Оценка степени манипулируемости правил коллективного выбора

Юлия Веселова

Международная научно-учебная лаборатория анализа и выбора решений,
НИУ ВШЭ; Институт проблем управления РАН

Семинар аспирантской школы по компьютерным наукам,
19 мая 2016

Введение

- Принятие коллективных решений - агрегирование множества индивидуальных предпочтений посредством процедуры голосования.
- Недостаток – манипулированию со стороны избирателей: избиратели, действуя стратегически, могут добиться более выгодного для них результата голосования, намеренно исказив свои предпочтения.
- Теорема, доказанная в (Gibbard, 1973) и (Satterthwaite, 1975), утверждает, что любая недиктаторская процедура, в которой участвуют хотя бы три кандидата, является манипулируемой.
- Как сравнить правила по степени манипулируемости?

Введение

- Существует два подхода к сравнению правил по степени манипулируемости:
  - Вычисление вероятности манипулирования: среди всех возможных ситуаций подсчитать долю тех, где возможно манипулирование.
  - Оценка вычислительной сложности задачи манипулирования: насколько трудоемка задача поиска стратегии манипулирования?

Модель: термины и обозначения

- $N = \{1, 2, ..., n\}$ - множество избирателей.
- X - множество альтернатив, $|X| = m$. $L(X)$ - множество всех линейных порядков на X. $W(X)$ - множество всех слабых порядков на X.
- $P_i \in L(X)$ - Предпочтения избирателя i на X .
- Если aP_ib, то альтернатива a более предпочтительна, чем альтернатива b для избирателя i.
- Профиль предпочтений - вектор $\vec{P} = (P_1, P_2, ..., P_n)$. Анонимный профиль предпочтений $\vec{p} = (n_1, n_2, ..., n_{m!})$.
- Вектор распределения позиций для a – $v(a, \vec{P}) = (v_1(a), ..., v_m(a))$, где $v_j(a)$ – количество избирателей, у которых a на j-ом месте в предпочтениях.
- P_M – мажоритарное отношение: aP_Mb если $|\{i \in N : aP_ib\}| > |\{i \in N : bP_ia\}|$.

Модель: термины и обозначения

- Взвешенный граф мажоритарного отношения $WMG(\vec{P})_{kl} = |\{i \in N : a_k P_i a_l\}|$
- Граф мажоритарного отношения

$$MG(\vec{P})_{kl} = \begin{cases} 1, & \text{if } WMG(\vec{P})_{kl} > WMG(\vec{P})_{lk}, \\ -1, & \text{otherwise.} \end{cases}$$

- $\vec{P}_I$ - профиль предпочтений членов коалиции, $I \subseteq N$.
- $\vec{P}_{-I}$ - профиль предпочтений всех избирателей, кроме членов коалиции I.
- $F : L(X)^N \to W(X)$ - функция общественного благосостояния.
- $C_F : L(X)^N \to 2^X$ - правило коллективного выбора.

Правила коллективного выбора

- Правила подсчета очков

$$c \in C_F(\vec{P}) \Leftrightarrow c \in \arg\max_{a \in A} \left(\left\langle s_F, v(a, \vec{P}) \right\rangle \right)$$

  - Правило относительного большинства $s_{Pl} = (1, 0, ..., 0)$,
  - Одобряющее голосование с квотой q (если $q = m - 1$, правило вето) $s_{App} = (\underbrace{1, ..., 1}_{q}, 0, ..., 0)$.
  - Правило Борда $s_B = (m - 1, m - 2, ..., 1, 0)$.
- Двухступенчатая мажоритарная система. Выбирается альтернатива, имеющая более 50 % первых мест. Если такой альтернативы нет, то проводится второй тур голосования, в котором участвуют две альтернативы, набравшие наибольшее количество голосов в первом туре.

