

Odporúčanie vo viacrozmerných dátach

Diplomový projekt
Bc. Michal Oláh
vedúci Ing. Ján Suchal

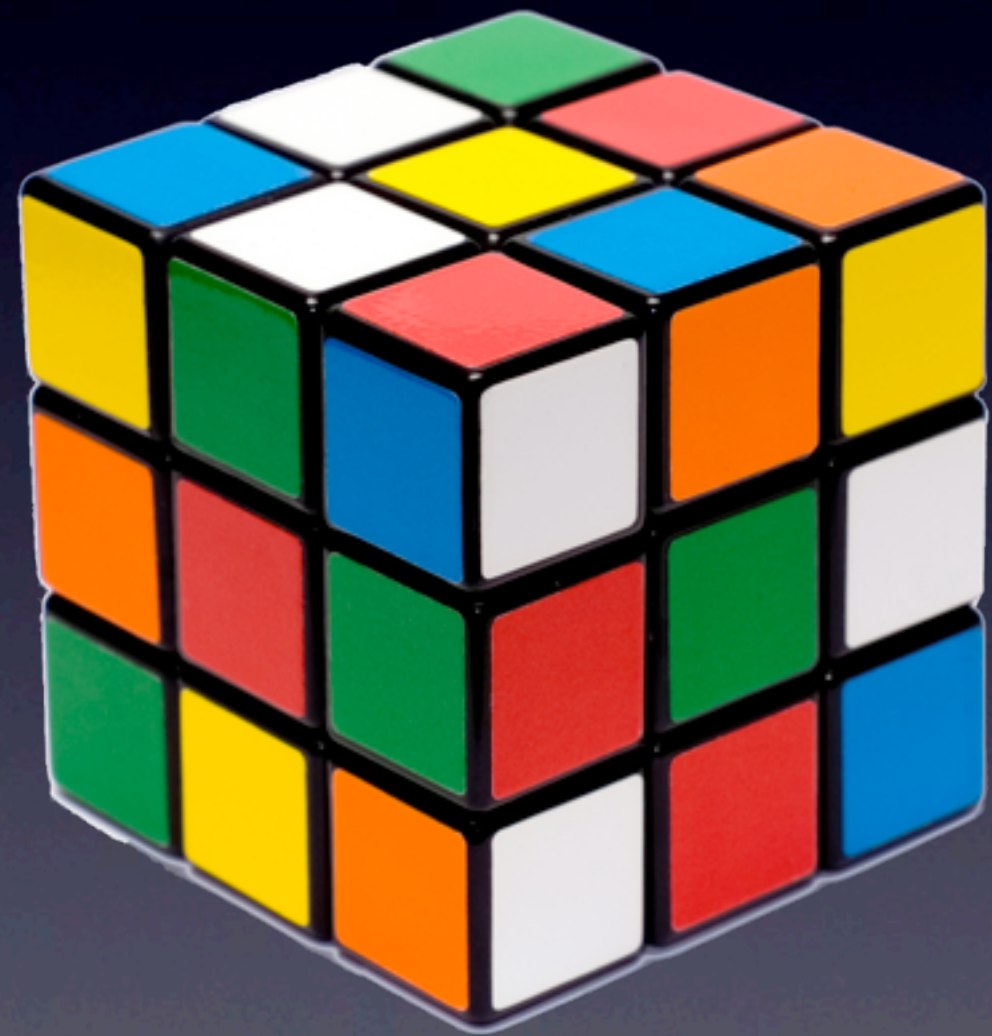
Odporúčanie

- mnohí z Vás sa mu venujú
- rôzne prístupy
 - HITS, PageRank etc.
- väčšina neberie do úvahy viac ako jeden typ prepojenia



