

Úvod k mnohोरukým banditům



Matej Končál a Andrej Zaťko

Vedúci: Ing. Peter Gašpar

Problém mnohorukého banditu

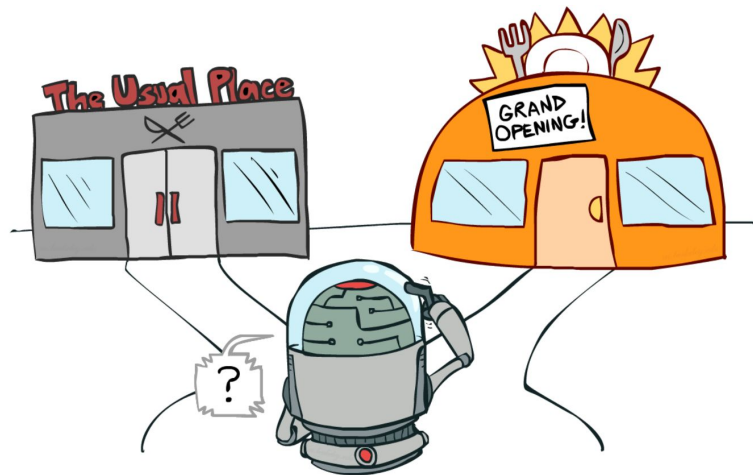
výber z N dostupných možností (automatov, pivovarov, ...)



maximalizovanie výhry (spokojnosti)



exploration vs exploitation



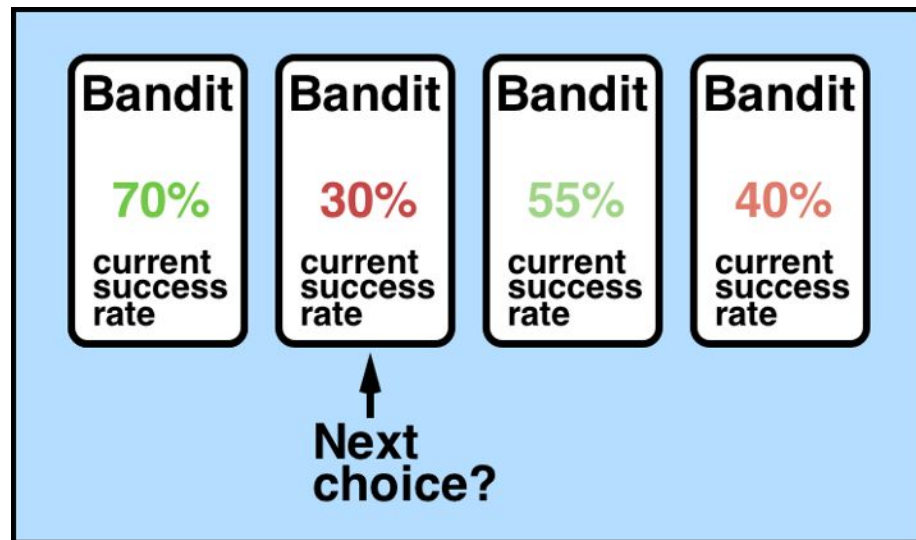
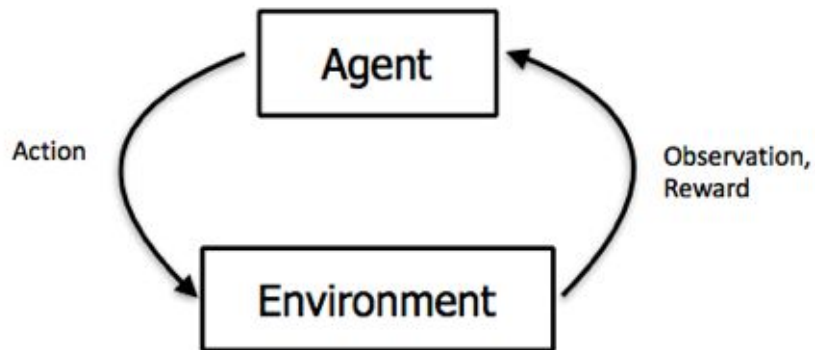
Ako vyhrať čo najviac?

