DOI: 10.25728/mlsd.2022.0106

О ПРОТИВОБОРСТВЕ НЕСКОЛЬКИХ ИСТОЧНИКОВ ИНФОРМАЦИИ, ФОРМИРУЮЩИХ МНЕНИЯ В СОЦИАЛЬНЫХ СЕТЯХ[1]

Губанов Д.А., Чхартишвили А.Г.
Институт проблем управления им. В.А. Трапезникова РАН,
Россия, г. Москва, ул. Профсоюзная д.65
dmitry.a.g@gmail.com, sandro_ch@mail.ru

Аннотация: Рассмотрена модель смешанных типов агентов социальных сетей, для которой решены задачи информационного управления и противоборства. Проведено имитационное моделирование для ситуации противоборства, рассчитаны равновесия игроков.

Ключевые слова: социальные сети, информационное управление, информационное противоборство.

Введение

В области моделирования динамики мнений в социальных сетях традиционно строятся и исследуются модели [1, 2], в которых мнения и действия агентов отождествляются. В последнее десятилетие ситуация меняется и предлагаются более сложные и реалистичные модели [3, 4]. В данной работе развивается ранее предложенная авторами модель [5, 6, 7] именно такого класса, где мнения агентов (или предпочтения) не наблюдаемы, а наблюдаемые действия не полностью отражают их мнения. Для такой модели формулируются и решаются задачи информационного управления и противоборства, в которых массовое воздействие оказывается на состояние агентов (например, посредством подконтрольных СМИ) для того, чтобы получить выгодные для центра действия в сети. Стратегией центра является выбор степени информационного воздействия. Для случая условно «атомизированной» сети (агенты в сети не доверяют друг другу) получены аналитические решения.

Далее в разделе 1 рассмотрена модель смешанных типов, характеризующих индивидов сети. В разделе 2 содержится постановка и решение задачи информационного управления, в разделе 3 – рассматривается теоретико-игровая модель противоборства и показывается существование доминантных стратегий у центров. В разделе 4 приводятся результаты имитационного моделирования.

1 Модель смешанных типов

Формальное определение предпочтений пользователей онлайновых социальных сетей имеет огромную важность при моделировании информационного управления и информационного противоборства (см., напр., [5, 8, 9, 10]). Эти предпочтения могут иметь идейно-политическую (какое политическое движение поддерживать), рыночную (продукцию какого производителя приобретать) или какую-либо другую природу. Одной из возникающих при этом сложностей является разнообразие типов человеческого сознания в различных его аспектах. Это разнообразие учитывалось при создании теорий и моделей путем введения в рассмотрение типизации индивидов, широко распространенной в психологических, социологических, политологических исследованиях (см. [6, 11, 12, 13] и др.).

Предположение о том, что каждый индивид характеризуется ровно одним типом, является весьма жестким поскольку реальный индивид может совершать действия в соответствии с различными типами в зависимости от внешних или внутренних факторов. В рамках рассматриваемой модели примем предположение о том, что каждый индивид характеризуется *смешанным типом* – стохастическим вектором $(p_1, \dots, p_k)$, j-я компонента которого характеризует склонность индивида совершать действия по типу j (чистый тип, для которого ровно одна компонента равна 1, а остальные равны 0, является частным случаем смешанного). Иными словами, p_j – это координата индивида в базисе чистых типов $(e_1, \dots, e_k)$. Величину p_j будем интерпретировать как вероятность совершения индивидом действия x_j из фиксированного множества действий $\{x_1, \dots, x_k\}$, каждое из которых ассоциирована с соответствующим типом. Например, чистому типу «сторонник такого-то политика» соответствует голосование за этого политика на выборах, а чистому типу «верный пользователь фирмы Apple» – покупка нового телефона именно этой фирмы.

