

Логические средства когнитивной социологии

М.А. Михеенкова, В.К. Финн

Когнитивная социология

- Результат взаимодействия идей классической социологии и современных методов интеллектуального анализа данных
- Получение нового знания на основе анализа эмпирических социологических данных
- **Обоснованная теория** (*grounded theory*) (B. Glaser, A. Strauss) – построение теории на основе эмпирических фактов с использованием индуктивного анализа, причем сбор информации происходит вплоть до насыщения выборки

Интеллектуальный анализ социологических данных

Решение задач качественного АСД средствами
Интеллектуального анализа данных (*knowledge
discovery vs. data mining*)

Извлечение интерпретируемых зависимостей между
различными факторами, неявно содержащимися в
массивах данных

ИАД: Обнаружение нового (относительно имеющихся БФ
и БЗ) знания, извлеченного из БФ и БЗ посредством ИС

ИС – конструктивная имитация (возможно, лишь до
некоторой степени) познавательных способностей
человека (В.К. Финн. Искусственный интеллект: методология,
применения, философия. – М.: URSS, 2010 (в печати))

Существенно: *наличие нечисловых параметров*

Интеллектуальный анализ социологических данных

- Задачи: первичная структуризация данных и знаний, формирование системы отношений (социум), выбор адекватных средств анализа, порождение новых отношений

(измерение – отображение системы отношений в числовые характеристики)

Согласование качественного анализа и количественных характеристик

- Причинная обусловленность – предрасположенность к совершению поведенческих актов при отсутствии противодействующих влияний
- Принцип: «сходство объектов (субъектов) порождает наличие эффекта и его повторяемость, а не наоборот»
(тип каузальности «структура – эффект»)

Качественный анализ социологических данных

- Qualitative Comparative Analysis (C.C. Ragin) – использование булевой алгебры для анализа причин социальных явлений. Причины – интерпретируемые комбинации качественных характеристик. Расширение (fsQCA) – использование теории нечетких множеств.
- Формализованный – средствами интеллектуального анализа данных

Задачи интеллектуального анализа социологических данных

- исследование индивидуального поведения (сравнение, типологизация)
- обнаружение детерминант поведения
- предсказание поведения
- учет влияния ситуации
- распознавание рациональности мнений (в т.ч. степени рациональности мнений данной социальной общности) и отклонений от нее

Рациональность – обобщенная целе-ценностная рациональность (М. Вебер) как аргументированное принятие решений

Интеллектуальные системы (ИС)

ИС = Решатель (задач) + [База фактов (БФ) + База знаний (БЗ)] + комфортный интерфейс

Решатель = Рассуждатель + Вычислитель + Синтезатор

ИС – компьютерная система описанной архитектуры, имитирующая способности естественного интеллекта:

в **автоматическом** режиме – рассуждение, аргументация, рефлексия, обучение, объяснение;

в **интерактивном** режиме – уточнение идей до понятий, адаптация, коррекция знаний и поведения.

Рассуждатель: **индукция, аналогия, абдукция, дедукция** – синтез познавательных процедур, реализующий различные классы эвристик решения задач.

ДСМ-метод автоматического порождения гипотез и логика ИС

- Условия применимости
- ДСМ-рассуждения
- Квазиаксиоматические теории (для баз знаний)
- Дедуктивная имитация правдоподобных рассуждений
- Интеллектуальные системы типа ДСМ

Анализ социологических данных средствами ДСМ-ИС

- Построение формального языка для структуризации данных и аргументации
- Порождение отношений «причина-следствие» (\Rightarrow_2) и «следствие-причина» (\leftarrow_3) из исходного отношения «объект-свойство» (\Rightarrow_1) (индуктивный вывод)
- Модификация БФ и прогнозирование отношения \Rightarrow_1 посредством выводов по аналогии
- Абдуктивное объяснение начального состояния БФ и управление ею
- Анализ рациональности мнений
- Построение квазиаксиоматической теории (открытой) КАТ для анализа социологических данных, систематизирующей факты и знания о социуме

