

Sociálne siete a fuzzy formálna konceptová analýza

Stanislav Krajči

UPJŠ Košice

Jana Krajčiová

Osemročné gymnázium, Alejová 1, Košice

- ▶ sociálna sieť – orientovaný graf
- ▶ vrcholy – osoby
- ▶ (orientované) hrany – ich vzájomné vzťahy:
 - ▶ crisp verzia – klasický (orientovaný) graf
 - ▶ fuzzy verzia – graf je úplný, ale jeho hrany sú vážené
- ▶ oba prípady sú popísateľné štvorcovou maticou
 - ▶ crisp verzia – dve hodnoty
 - ▶ fuzzy verzia – viac hodnôt, obvykle z $[0, 1]$
- ▶ ideálna situácia pre relačno-dátovú analýzu

- ▶ **pozorovanie:**
ľudia s podobnými charakteristikami sú priatelia
(ak majú dostatok času a príležitostí spriatelíť sa)
v zhode s príslovím „Vrana k vrane sadá“
- ▶ **cieľ:**
získať informácie o priateľstvách medzi členmi sociálnej siete
- ▶ **vstup:**
graf vzájomných vzťahov medzi členmi
je popísaný jeho maticou
- ▶ **výstup:**
zhluky členov, ktorí sú podobne vnímaní ostatnými,
t. j. **skupiny priateľov**

- ▶ školská trieda (sekunda) s okolo 11-ročnými deťmi, ktoré sa navzájom poznajú aspoň 1 rok
- ▶ 34 spolužiakov – 23 chlapcov, 11 dievčat
- ▶ spoluautorka JK bola ich triednou učiteľkou
- ▶ každý študent vyjadril svoj vzťah ku každému inému spolužiakovi prostredníctvom nasledujúcej ankety

- ▶ 7 (lineárne usporiadaných) hodnôt:

hodnota	vysvetlenie
3	je to môj veľmi dobrý priateľ
2	je to môj priateľ
1	je mi sympatický
0	nie je mi ani sympatický, ani nesympatický
-1	je mi nesympatický
-2	nemám ho rád
-3	neznášam ho

- ▶ v zhode s Likertovou stupnicou

- ▶ **riadok** – hodnotený spolužiak,
- stĺpec** – hodnotiaci spolužiak:

	kód	1	2	3	4	5	6	7	8	9	10	...	34
1	Ch1	3	-2	1	1	3	0	-1	0	3	0	...	0
2	Ch2	1	3	3	0	0	0	3	3	0	1	...	2
3	Ch3	0	3	3	2	-1	-2	3	2	1	3	...	1
4	Ch4	-1	-1	0	3	-1	-2	-2	0	0	0	...	-2
5	Ch5	3	0	2	-1	3	0	-1	0	2	0	...	1
6	D6	-1	-1	-1	-1	-1	3	-3	0	-1	0	...	2
7	Ch7	0	3	3	-1	-2	0	3	3	-3	0	...	1
8	Ch8	1	2	3	1	1	0	2	3	1	1	...	2
9	Ch9	3	1	2	2	3	-1	-1	-1	3	1	...	1
10	Ch11	0	0	3	1	0	-2	0	2	1	3	...	1
...
34	D11	0	0	1	-2	0	2	-2	1	0	0	...	3

- ▶ **hodnoty** môžu byť transformované do $\{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1\} \subseteq [0, 1]$ funkciou $x \mapsto \frac{x+3}{6}$

- ▶ fuzzy kontext:
 - ▶ A – atribúty, $A \neq \emptyset$
 - ▶ B – objekty, $B \neq \emptyset$
 - ▶ R – fuzzy relácia na $A \times B$, t. j. $R : A \times B \rightarrow [0, 1]$
- ▶ R – tabuľka, A – stĺpce, B – riadky
- ▶ $R(a, b)$ – stupeň,
v akom má objekt b atribút a