10 x potiahnutie každej ruky

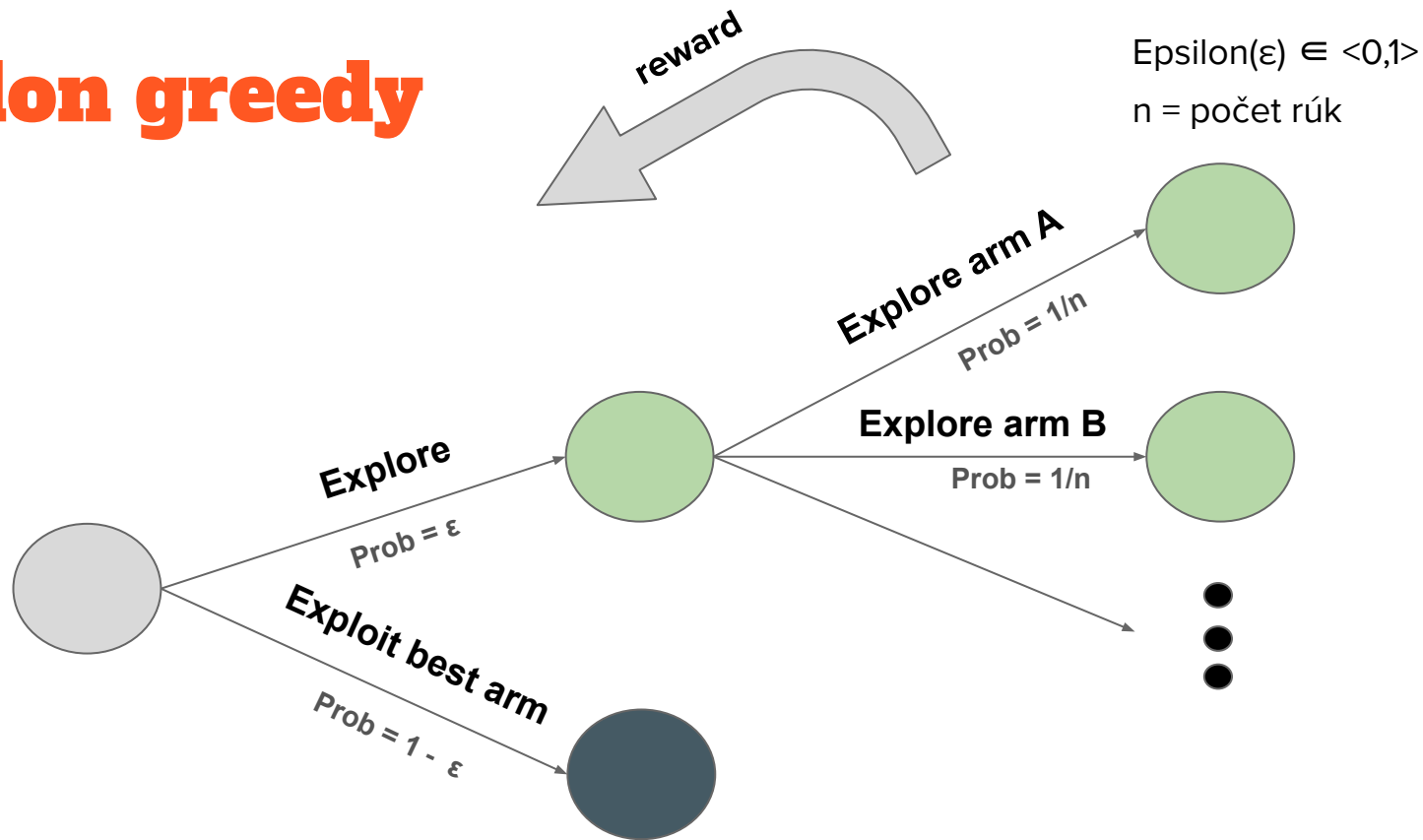
Exploit or Explore?