---

[1] *Исследование выполнено при частичной финансовой поддержке РФФИ в рамках научного проекта № 20-07-00190 (Губанов Д.А.)*

2 Управление динамикой мнений агента

Будем считать, что управляющий орган (центр) может управлять мнением агента, сдвигая его в сторону одного из полюсов и затрачивая при этом определенный ресурс. Пусть управлением центра являются затраты объемом $u \geq 0$, посредством которых центр стремится сдвинуть смешанный тип агента в сторону первого (для определенности) полюса, задаваемого единичным вектором $e_1 = (1,0,\dots,0)$. При этом задана монотонно возрастающая на полуоси $[0,+\infty)$ непрерывно дифференцируемая *функция воздействия* $\sigma(u)$, которая каждому управляющему воздействию ставит в соответствие коэффициент сдвига состояний агентов в сторону полюса:

$$\sigma(0) = 0,\ \ \sigma(u) \to 1 \text{ при } u \to +\infty.$$

Примером такой функции является

$$\sigma(u) = \frac{u}{u+\gamma}, \gamma > 0, \tag{1}$$

где параметр γ тем больше, чем менее эффективным является воздействие.

При оказании управляющего воздействия со стороны центра смешанный тип агента $p = (p_1, \dots, p_k)$ меняется следующим образом:

$$p \ \to p^*(u) = \big(1 - \sigma(u)\big)p + \sigma(u)e_1, \tag{2}$$

т.е. новое мнение является выпуклой комбинацией текущего мнения и 1-го полюса.

Интересы центра заданы функцией полезности, определенной на множестве действий агента: $F\big(x_j\big), j \in K$, при этом максимальное значение эта функция принимает при совершении агентом действия (для определенности) x_1, т.е. $F(x_1) > F\big(x_j\big)$, $j \in K \backslash \{1\}$. Тогда решаемая центром задача максимизации ожидаемой полезности имеет следующий вид:

$$\Phi(u) = \sum_{j=1}^{k} p_j^*(u) F\big(x_j\big) - u \underset{u \geq 0}{\longrightarrow} \max. \tag{3}$$

Введем обозначение $F_j = F\big(x_j\big), j \in K$. Тогда, с учетом (2), целевую функцию (3) можно записать следующим образом:

$$\Phi(u) = \sigma(u) F_1 + \big(1 - \sigma(u)\big) \sum_{j=1}^{k} p_j F_j - u \ = \Phi(0) + \Delta_1 \sigma(u) - u, \tag{4}$$

где

$$\Delta_1 = F_1 - \sum_{j=1}^{k} p_j F_j.$$

Величина Δ_1 имеет простой содержательный смысл: это прирост полезности центра при сдвиге состояния агента от текущего к максимально выгодному («идеальному») для центра. Ясно, что $\Delta_1 > 0$ (за исключением случая, когда текущее состояние агента уже является идеальным и $\Delta_1 = 0$).

Нетрудно видеть, что при вогнутой функции воздействия максимум функции $\Phi(u)$ достигается ровно при одном значении u, которое либо является решением уравнения

$$\Phi'(u) = \Delta_1 \sigma'(u) - 1 = 0, \tag{5}$$

либо, в случае $\Phi'(0) < 0$, равно нулю (последний случай означает, что центру невыгодно как-либо воздействовать на агента, поскольку увеличение полезности не компенсирует затраты).

Рассмотрим случай, когда функция воздействия имеет вид (1). Тогда

$$\Phi'(u) = \Delta_1 \frac{\gamma}{(u+\gamma)^2} - 1 = 0 \tag{6}$$

при

$$u = \sqrt{\Delta_1 \gamma} - \gamma, \tag{7}$$

и выбор ненулевого воздействия оптимален для центра при выполнении условия $\Delta_1 > \gamma$. Содержательно это условие означает, что различие между идеальным для центра и текущим состояниями достаточно велико и (или) воздействие на агента является достаточно эффективным.