Формализация m -значного закрытого опроса

- T – тема опроса
- $P = \{p_1, \dots, p_n\}$ – каркас темы T
- J_m – m -значная логика, ИЭФ- J_m и ИЭФ*- J_m
- $V_m = \{0, 1/(m-1), \dots, (m-2)/(m-1), 1\}$ – множество оценок переменных
- Заданы атомарные оценки $v^{(i)}[p_j] = v_j^{(i)}$, $i = 1, \dots, m^n$, $j = 1, \dots, n$
- Задана функция оценки квазиформул и формул ИЭФ- J_m и ИЭФ*- J_m
- $?p_j$ – вопрос «какова оценка v корня вопроса p_j ?», $v \in V_m$, ответ – $J_v p_j$
- Ответ респондента b_i по теме T – J_m -максимальная конъюнкция $C_i \stackrel{\square}{=} J_{v_1(i)} p_1 \& \dots \& J_{v_n(i)} p_n$, $i = 1, \dots, m^n, \dots, r$.

Формализация m -значного закрытого опроса

$O_m = \langle J_m, P, \Sigma, K', R \rangle$, где

J_m – m -значная логика ($m \geq 3$),

$P = \{p_1, \dots, p_n\}$ – каркас темы опроса T ,

$\Sigma = \{\psi_1, \dots, \psi_s\}$ – логические зависимости между элементами каркаса, формулы логики J_m ,

K' – стабилизированное множество ответов,

$K' \subseteq K$, K – множество всех возможных ответов,

$|K| = m^n$

$R = \{b_1, \dots, b_r\}$ – множество респондентов (соответствующее K')

$R = \{X \mid \exists \phi J_{\langle v, 0 \rangle} (X \Rightarrow_1 [\phi]) \& \phi \in K'\} ([\phi] = \{J_{v_1} p_1, \dots, J_{v_n} p_n\}; v, v_1, \dots, v_n \in V_m)$

Аргументация

- A – множество доводов (аргументов и контраргументов)

Эмпирические оценки – функции

- $g_j^+: P \rightarrow 2^A, g_j^+(p_i) \subseteq A$

- $g_j^-: P \rightarrow 2^A, g_j^-(p_i) \subseteq A,$

$i = 1, \dots, n, j = 1, \dots, r (r = |R|), g_j^\sigma$ – индивидуальные функции выбора аргументации, $\sigma \in \{+, -\}$

- $G^+ = \{g_1^+, \dots, g_r^+\},$

- $G^- = \{g_1^-, \dots, g_r^-\}.$

- $O_m^a = \langle \langle J_m, P, \Sigma, K', R \rangle, \langle A, G^+, G^- \rangle \rangle$ – m -значный опрос с аргументационной семантикой

Цель – понимание рациональности мнений

Аргументация

- $A \neq \emptyset$, $g^+(p_i) \cap g^-(p_i) = \emptyset$, $g^\sigma(p_i) \subseteq A$, $\sigma \in \{+, -\}$

Атомарная оценка:

- $v[p_i] = 1 \leftrightarrow g^+(p_i) \neq \emptyset, g^-(p_i) = \emptyset$;
- $v[p_i] = -1 \leftrightarrow g^+(p_i) = \emptyset, g^-(p_i) \neq \emptyset$;
- $v[p_i] = 0 \leftrightarrow g^+(p_i) \neq \emptyset, g^-(p_i) \neq \emptyset$;
- $v[p_i] = \tau \leftrightarrow g^+(p_i) = g^-(p_i) = \emptyset$;

Аргументационная семантика возможна и для $m=2, 3$

Логика JA_4

Логика JA_4 – логика распознавания рациональности

- $v[J_1 p]=t \leftrightarrow g^+(p) \neq \emptyset$ и $g^-(p) = \emptyset$
- $v[J_{-1} p]=t \leftrightarrow g^+(p) = \emptyset$ и $g^-(p) \neq \emptyset$
- $v[J_0 p]=t \leftrightarrow g^+(p) \neq \emptyset$ и $g^-(p) \neq \emptyset$
- $v[J_{\top} p]=t \leftrightarrow g^+(p) = g^-(p) = \emptyset$

(частный случай m -значной логики)