Правила коллективного выбора

- Правило Коупленда. Выбирается альтернатива, для которой количество побед минус количество поражений по мажоритарному отношению наибольшее, т.е.

$$CS(a, \vec{P}) = |\{b \in A | aP_M b\}| - |\{b \in A | bP_M a\}|,$$

$$c \in C_{Copeland1}(\vec{P}) \Leftrightarrow c \in \arg\max_{a \in A} CS(a, \vec{P}).$$

- Правило передачи голосов. Выбирается альтернатива, имеющая более 50 % первых мест. Если таковой нет, то из профиля исключается альтернатива, имеющая наименьшее число первых мест. Процедура повторяется до тех пор, пока не будет выбрана какая-либо альтернатива.

Множественный выбор

- Правило устранения несравнимости $T : 2^X \setminus \emptyset \to X$.
- Алфавитное правило: предполагаем заданным некоторый линейный порядок на X, $aP_T bP_T c\dots$, и из данного множества A
- Методы расширения предпочтений: Leximin и Leximax. Если предпочтения избирателя $xP_i yP_i z$, то расширенные предпочтения, EP_i
  - согласно Leximin: $\{x\}EP_i\{x,y\}EP_i\{y\}EP_i\{x,z\}EP_i\{x,y,z\}EP_i\{y,z\}EP_i\{z\}$.
  - согласно Leximax: $\{x\}EP_i\{x,y\}EP_i\{x,y,z\}EP_i\{x,z\}EP_i\{y\}EP_i\{y,z\}EP_i\{z\}$.

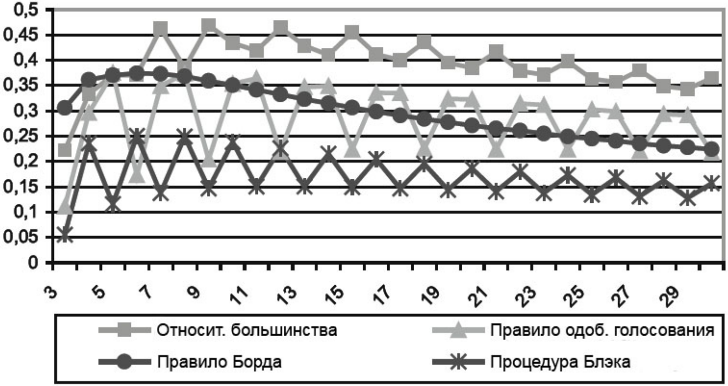
Манипулирование

- Правило коллективного выбора C_F называется манипулируемым, если $\exists \vec{P} \in L(X)^N$, $\exists i \in N$ и $P'_i \neq P_i$ т.ч. $C_F(\widetilde{P}_i, \vec{P'}_{-i})\ EP_i\ C_F(\vec{P})$.
- Индекс манипулируемости Нитцана-Келли – вероятность того, что в случайно выбранном профиле возможно манипулирование (Nitzan, 1985), (Kelly, 1988). Пространством элементарных событий может быть:
  - Множество профилей предпочтений ($\vec{P}$) $\to$ Impartial Culture Model (IC)
  - Множество анонимных профилей предпочтений ($\vec{p}$) $\to$ Impartial Anonymous Culture Model (IAC)
- Сравнению вероятностных моделей посвящена работа (Veselova, 2016).

Вероятность манипулирования

- Индивидуальное манипулирование
  - IC: (Kelly, 1993), (Aleskerov, Kurbanov, 1999), (Aleskerov et al. 2011)
  - IAC: (Pritchard, Wilson, 2007), (Aleskerov et al. 2015)
  - IANC: (Veselova, 2016)
- Коалиционное манипулирование
  - IC: (Pritchard, Wilson, 2007)
  - IAC: (Pritchard, Wilson, 2007), (Lepelley, Valognes, 2003), Slinko (2006).
- Игровое взаимодействие участников: (Favardin, Lepelley, 2006).