3 Теоретико-игровая модель противоборства

В предыдущем разделе была рассмотрена модель с одним агентом и одним центром. Теперь обобщим ее на случай, когда имеется множество агентов N, пронумерованных числами от 1 до n (будем обозначать их индексом i), и множество центров M, пронумерованных числами от 1 до m (будем обозначать их индексами l и q). Будем считать, что выполнено условие $m \le k$, т.е. количество центров не превосходит количества чистых типов (полюсов), при этом для l-го центра наиболее предпочтительным является сдвиг состояний агентов в сторону l-го полюса.

При оказании управляющего воздействия со стороны центров смешанный тип i-го агента $p_i = (p_{i1}, \ \dots, \ p_{ik})$ меняется следующим образом:

$$p_i \ \to p_i^*(u_1, \dots, u_m) = \left(1 - \frac{1}{m}\sum_{q=1}^{m} \sigma_q(u_q)\right)p_i + \frac{1}{m}\sum_{q=1}^{m} \sigma_q(u_q)\, e_q, \tag{8}$$

т.е. новое мнение является выпуклой комбинацией текущего мнения и полюсов из множества M (здесь единичный вектор $e_q = (0, \dots, 1, \dots, 0)$, где единица находится на q-м месте, задает q-й полюс).

Интересы l-го центра заданы функцией полезности, определенной на множестве действий каждого агента: $F^l(x_j), j \in K$, при этом максимальное значение эта функция принимает при совершении агентом действия x_l, т.е. $F(x_l) > F(x_j)$, $j \in K \backslash \{l\}$. Таким образом, l-й центр стремится максимизировать суммарную полезность от действий всех агентов, т.е. его целевая функция имеет следующий вид:

$$\Phi_l(u_1, \dots, u_m) = \sum_{i=1}^{n}\sum_{j=1}^{k} p_{ij}^*(u_1, \dots, u_m)F_j^l - u_l \tag{9}$$

Введем обозначение $F_j^l = F^l(x_j), j \in K$. Тогда, с учетом (8), целевую функцию (9) можно преобразовать следующим образом:

$$\Phi_l(u_1, \dots, u_m) = \sum_{i=1}^{n}\left[\sum_{j=1}^{k}\left(1 - \frac{1}{m}\sum_{q=1}^{m}\sigma_q(u_q)\right)p_{ij}F_j^l + \frac{1}{m}\sum_{q=1}^{m}\sigma_q(u_q)\, F_q^l\right] - u_l =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{k} p_{ij}F_j^l + \frac{1}{m}\sum_{q=1}^{m}\left(nF_q^l - \sum_{i=1}^{n}\sum_{j=1}^{k} p_{ij}F_j^l\right)\sigma_q(u_q) - u_l.$$

В результате получаем

$$\Phi_l(u_1, \dots, u_m) = \Phi_l(0, \dots, 0) + \sum_{q=1}^{m} \Delta_q^l \sigma_q(u_q) - u_l, \tag{10}$$

где

$$\Delta_q^l = \frac{1}{m}\left(nF_q^l - \sum_{i=1}^{n}\sum_{j=1}^{k} p_{ij}F_j^l\right).$$

Величина Δ_q^l означает разность, с точки зрения интересов l-го центра, между состоянием, в котором все агенты выбирают q-е действие, и текущим состоянием. Справедливо неравенство $\Delta_q^l = \Delta_l^l > 0$, поскольку действия агентов по l-му типу являются наиболее предпочтительными для l-го центра (за исключением случая, когда текущие состояния всех агентов уже являются наиболее предпочтительными для l-го центра и $\Delta_q^l = 0$).

Из вида целевой функции (10) ясно, что у каждого центра существует доминантная стратегия, максимизирующая (10) при любых действиях остальных центров. Поэтому в игре центров существует равновесие в доминантных стратегиях. Аналогично рассмотренному в предыдущем разделе случаю оптимальное воздействие при вогнутой функции $\sigma_l(u_l)$ достигается ровно при одном значении u_l, которое либо является решением уравнения

$$\frac{\partial \Phi_l(u_1, \ldots, u_m)}{\partial u_l} = \Delta_l^l \sigma_l'(u_l) - 1 = 0,$$

либо, в случае $\Delta_l^l \sigma_l'(0) - 1 < 0$, равно нулю (последний случай означает, напомним, что центру невыгодно как-либо воздействовать на агента, поскольку увеличение полезности не компенсирует затраты).