Двухуровневый m, l -значный опрос

- Оценки относительно элементов P и темы T формируются независимо
- опрос для каркаса – «внутренний» – m -значный; соответственно, используется логика J_m ,
- опрос по теме – «внешний» – l -значный; логика J_l .
- Опрос определяется расширенно: $O_{m,l} = \langle J_m, J_l, P, \Sigma, K', R \rangle$.

$$R = \{X \mid \exists \phi J_{\langle v, 0 \rangle} (X \Rightarrow_1 [\phi]) \& \phi \in K'\} \quad ([\phi] = \{J_{v_1} p_1, \dots, J_{v_n} p_n\}; \\ v \in V_l, v_1, \dots, v_n \in V_m).$$

$O_{m,l}^a = \langle O_{m,l}, \langle A, G^+, G^- \rangle \rangle$ – m, l -значный опрос с аргументационной семантикой

Непротиворечивость m -значного опроса

- Ответы респондентов $\phi_j = J_{v1(j)} p_1 \& \dots \& J_{vn(j)} p_n$ – J -максимальные конъюнкции логики J_m
- $consis(\Sigma \cup \{\phi\})$ – метаязык непротиворечивости множества формул $(\Sigma \cup \{\phi\})$ (метод аналитических таблиц)
- $K^+ = \{\phi | consis(\Sigma \cup \{\phi\}) \& (\phi \in K)\}$ – множество всех ϕ , не противоречащих Σ
- $\Delta = \{\phi | \neg consis(\Sigma \cup \{\phi\}) \& (\phi \in K)\}$ – множество ответов ϕ , противоречащих Σ (множество “запрещенных максимальных конъюнкций” относительно Σ).

Критерии рациональности опроса

Степень непротиворечивости опроса δ

- $|K'| = r$, функция $\delta(K', \Delta) = 1 - (|K' \cap \Delta| / |K'|)$ есть *степень непротиворечивости* опроса соответствующего множества респондентов.
- Если $K' \cap \Delta = \emptyset$, то опрос *непротиворечив* ($\delta = 1$);
- если $K' \subseteq \Delta$, – *противоречив* ($\delta = 0$);
- если $K' \cap \Delta \neq \emptyset$ и $\neg(K' \subseteq \Delta)$, то $0 < \delta < 1$

K' – стабилизированное множество ответов

Δ – множество “запрещенных максимальных конъюнкций”

Критерии рациональности опроса

Близость к «идеальному мнению» ρ

- i -я партия, «идеальное партийное мнение» $\phi_j = J_{v1} p_1 \& \dots \& J_{vn} p_n$;
- $\alpha_k^{(i)}$ – число респондентов с $(n - k)$ пересечений с идеальным мнением ϕ_j (сходство в $(n - k)$ пунктах программы, $\alpha_k^{(i)}$ – число сходств), $\beta^{(i)}$ – число сторонников i -й партии при опросе;
- $p^{(i)} = \alpha_0^{(i)} \cdot n + \alpha_1^{(i)} \cdot (n - 1) + \dots + \alpha_k^{(i)} \cdot (n - k) + \dots + \alpha_{(n-1)}^{(i)} \cdot 1$
- $\rho^{(i)} = p^{(i)} / n \beta^{(i)}$

Критерии рациональности опроса «Степень согласованности» λ

- Для фиксированной темы T^* (сторонники одной партии)
- Множество мнений $\Phi = \{\phi_1, \dots, \phi_s\}$
- Множество респондентов, имеющих мнение ϕ_j : $B_j = \{X \mid J_{\langle 1, 0 \rangle}(X \Rightarrow_1 \langle [\phi_j], T^* \rangle)\}$.
- Число респондентов, имеющих мнение ϕ_j : $|B_j| = \alpha_j, j = 1, \dots, s$.
- $[\phi_i] \cap [\phi_j] = [\phi_{ij}], j = i+1, \dots, s, |[\phi_{ij}]| = m_{ij}, \alpha_{ij} = \min(\alpha_i, \alpha_j)$.
- $\gamma_j = (\alpha_{j1} m_{j1} + \dots + \alpha_{js} m_{js}), \gamma = (\gamma_1 + \dots + \gamma_{s-1}), \gamma' = (n-1) \sum \alpha_j (s-j), j = 1, \dots, s-1$.
- Степень согласованности $\lambda = \gamma/\gamma'$.