Индивидуальное манипулирование в IC



Вычислительная сложность манипулирования

- В [Bartholdi et al. 1989] был впервые рассмотрен вопрос: насколько трудоемка (в смысле вычислительной сложности) задача поиска стратегии манипулирования для избирателя?
- Типы задач манипулирования
  1. Индивидуальное манипулирование или коалиционное?
  2. Конструктивное или деструктивное манипулирование?
  3. Манипулирование: сделать любимого кандидата победителем или добиться более выгодного результата голосования?
  4. Полная или неполная информация?
  5. Одинаковый ли вес имеют голоса избирателей?
  6. Ограничено ли множество кандидатов?

Вычислительная сложность манипулирования

Задача коалиционного манипулирования для правила F (КМ-F)

Даны предпочтения всех избирателей, кроме манипулирующей коалиции, $\vec{P}_{-I}$, и кандидат c. Требуется найти такой профиль предпочтений $\vec{P}_I$ для манипулирующей коалиции I, при котором $c = C_F(\vec{P}_I, \vec{P}_{-I})$ или доказать, что его нет.

F	ИМ-F	КМ-F	КМВ-F	ДКМВ-F
Правило относ. большинства	P [1]	P [2]	P [2]	P [2]
Правило Борда	P [1]	NP-т [5]	NP-п [2]	P [2]
Правило вето	P [1]	P [6]	NP-п [2]	P [2]
Правило передачи голосов	NP-п [4]	NP-п [4]	NP-п [2]	NP-п [2]
Двухст. маж. система	P [6]	P [6]	NP-п [2]	NP-п [2]
Правило Коупленда	P [1]	P [3]	NP-п [2]	P [2]

[1] (Bartholdi et al, 1989), [2] (Conitzer et al, 2007), [3] (Faliszewski et al, 2008), [4] (Bartholdi, Orlin, 1991), [5] (Davies et al, 2011), [6] (Zuckerman et al, 2008).

Случай с неполной информацией

- Функция публичной информации (ФПИ) (poll information function), введенная в (Reijngoud, Endriss, 2012), может быть интерпретирована как результат предварительного опроса избирателей, оглашаемый перед выборами.
  1. Профиль $\pi(\vec{P}) = \vec{P}$.
  2. Анонимный профиль: $\pi(\vec{P}) = \vec{p}(\vec{P}) = (n_1, ..., n_{m!})$.
  3. Позиции: $\pi(\vec{P}) = \vec{v}(\vec{P}) = (v(a_1), ..., v(a_m))$.
  4. Очки: $\pi(\vec{P}) = \vec{S}(\vec{P}) = (S(a_1), ..., S(a_m))$ определяется согласно правилу F.
  5. Ранжирование: $\pi(\vec{P}) = F(\vec{P})$.
  6. Победитель: $\pi(\vec{P}) = C_F(\vec{P})$.
  7. Единственный победитель: $\pi(\vec{P}) = T(C_F(\vec{P}))$
  8. Взвешенный граф мажоритарного отношения: $\pi(\vec{P}) = WMG(\vec{P})$.
  9. Граф мажоритарного отношения: $\pi(\vec{P}) = MG(\vec{P})$.

Информационное множество

- Располагая информацией о профиле $\pi(\vec{P})$ и зная свои предпочтения, избиратель i строит свое информациооное множество

$$W_i^{\pi(\vec{P})} = \{\vec{P'_{-i}} \in L(X)^{N\setminus\{i\}} : \pi(P_i, \vec{P'_{-i}}) = \pi(\vec{P})\}$$

.

- Если $\forall \vec{P} \in L(X)^N \; \forall i \in N \; W_i^{\pi(\vec{P})} \subseteq W_i^{\pi'(\vec{P})}$, то π не менне информативна, чем π'.
- Правило C_F вычислимо из π, если существует функция $H : I \to 2^X \setminus \{\emptyset\}$, т.ч. $C_F = H \circ \pi$.
- Правило C_F сильно вычислимо из π, если информации π достаточно избирателю i, чтобы вычислить результат C_F для любого предпочтения $\widetilde{P}_i \in L(X)$.