- ▶ $\uparrow : \mathcal{P}(B) \rightarrow [0, 1]^A$:

$$(\uparrow(X))(a) = \inf\{R(a, b) : b \in X\}$$

t. j. $\uparrow(X)$ je fuzzy množina atribútov,
jej hodnota v stĺpci $a \in A$ je infimum hodnôt tohto stĺpca
v riadkoch z X

- ▶ $\downarrow : [0, 1]^A \rightarrow \mathcal{P}(B)$:

$$\downarrow(f) = \{b \in B : (\forall a \in A) R(a, b) \geq f(a)\}$$

t. j. $\downarrow(f)$ je množina objektov,
ktorých funkcie príslušnosti dominujú nad f

- ▶ $\langle \uparrow, \downarrow, \subseteq, \leq \rangle$ je **Galois-ova konexia**:
 - ▶ ak $X_1 \subseteq X_2$, tak $\uparrow(X_1) \geq \uparrow(X_2)$
 - ▶ ak $f_1 \leq f_2$, tak $\downarrow(f_1) \supseteq \downarrow(f_2)$
 - ▶ $X \subseteq \downarrow(\uparrow(X))$
 - ▶ $f \leq \uparrow(\downarrow(f))$
- ▶ alebo ekvivalentne

$$f \leq \uparrow(X), \quad \text{akk} \quad X \subseteq \downarrow(f)$$

- ▶ zobrazenie $\text{cl} : \mathcal{P}(B) \rightarrow \mathcal{P}(B)$

$$\text{cl}(X) = \downarrow(\uparrow(X))$$

- ▶ cl je **operátor uzáveru**:
 - ▶ $X \subseteq \text{cl}(X)$
 - ▶ ak $X_1 \subseteq X_2$, tak $\text{cl}(X_1) \subseteq \text{cl}(X_2)$
 - ▶ $\text{cl}(X) = \text{cl}(\text{cl}(X))$

- ▶ ak $X = \text{cl}(X)$,
tak dvojica $\langle X, \uparrow(X) \rangle$ sa nazýva **jednostranne fuzzy koncept**
- ▶ **jednostranne fuzzy**, lebo:
 - ▶ (extent) X je **crisp** množina objektov
 - ▶ (intent) $\uparrow(X)$ je **fuzzy** množina atribútov
- ▶ obe súradnice konceptu sú navzájom odvoditeľné;
stačí uvažovať prvú
- ▶ $\langle \{X \in \mathcal{P}(B) : X = \text{cl}(X)\}, \subseteq \rangle$ je (úplný) zväz,
ktorého operácie sú:
 - ▶ $X_1 \wedge X_2 = X_1 \cap X_2$
 - ▶ $X_1 \vee X_2 = \text{cl}(X_1 \cup X_2)$
- ▶ tento zväz sa nazýva **jednostranne fuzzy konceptový zväz**

Príliš veľa zhlukov na posudzovanie

- ▶ náš konceptový zväz má viac než 25 000 konceptov
 - nemožné na praktické použitie!
- ▶ potreba redukcie tohto počtu nejakou metódou:
 - ▶ Radim Bělohlávek et al.:
použitím nejakej **dodatočnej informácie** o objektoch alebo atribútoch,
napr. ich usporiadania alebo inej relácie
(ale: my nemáme žiadnu takú informáciu)
 - ▶ Radim Bělohlávek, Vilém Vychodil:
redukcia kardinality množiny hodnôt použitím
tzv. **zdôrazňovačov pravdy**
(ale: aj redukcia na 2 hodnoty vedie k viac než 2000 zhlukom)
 - ▶ použitím nejakých **metrických vlastností**
(napr. Rice-ovho a Siff-ovho algoritmu)

- ▶ $\rho : \mathcal{P}(B) \times \mathcal{P}(B) \rightarrow \mathbb{R}$:

$$\rho(X_1, X_2) = 1 - \frac{\sum_{a \in A} \min\{\uparrow(X_1)(a), \uparrow(X_2)(a)\}}{\sum_{a \in A} \max\{\uparrow(X_1)(a), \uparrow(X_2)(a)\}}$$