Задача изучения мнений

Исходные элементы БФ $J_{\bar{v},0c}(C_j \Rightarrow_1 [\phi_j])$ – «субъект C_j имеет мнение ϕ_j »

Гипотезы $J_{\bar{v},nc}([\psi_j]_3 \Leftarrow C'_j)$ – «мнение ψ_j есть следствие характеристик субъекта C'_j » ($n > 0$)

$\langle v, n \rangle$ – истинностное значение (оценка), полученная применением ДСМ-метода АПГ, где $v \in \{\pm 1, 0, \tau\}$, а n – число применений ДСМ-правил правдоподобного вывода

Изучение электорального поведения

Анализ и предсказание электоральных предпочтений студентов старших курсов РГГУ накануне декабрьских (2003, 2007 гг.) выборов в Государственную Думу

описание респондента («портрет личности») \Rightarrow мнение (выбор партийных программных установок) + выбор электорального действия (конкретная партия, другие партии, против всех, неучастие в выборах): $C_i \Rightarrow_1 \langle [\phi_i], T_j^* \rangle$

Эксперимент

порождение детерминант электорального поведения (мнение, выбор действия), где «мнение» есть выбор программных установок (без указания в исходных данных их принадлежности конкретной партии), а «действие» – выбор одной из шести заданных партий или другой, или выбор «против всех» или отказ от участия в выборах;

при этом (–)-примерами для каждого действия (демонстрации партийных предпочтений) оказываются голосующие за все другие партии.

Формирование мнения

p_1, \dots, p_n – программные установки (по проблемам «земля», «приватизация», «СМИ», «внешняя политика», «экономика», «свобода личности», «социальная политика» и.т.п.)

p_{57} : «Земля всецело должна быть в государственной собственности» (КПРФ)

p_{61} : «Социальная рыночная экономика наиболее эффективна в условиях России» (Яблоко)

p_{68} : «Необходимо узаконить итоги приватизации» (СПС)

p_{85} : «Основная угроза России исходит от США» (ЛДПР)

Мнение – ответы на предложенные вопросы с соответствующими оценками (из множества $\{1, -1, 0, \tau\}$)

Эксперимент

Анализ рациональности выбора программных установок и электорального действия (выбора политической партии в том числе).

Рациональность

57) Земля всецело должна быть в государственной собственности.

67) Продажа земли в частные руки должна быть строго ограниченной.

84) Земля должна постепенно приватизироваться, но при условии строгого соблюдения законности.

