди-

тивной. Значимыми считаются посты и комментарии форума politics за 2010 год, ценность каждого поста и комментария определяется одной и той же положительной константой (без ограничения общности будем считать ее равной 1).

Различия между рассчитанными вариантами влияния A1 и A2 (см. раздел 3 выше) приведены на рис. 2.

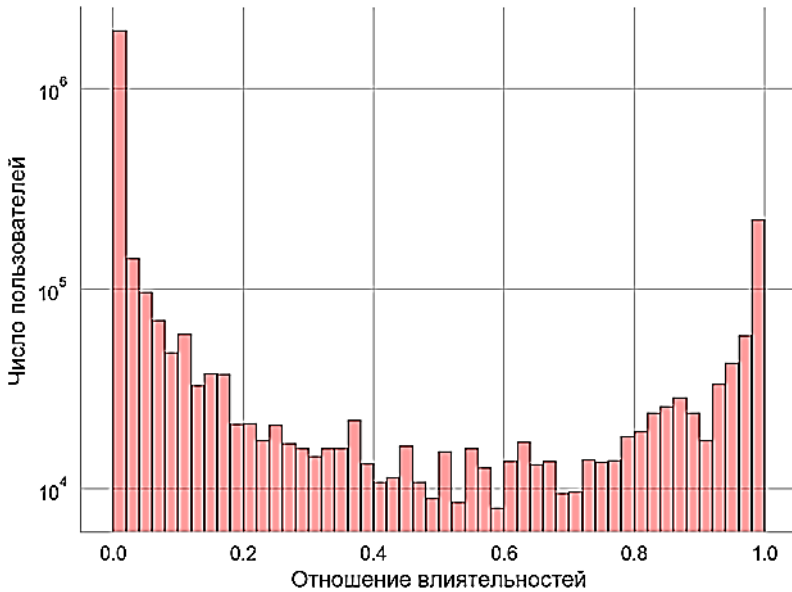


Рис. 2. Распределение числа пользователей по рассчитанному значению отношения влиятельности по начальным действиям к влиятельности по всем действиям

Из рис. 2 можно увидеть довольно четкое разделение пользователей на «писателей» (влиятельность которых обусловлена их начальными действиями) и «комментаторов» (влиятельность которых обусловлена их действиями, не являющихся начальными).

Далее приведем примеры различий между вариантами формализации влияния. На рис. 3–5 представлены диаграммы рассеяния, точками которых являются выявленные связи влияния

между агентами (пользователями Reddit). По осям откладываются значения различных вариантов влияния.

На рис. 3 показаны различия между вариантами A1, B2, C2 и A2, B2, C2. В данном случае различие между ними обусловлено тем, что первый вариант расчета влияния учитывает все действия агента, оказывающего влияние, а второй – только его начальные действия. Коэффициент ранговой корреляции Спирмена составляет 0,066.

На рис. 4 показаны различия между вариантами A1, B2, C2 и A1, B1, C2. Первый вариант расчета влияния рассматривает воздействие на действия агента, а второй – только воздействие на последствия его действий. Коэффициент ранговой корреляции Спирмена составляет 0,314.

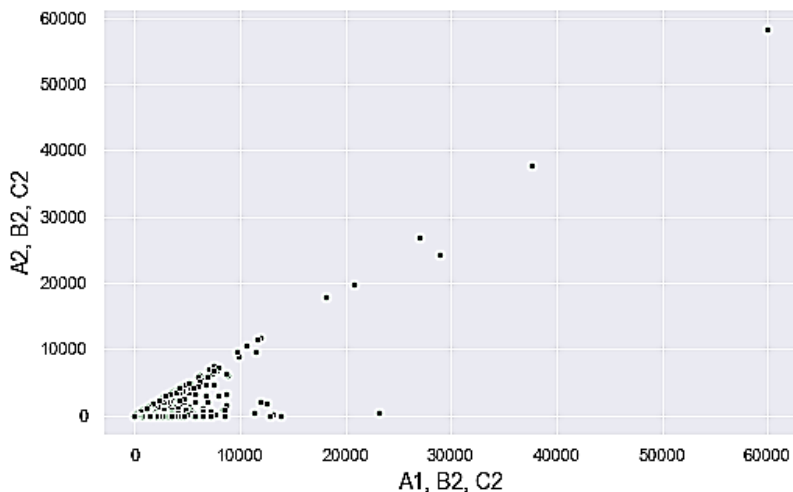


Рис. 3. Диаграмма рассеяния для вариантов A1, B2, C2 и A2, B2, C2

На рис. 5 различия между вариантами A1, B2, C2 и A1, B2, C1 обусловлены тем, что первый вариант рассчитывает ненормированное влияние, а второй – нормированное. Коэффициент ранговой корреляции Спирмена равен 0,312. Необходимо отметить, что здесь основная масса точек сосредоточена в левой части Г-образного графика, в правой части располагаются точки, соответствующие таким связям влияния, для которых испыты-

вающий влияние агент подвергается воздействию ровно со стороны одного агента и последствия такого воздействия значимы.