4 Анализ имитационной модели противоборства в сети

Рассмотрим ситуацию противоборства, в которой управляющие воздействия оказываются на агентов, между которыми отсутствует информационное взаимодействие. Множество участников сети является конечным $N = \{1, \ldots, 100\}$, $n = |N|$. Будем считать, что смешанный тип агента (мнение) представляет собой вектор из $k=3$ компонент, мнение каждого агента задается вектором $x \sim \mathrm{Dir}(\alpha_{dir})$ (вероятностное распределение Дирихле), где вектор параметров $\alpha_{dir} = (1,1,1)$ определяет априорную значимость компонент. Варьируя значения параметров распределения Дирихле, можно получить желаемые распределения мнений в социальной сети: от «равномерно» случайного распределения до распределения с пиками в полюсах (высокая поляризация) и однопикового распределения (низкая поляризация - консенсус). В данном случае используется равномерное распределение. В качестве иллюстрации на рис. 1 приведена тепловая карта на стандартном симплексе мнений пользователей (каждая точка на симплексе характеризует соотношение между компонентами вектора смешанного типа/мнения агента, интенсивностью цвета показаны частоты встречаемости агентов, попадающих в шестиугольники симплекса). Множество центров $M = \{1, 2\}$, функция полезности первого центра $F^1 = (10,0,3)$, а второго – $F^2 = (0,10,3)$, параметр $\gamma = 100$. Управляющее воздействие u варьируется от 0 до 1000 с шагом 100. Число запусков для каждой конфигурации равно 50.

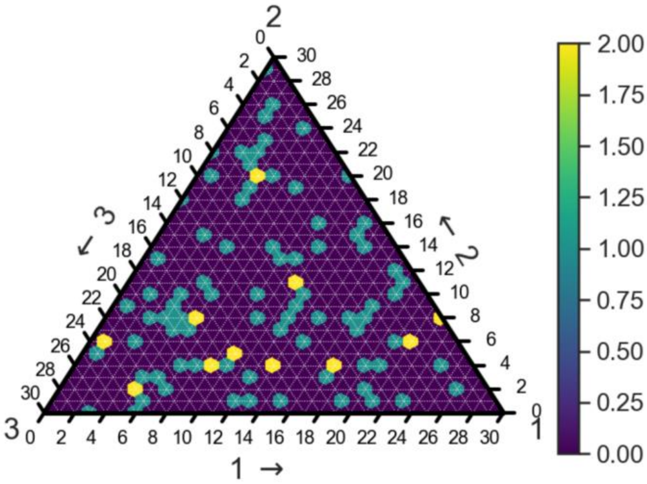
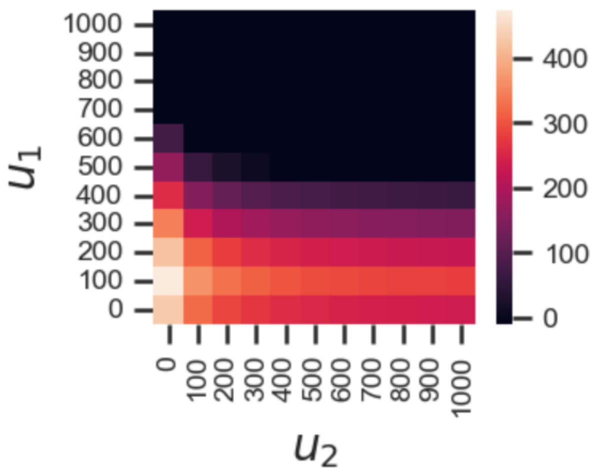


Рис. 1. Тепловая карта мнений агентов

На рисунке 2 приведены рассчитанные матрицы выигрышей центров (точнее тепловые карты выигрышей, при этом выигрыши меньшие нуля обозначены одним цветом).