Функции публичной информации

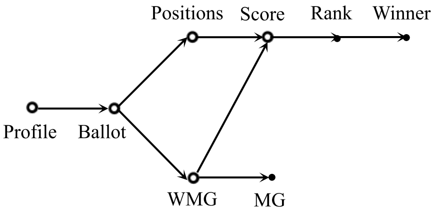


Рис. : Информативность и вычислимость ФПИ для правила Борда

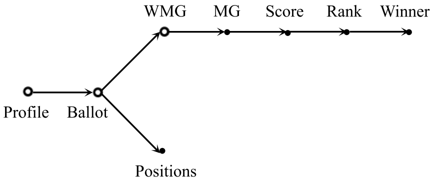


Рис. : Информативность и вычислимость ФПИ для правила Коупленда

Манипулирование при ФПИ π

- Дана функция F и профиль предпочтений $\vec{P}$. Избиратель i имеет стимул к манипулированию при ФПИ π, если $\exists \widetilde{P}_i$ т.ч.
  i) $\forall \vec{P'_{-i}} \in W_i^{\pi(\vec{P})}$ $C_F(\widetilde{P}_i, \vec{P'_{-i}})\ EP_i\ C_F(\vec{P})$ или $C_F(\widetilde{P}_i, \vec{P'_{-i}})\ EI_i\ C_F(\vec{P})$;
  ii) $\exists \vec{P'_{-i}} \in W^{\pi(\vec{P})}$, s.t. $C_F(\widetilde{P}_i, \vec{P'_{-i}})\ EP_i\ C_F(\vec{P})$.
- Функция F называется подвереженной манипулированию при ФПИ π, если $\exists \vec{P} \in L(X)^N$ и $\exists i \in N$, который имеет стимул манипулировать при ФПИ π в профиле $\vec{P}$.
- $I_1(m, n, \pi, F)$ - вероятность того, что в профиле предпочтений, случайно выбранном из $L(X)^N$, есть хотя бы один избиратель, имеющий стимул манипулировать при функции F и ФПИ π.

Теоретические результаты

Теорема 1

Если функция F сильно вычислима из π, то
$I_1(m, n, \pi, F) = I_1(m, n, Profile, F)$.

Теоретические результаты

Теорема 1

Если функция F сильно вычислима из π, то $I_1(m, n, \pi, F) = I_1(m, n, Profile, F)$.

Теорема 2

$\lim_{n\to\infty} I_1(m, n, Winner, Plurality) = 1$ при Leximin и Leximax.
$\lim_{n\to\infty} I_1(m, n, 1Winner, Plurality) = 1$ при алфавитном правиле устранения несравнимости.

Теоретические результаты

Теорема 1

Если функция F сильно вычислима из π, то
$I_1(m, n, \pi, F) = I_1(m, n, Profile, F)$.

Теорема 2

$\lim_{n\to\infty} I_1(m, n, Winner, Plurality) = 1$ при Leximin и Leximax.
$\lim_{n\to\infty} I_1(m, n, 1Winner, Plurality) = 1$ при алфавитном правиле устранения несравнимости.

Теорема 3

При Leximin, $I_1(3, 3, MG, Borda) = 1$.

Теоретические результаты

Теорема 1

Если функция F сильно вычислима из π, то $I_1(m, n, \pi, F) = I_1(m, n, Profile, F)$.

Теорема 2

$\lim_{n\to\infty} I_1(m, n, Winner, Plurality) = 1$ при Leximin и Leximax.
$\lim_{n\to\infty} I_1(m, n, 1Winner, Plurality) = 1$ при алфавитном правиле устранения несравнимости.

Теорема 3

При Leximin, $I_1(3, 3, MG, Borda) = 1$.

Теорема 4

При Leximax, $I_1(3, n, Winner, Copeland) = 0$ для нечетного числа избирателей.

Вычислительные эксперименты

Прведена серия вычислительных экспериментов в MetLab для 6 правил коллективного выбора, 8 ФПИ, для Leximin, Leximax и алфавитного правила устранения несравнимости. Число альтарнатив - 3, число избирателей - от 3 до 15.