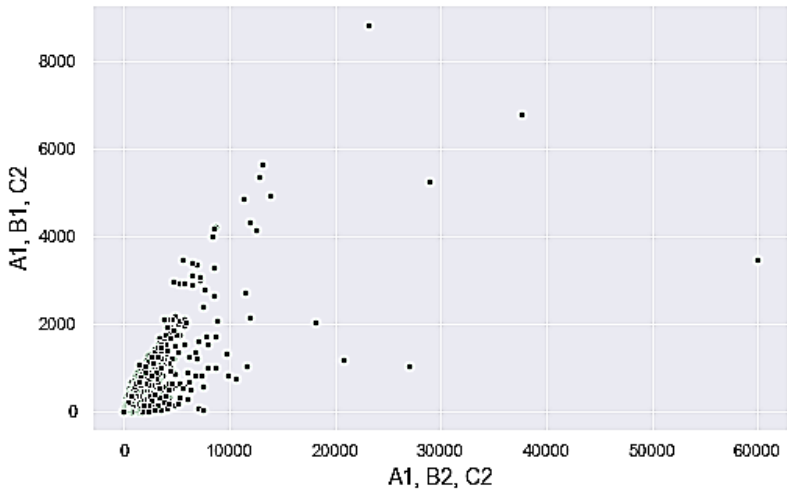


Рис. 4. Диаграмма рассеяния для вариантов $A1, B2, C2$ и $A1, B1, C2$

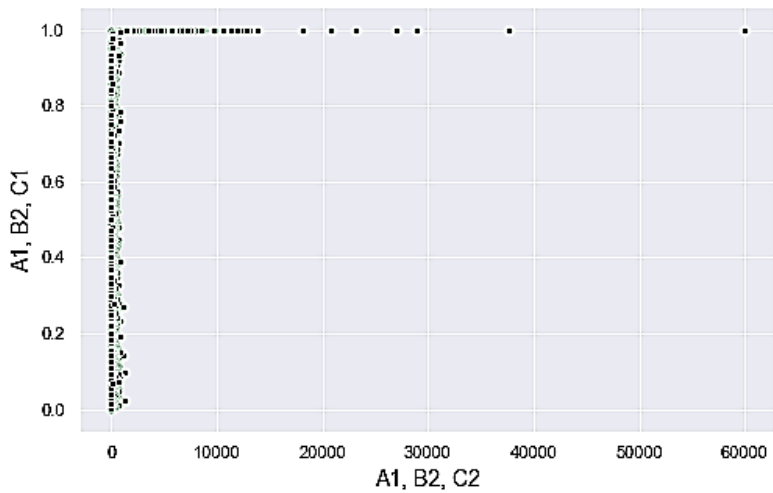


Рис. 5. Диаграмма рассеяния для вариантов $A1, B2, C2$ и $A1, B2, C1$

Таким образом на примере сети Reddit можно заключить, что между предложенными вариантами влияния и влиятельности на практике имеются существенные различия.

6. Заключение

Предложено обобщение разработанных ранее вариантов формализации влияния и влиятельности пользователей онлайн-социальных сетей на основе акциональной модели влияния. Рассмотрены различные случаи формализации влияния и влиятельности мета-агентов сети, доказан ряд свойств функций влияния и влиятельности. Рассмотрены примеры расчета влиятельности пользователей онлайн-социальной сети, проведено сравнение разных вариантов формализации влиятельности и продемонстрированы различия между ними.

Литература

1. ГУБАНОВ Д.А., НОВИКОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Социальные сети: модели информационного влияния, управления и противоборства*. – М.: МЦНМО, 2018. – 224 с.
2. ГУБАНОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Акциональная модель влиятельности пользователей социальной сети // Проблемы управления*. – 2014. – №4. – С. 20–25.
3. ГУБАНОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Влиятельность пользователей и метапользователей социальной сети // Проблемы управления*. – 2016. – №6. – С. 12–17.
4. ГУБАНОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Концептуальный подход к анализу онлайн-социальных сетей // Управление большими системами*. – 2013. – Вып. 45. – С. 222–236.
5. ГУБАНОВ Д.А., ЧХАРТИШВИЛИ А.Г. *Связи дружбы и комментирования пользователей социальной сети Facebook // Управление большими системами*. – 2014. – Вып. 52. – С. 69–84.
6. КОЗИЦИН И.В., ЧХАРТИШВИЛИ А.Г., МАРЧЕНКО А.М., НОРКИН Д.О., ОСИПОВ С.Д., УТЕШЕВ И.А., ГОЙКО В.Л., ПАЛКИН Р.В., МЯГКОВ М.Г. *Моделирование политических взглядов российских пользователей социаль-*