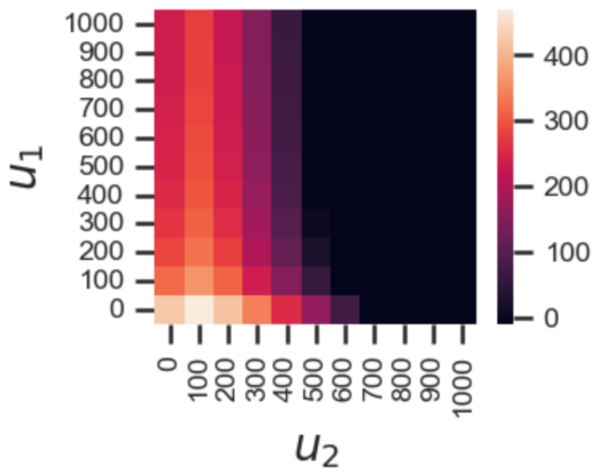


Рис. 2. Матрицы выигрышей центров (слева – первого центра, справа – второго)

Видно, что при любой стратегии оппонента центру выгодно оказывать одно и то же воздействие $u = 100$, это его доминантная стратегия (что соответствует теоретическим результатам). Выигрыши центров при этом практически одинаковые: 366 и 363.

Заключение

В работе рассмотрена модель смешанных типов агентов социальных сетей, для которой решены задачи информационного управления (центр выбирает уровень управляющего воздействия на мнения агентов сети) и информационного противоборства (в ситуации несколько центров, каждый из которых преследует свои интересы). Проведено имитационное моделирование ситуации без информационного взаимодействия агентов сети, для которой рассчитанное равновесие Нэша соответствует теоретическим расчетам.

Литература

1. *DeGroot M.H.* Reaching a Consensus // Journal of American Statistical Assotiation, 1974. – No. 69. – P. 118–121.
2. *Flache A., Mäs M. et al*. Models of Social Influence: Towards the Next Frontiers // The Journal of Artificial Societies and Social Simulation, 2017. – Vol. 20, No. 4.
3. *Новиков Д.А.* Модели динамики психических и поведенческих компонент деятельности в коллективном принятии решений // Управление большими системами, 2020. – Вып. 85. – С. 206–237.
4. *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.
5. *Губанов Д.А., Петров И.В*. О модели поляризации мнений в социальных сетях / Материалы 12-й междунар. конф. «Управление развитием крупномасштабных систем» (MLSD'2019, Москва, ИПУ РАН). – М., 2019. – С. 1200–1202.
6. *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.
7. *Чхартишвили А.Г.* Задача нахождения медианного предпочтения индивидов в стохастической модели // Автоматика и телемеханика, 2021. – № 5. – С. 139–150.
8. *Chkhartishvili A.G., Gubanov D.A., Novikov D.A.* Social Networks: Models of information influence, control and confrontation. – Cham, Switzerland: Springer International Publishing, 2019.
9. *Perra N., Rocha L. E. C.* Modelling opinion dynamics in the age of algorithmic personalisation // Scientific reports, 2019. – Vol. 9. – № 1. – P. 1–11.
10. *Petrov A., Proncheva O.* Modeling Propaganda Battle: Decision-Making, Homophily, and Echo Chambers / Conference on Artificial Intelligence and Natural Language, 2018. – P. 197–209.
11. *Goldberg L.R.* The development of markers for the Big-Five factor structure // Psychological Assessment, 1992. – No. 4. – P. 26-42.
12. *Eysenck S.B.G., Eysenck H.J., Barrett P.* A revised version of the psychoticism scale // Personality and Individual Difference, 1985. – Vol. 6. – No. 1. – P. 21-29.
13. *Keirsey D., Bate M.* Please Understand Me: Character and Temperament Types. Del Mar, CA: Prometheus Nemesis books, 1984.