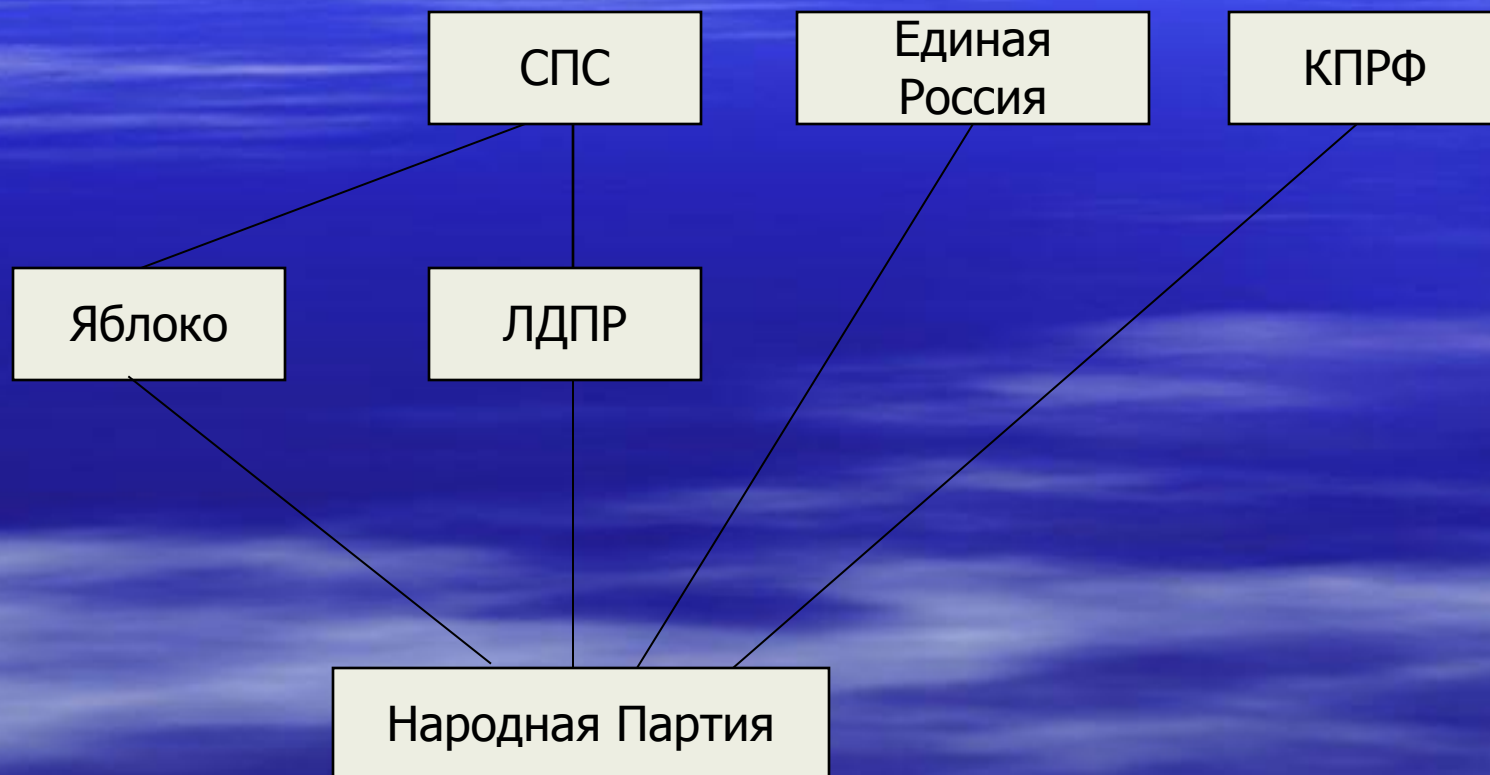
88) Необходима свободная без ограничений продажа сельскохозяйственных угодий.

- $J_1 p_{57} \rightarrow J_{-1} p_{88}$
- $J_{-1} p_{57} \rightarrow J_1 p_{88}$
- $J_1 p_{57} \rightarrow J_{-1} p_{84}$
- $J_1 p_{67} \rightarrow J_{-1} p_{88}$

Рациональность (непротиворечивость)

	Всего	Противор.	Непротив.	Степень непротив.
Все респонден.	231	185	46	0,199
СПС	27	21	6	0,222
Против всех	57	45	12	0,211
ЛДПР	5	4	1	0,2
Яблоко	27	22	5	0,185
Другое	11	9	2	0,182
Ед.Рос.	21	19	2	0,095
КПРФ	5	5	0	0
Народ. пар.	6	6	0	0

Рациональность (3 критерия) (2003)



$$\mathbf{R} = \langle R, \leq \rangle, R = \langle \delta, \lambda, \rho \rangle, R^{(i)} \leq R^{(j)} \leftrightarrow ((\delta^{(i)} \leq \delta^{(j)}) \& (\lambda^{(i)} \leq \lambda^{(j)}) \& (\rho^{(i)} \leq \rho^{(j)}))$$

Рациональность (2 критерия) (2003)

	Всего	Близость к идеальному мнению р	Степень согласованности λ
Ед. Россия	21	0,680	0,625
СПС	25	0,571	0,598
КПРФ	5	0,700	0,540
Яблоко	25	0,533	0,472
ЛДПР	5	0,350	0,329
Нар. партия	5	0,233	0,240



Рациональность (2 критерия) (2007)

	Всего	Близость к идеальному у мнению р	Степень согласованности λ
СПС	14	0,673	0,559
Яблоко	17	0,412	0,447
Гражд. сила	4	0,571	0,444
ЛДПР	8	0,5	0,441
Ед. Россия	31	0,407	0,406
КПРФ	8	0,429	0,351
Справ. Россия	3	0,333	0,333



Предсказательный опрос

- Расширение базы фактов БФ на основе абдуктивного объяснения
- Пополнение БЗ доопределенными примерами отношения

K'' – стабилизированное множество ответов по завершении вычислений, $K'' = K' \cup K^*$, где K^* - множество доопределенных ответов.