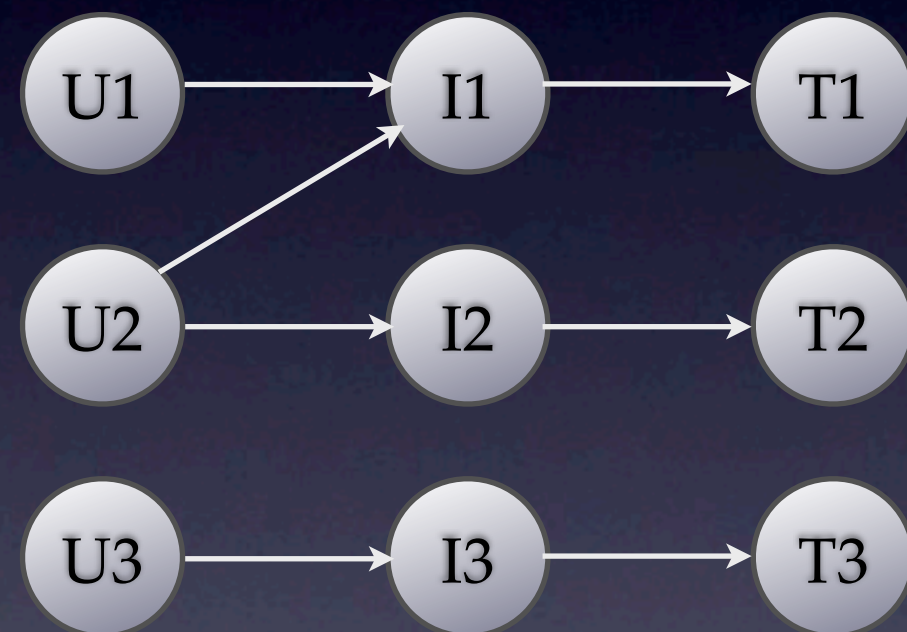
Multidimenzionálne odporúčanie

- Sleduje viac ako iba 1 aspekt
- Zvyčajne sa reprezentuje ako kocka (tenzor)
- Cieľ je využiť viac informácií pri odporúčaní