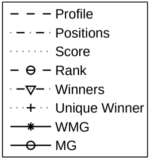


Рис. : It’s a legend

I_1 при алфавитном правиле устранения несравнимости

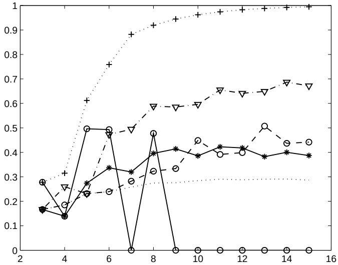


Рис. : Правило относительного большинства

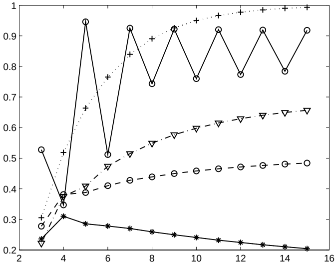


Рис. : Правило Борда

I_1 при алфавитном правиле устранения несравнимости

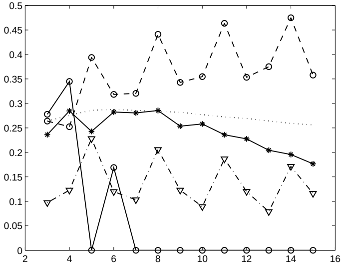


Рис. : Правило одобряющего голосования с квотой 2

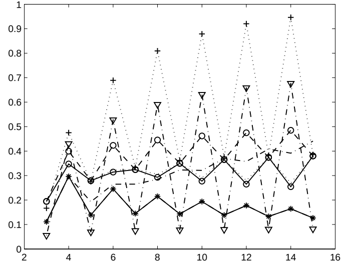


Рис. : Правило Коупленда

Результаты вычислительных экспериментов

- Манипулируемость не уменьшается, когда мы рассматриваем менее информативные ФПИ.
- Для правила относительного большинства и правила Борда максимальные значения I_1 соответствуют ФПИ $1Winner$, наименее информативной ФПИ, из которой правила вычислимы.
- $I_1(3, n, 1Winner, F)$ очень быстро стремится к 1 с ростом числа избирателей для правила относительного большинства и правила Борда.

Результаты вычислительных экспериментов

- Манипулируемость слабо возрастает с уменьшением информативности ФПИ для правила Коупленда. Максимальные значения $I_1(3, n, 1\mathit{Winner}, \mathit{Copeland})$ стремятся к 1.
- Максимальные значения I_1 для правила одобряющего голосования с квотой 2 соответствуют ФПИ *Rank*. При алфавитном правиле устранения несравнимости это правило защищено от манипулирования при ФПИ $1\mathit{Winner}$
- (!) В большинстве случаев манипулируемость увеличилась по сравнению со случаем полной информации.

Текущие задачи исследования

- Рассматривается индекс I_2 - вероятность успеха манипулирования.
- I_3 - агрегированный индекс стимула к манипулированию.

Текущие задачи исследования

- Рассматривается индекс I_2 - вероятность успеха манипулирования.
- I_3 - агрегированный индекс стимула к манипулированию.
- Гипотеза 1: вероятность успеха манипулирования не меняется, если правило вычислимо из ФПИ π.

Текущие задачи исследования

- Рассматривается индекс I_2 - вероятность успеха манипулирования.
- I_3 - агрегированный индекс стимула к манипулированию.
- Гипотеза 1: вероятность успеха манипулирования не меняется, если правило вычислимо из ФПИ π.
- Гипотеза 2: стимул к манипулированию уменьшается при уменьшении информативности ФПИ.

Текущие задачи исследования

- Рассматривается индекс I_2 - вероятность успеха манипулирования.
- I_3 - агрегированный индекс стимула к манипулированию.
- Гипотеза 1: вероятность успеха манипулирования не меняется, если правило вычислимо из ФПИ π.
- Гипотеза 2: стимул к манипулированию уменьшается при уменьшении информативности ФПИ.
- Определение сложности задачи манипулирования при неполной информации различных типов.

Спасибо за внимание!