- ▶ je to **metrika** na množine $\{X \subseteq B : \text{cl}(X) = X\}$ všetkých extentov
- ▶ ρ je generalizácia Rice-om a Siff-om použitej funkcie vzdialenosti

$$\rho'(X_1, X_2) = 1 - \frac{|X_1 \cap X_2|}{|X_1 \cup X_2|}$$

Modifikovaný Rice-ov a Siff-ov algoritmus

▶ algoritmus:

vstup: A, B, R

$\mathcal{C} \leftarrow \mathcal{D} \leftarrow \{\text{cl}(\{b\}) : b \in B\}$

kým ($|\mathcal{D}| > 1$), rob:

$m \leftarrow \min\{\rho(X_1, X_2) : X_1, X_2 \in \mathcal{D}, X_1 \neq X_2\}$

$\mathcal{E} \leftarrow \{\langle X_1, X_2 \rangle \in \mathcal{D} \times \mathcal{D} : \rho(X_1, X_2) = m\}$

$\mathcal{V} \leftarrow \{X \in \mathcal{D} : (\exists Y \in \mathcal{D}) \langle X, Y \rangle \in \mathcal{E}\}$

$\mathcal{N} \leftarrow \{\text{cl}(X_1 \cup X_2) : \langle X_1, X_2 \rangle \in \mathcal{E}\}$

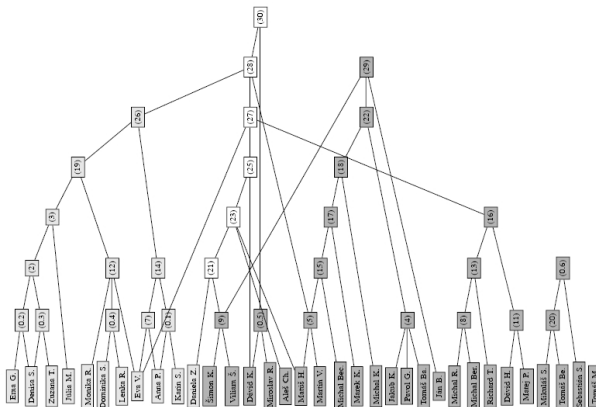
$\mathcal{D} \leftarrow (\mathcal{D} \setminus \mathcal{V}) \cup \mathcal{N}$

$\mathcal{C} \leftarrow \mathcal{C} \cup \mathcal{N}$

výstup: \mathcal{C}

- ▶ je teoreticky exponenciálny,
ale obvykle vracia okolo $2|B|$ zhlučkov

- ▶ väčšina zhlukov:



- ▶ svetlé – dievčenské zhluky, tmavé – chlapčenské zhluky
- ▶ menší zhluk – silnejšie priateľstvo

- ▶ zaujímavé pozorovania (**verifikované** triednym učiteľom):
 - ▶ dosť striktné delenie na chlapcov a dievčatá;
výnimka – jedno dievča s chlapčenskými záujmami
 - ▶ väčšina dvojíc sediacich v jednej lavici
 - ▶ kráska triedy (nie je na diagrame)
 - ▶ exhibicionista triedy
 - ▶ traja podceňovaní chlapci
- ▶ dnes je už možno situácia v triede iná...
- ▶ existujú aj nejaké iné zaujímavé skupiny, ktoré naša metóda nepodchytila?

- ▶ po našom experimente **sme presvedčení**, že (fuzzy) **formálna konceptová analýza** môže byť **užitočná**, ak je **kombinovaná** s nejakými **metrickými vlastnosťami** (množstvo zhlukov je vhodné na interpretovanie)
- ▶ táto metóda môže byť aplikovaná na ľubovoľnú relatívne uzavretú skupinu ľudí a môže pomôcť jej manažérovi lepšie **pochopiť štruktúru** manažovanej **sociálnej siete**

Ďakujeme za pozornosť

stanislav.krajci@upjs.sk
jana.krajciova@pobox.sk