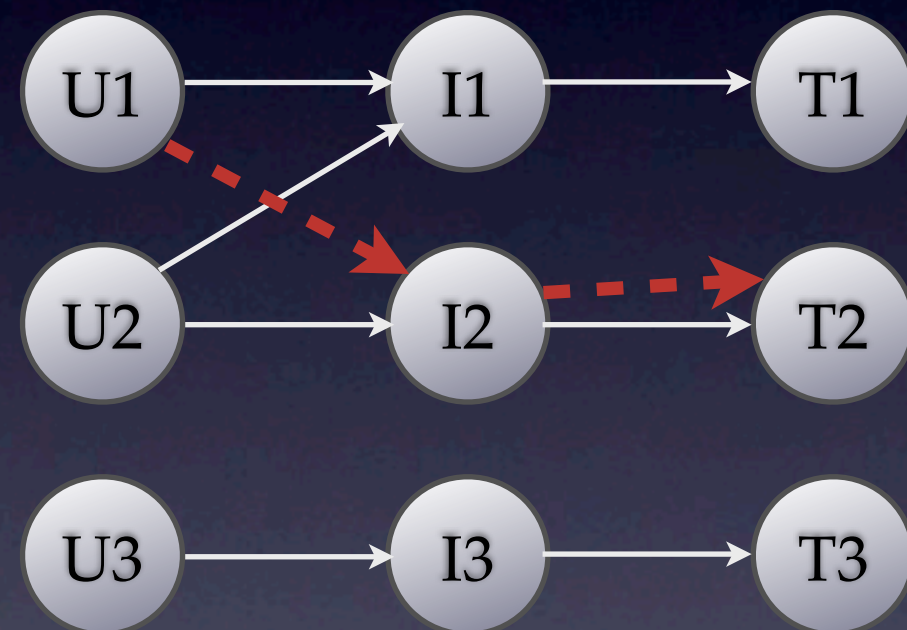
Príklad

- používatelia U1-U3
- predmety I1-I3
- značky T1-T3



Príklad

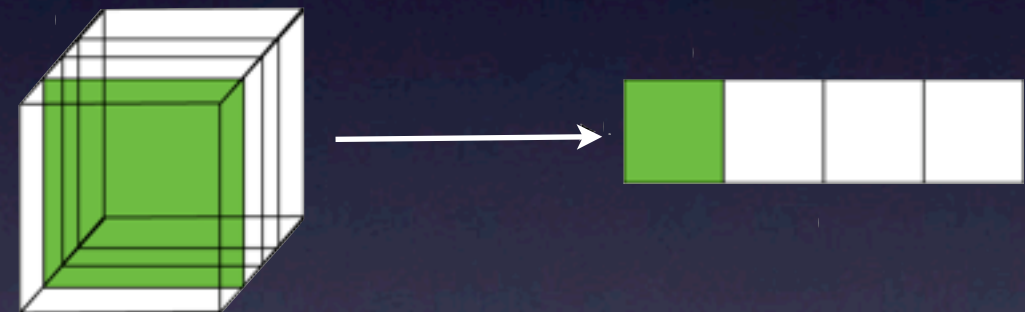
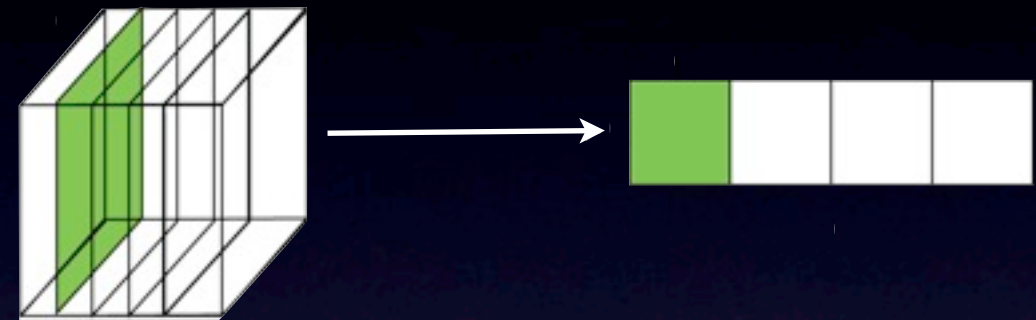
- používatelia U1-U3
- predmety I1-I3
- značky T1-T3
- nové prepojenie
 $U1 \Rightarrow I2 \Rightarrow T2$



Rozklady tenzorov

HOSVD

- použitá technika: SVD (singular value decomposition)
- rozklad na dvojrozmerné matice
- aplikácia SVD na každý rozmer
- sledovanie rozdielov medzi pôvodným a redukovaným tenzorom



Využitie

- odporúčanie pri vyhľadávaní
- sledovanie zmien (napr. v cene)
- všeobecné odporúčanie na základe viacrozmerných prepojení

Odporúčanie pri vyhľadávaní

- realitné dáta
- používateľ vyhľadáva (tag) podľa nejakých kritérií a nájde dobrý výsledok a pridá do bookmarkov (item). Tým vytvorí prepojenie
- nový používateľ keď vyhľadáva podobne (podobný tag) dostane odporúčanie na základe iných

Odporúčanie na základe meniacich sa atribútov

- cena
 - v čase sa mení
 - ak sa zmení cena na jednom prvku, a druhý v danej lokalite nie, tak to vieme identifikovať a odporučiť na základe toho

SME dáta

- je ich veľa
- sú riedke
- existuje odporúčač, s ktorým je možné porovnávať
- odporúčanie môže ísť cez autorov, sekciu, kategóriu alebo akýkoľvek iný atribút, ktorý spája články do ďalšej dimenzie...

GitHub

- súťaž, ako dobrá referencia pre presnosť
 - prepojenie napr. cez sledované repozitáre $\Leftrightarrow$ autor $\Leftrightarrow$ repozitár
- opäť veľké a riedke dáta

Ďalšia práca

- Použitie alternatívnych metód spracovania matíc s nižšími nárokmi na výpočtový výkon a pamäť