- Предсказание ответов
- Определение степени рациональности опроса (по трем критериям)
- Определение существенности параметров опроса (элементов P)

Предсказательный опрос

- $O_{4,\infty}$ – предсказательный опрос, $V_{in}^{(\infty)} = \{\langle v, n \rangle \mid (v \in \{\pm 1, 0, \tau\}) \& (n \in \mathbb{N})\}$
- $\forall X \forall \phi ((X \in B^*) \rightarrow (J_{\langle \tau, 0 \rangle} (X \Rightarrow_1 [\phi]) \& (\phi \in K)))$
- $\check{K} = K \setminus K'$
- $K_{v,n}^* = \{\phi \mid \exists X J_{\langle v, n \rangle} (X \Rightarrow_1 [\phi]) \& (X \in B^*) \& (v \in \{\pm 1, 0, \tau\}) \& (n \geq 1)\}$
- $K_n^* = K_{1,n}^* \cup K_{-1,n}^* \cup K_{0,n}^* \cup K_{\tau,n}^*$
- $\check{K}_n^* = K_n^* \cap \check{K}$ – множество предсказанных мнений
- $\theta(K', K^+) = |\check{K}_n^* \cap K'| / |K^+|$ – (рациональная) полнота предсказания

Задачи

- Предсказание ответов – любая реализация эвристической схемы
- Определение степени непротиворечивости опроса
- Определение существенности параметров опроса (элементов P)

*Т. Парсонс «О теории и метатеории»
(Теоретическая социология, Антология,
Т.2, с. 44 – 45)*

«Во-первых, Вебер определял социологию как научную дисциплину, которая, в первую очередь, должна попытаться понять действия индивидов, особенно в их социальных отношениях.

...Во-вторых, Вебер считал, что социология помимо субъективных мотивов должна развивать каузальные объяснения процесса действия, его направления и последствий».

ИАД как инструмент когнитивной СОЦИОЛОГИИ

Анализ данных – предсказание – объяснение –
(новая проблема)

(язык представления знаний с аргументативной и
дескриптивной функцией)

Сходство – аналогия – абдукция:

- алгебра логики – аналогия – абдукция, *AlAnAb*
(ср. *Ragin C.C.*) (закрытый мир);
- индукция – аналогия – абдукция, *InAnAb* (ср.
ДСМ-метод АПГ) (открытый ми и обучение).

- *Когнитивная социология* – условия применимости (онтология для социологических данных) + эпистемология (ИАД для социологических данных)
- *Формализованные средства* – ДСМ-метод АПГ и соответствующие ему логики
- *Инструмент анализа* – интеллектуальные системы типа ДСМ (ИС-ДСМ).

- ***Когнитивная социология*** – взаимодействие идей классической социологии и возможностей ИАД
- Ответ на вызов качественной социологии (потребность в формализации КАД)
- Когнитивная социология vs. Социальная философия и статистические вычисления

Некоторые публикации

- *Страусс А., Корбин Дж.* Основы качественного исследования. Обоснованная теория. Процедуры и техники. М.: КомКнига. – 2007.
- *Ядов В.А.* Стратегия социологического исследования. М.: Добросвет, 2003.
- *Ragin C.C.* The Comparative Method: Moving beyond Qualitative and Quantitative Strategies // Berkley, Los Angeles and London: University of California Press, 1987.
- *Rihoux B.* Qualitative Comparative Analysis and Related Systematic Comparative Methods // International Sociology, v. 21 (5), September 2006.
- *Финн В.К.* Интеллектуальные системы и общество. Изд. 3-е. М.: КомКнига, 2007.
- Автоматическое порождение гипотез в интеллектуальных системах, М.: Книжный дом «Либроком», 2009, Введение, Гл.1, Гл.4; Часть III.