- ной сети ВКонтакте // Математическое моделирование. – 2019. – Т. 31, №8. – С. 3–20.
7. НОВИКОВ Д.А., БРЕЕР В.В., РОГАТКИН А.Д. *Управление толпой: математические модели порогового коллективного поведения.* – М.: ЛЕНАНД, 2016. – 168 с.
 8. AGGARWAL С.С. *Social Network Data Analytics.* – Boston/Dordrecht/London: Kluwer Academic Publishers, 2011. – 502 p.
 9. *CentiServer. The most comprehensive centrality resource and web application for centrality measures calculation.* – URL: <http://www.centiserver.org/> (дата обращения: 03.01.2020).
 10. CSATÓ L. *Measuring centrality by a generalization of degree* // Central European Journal of Operations Research. – 2017. – Vol. 25, No. 4. – P. 771–790.
 11. ERKOL Ş., CASTELLANO С., RADICCHI F. *Systematic comparison between methods for the detection of influential spreaders in complex networks* // Scientific Reports. – 2019. – No. 1(9). – P. 1–11
 12. 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.
 13. JUNG NICKEL K. *New methods of measuring opinion leadership: a systematic, interdisciplinary literature analysis* // Int. Journal of Communication. – 2018. – Vol. 12. – P. 2702–2724.
 14. 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.
 15. KOZITSIN I.V. et al. *Symmetric Convex Mechanism of Opinion Formation Predicts Directions of Users' Opinions Trajectories* //2019 Twelfth International Conference" Management of large-scale system development"(MLSD). – IEEE, 2019. – P. 1–5.

16. MATHIOUDAKIS M. et al. *Sparsification of influence networks* // Proc. of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining – KDD '11. – San Diego, California, USA: ACM Press, 2011. – P. 529–537.
17. PARSEGOV S.E., PROSKURNIKOV A.V., TEMPO R., FRIEDKIN N.E. *Novel multidimensional models of opinion dynamics in social networks* // IEEE Trans. Autom. Control. – 2017. – Vol. 62, No. 5. – P. 2270–2285.
18. RAO A. et al. *Clout score: Measuring influence across multiple social networks* // IEEE Int. Conference on Big Data (Big Data) – 2015. – P. 2282–2289.
19. RIQUELME F., GONZÁLEZ-CANTERGIANI P. *Measuring user influence on Twitter: A survey* // Information Processing & Management. – 2016. – No. 5(52). – P. 949–975.
20. SKIBSKI O., RAHWAN T. *Attachment Centrality: An Axiomatic Approach to Connectivity in Networks* // Proc. of the 15th Int. Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016). – 2016. – P. 168–176.
21. *Stuck_In_the_Matrix. I have every publicly available Reddit comment for research. ~ 1.7 billion comments @ 250 GB compressed. Any interest in this?* – URL: https://www.reddit.com/r/datasets/comments/3bxlg7/i_have_every_publicly_available_reddit_comment/ (дата обращения: 05.01.2020).
22. WENG J., LIM E.-P., JIANG J., HE Q. *Twitterrank: finding topic-sensitive influential twitterers* // Proc. of the Third Int. Conf. on Web Search and Web Data Mining. – 2010. – P. 261–270.
23. ZHILYAKOVA L.YU, GUBANOV D.A. *Double-threshold Model of the Activity Spreading in a Social Network* // Proc. of the 11th IEEE Int. Conference on Application of Information and Communication Technologies (AICT2017, Moscow). – Moscow: IEEE, 2017. – Vol. 2. – P. 267–270.

INFLUENCE IN SOCIAL NETWORKS: FORMALIZATION VARIANTS

Dmitry Gubanov, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Cand.Sc., senior researcher (dmitry.a.g@gmail.com).

Abstract: This paper suggests a generalization of previously proposed cases of formalizing the influence of users in online social networks based on the actional model of influence. We consider a number of questions, depending on the answer to each of which the notion of agent's influence should be formalized in slightly different ways: whether all the agent's actions have an impact or only the initial ones; whether the Principal (a control subject) is interested in the impact on the agent's actions or the impact on the consequences of the agent's actions; whether the Principal believes that the maximum possible influence on any agent takes the same value or not. We consider various combinations of answers to these questions and provide appropriate variants to formalize the notion of influence. We study the properties of the proposed variants of influence formalization (in particular, monotonicity and additivity of the influence function). Then we give examples of calculating the influence of users of the online social network Reddit and compare different variants of influence formalization.

Keywords: social network, influence, actional model, influence cases formalization.

УДК 519.8

ББК 22.18

DOI: 10.25728/ubs.2020.85.3

*Статья представлена к публикации
членом редакционной коллегии Д.А. Новиковым.*

Поступила в редакцию 13.04.2020.

Опубликована 31.05.2020.