Učenie s odmenou a trestom

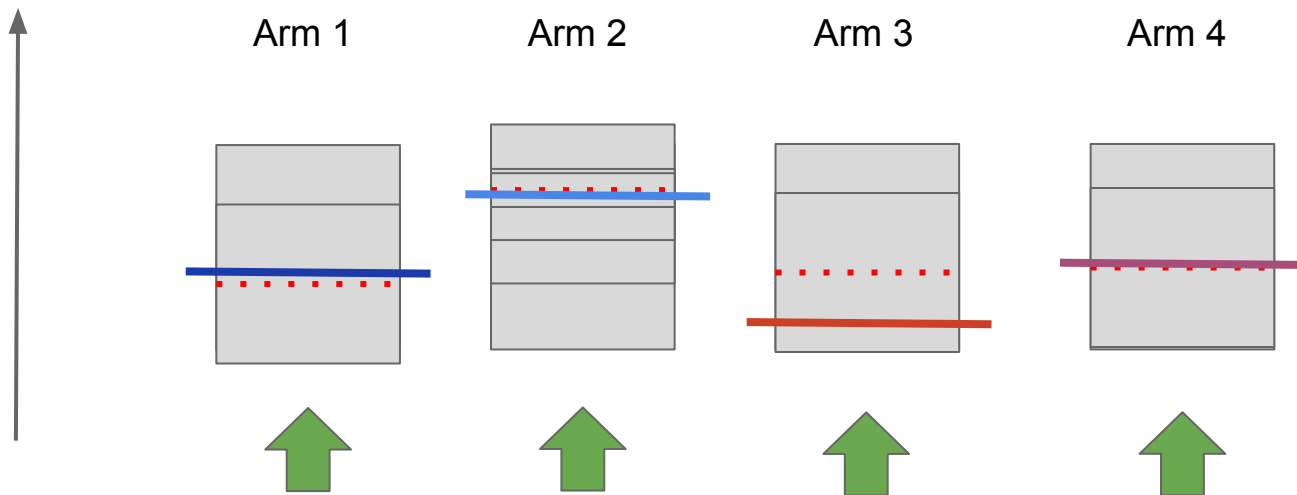


Epsilon greedy



Upper confidence bound

čím viac sme si neistí efektivitou ruky, tým je potrebné preskúmať ju



Zohl'adnenie stavu (kontextu)

LinUCB[1], Thompson Sampling, ...



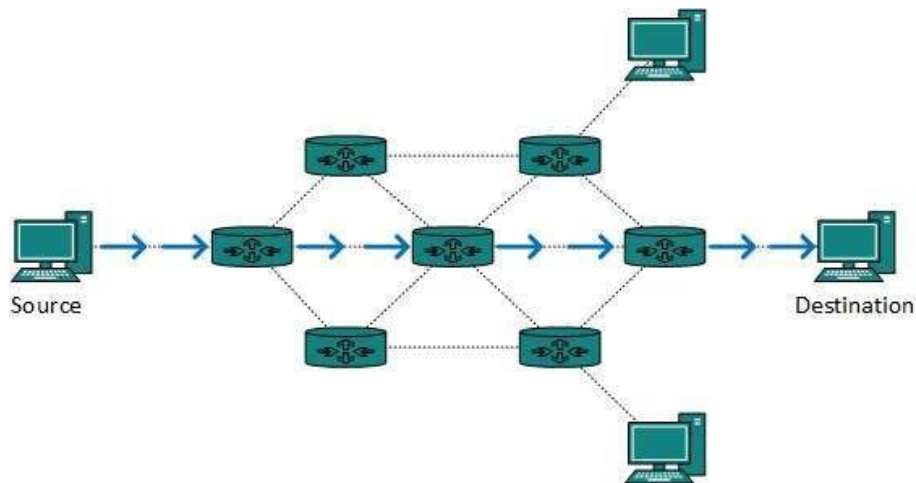
Multi-armed Bandit



Contextual Bandit

Použitie v praxi

- odporúčanie
- smerovanie v sieti
- medicína
- alternatíva k A/B testingu



A/B testovanie



Welcome to our website

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

[Learn more](#)

Click rate: 52 %



Welcome to our website

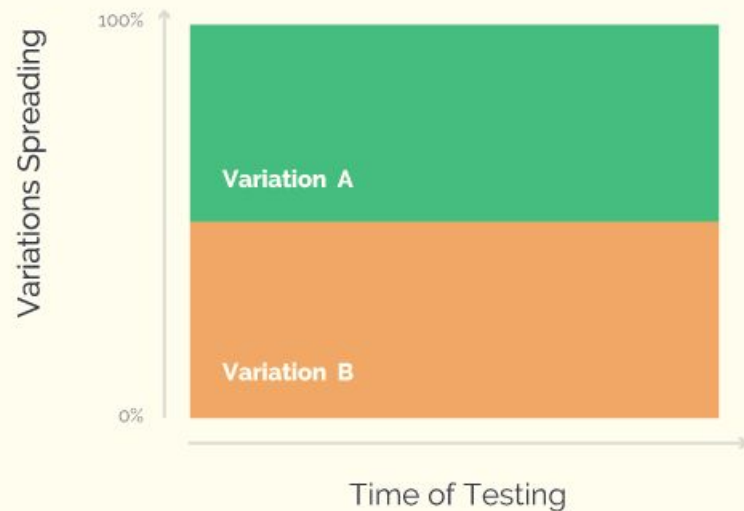
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

[→ Learn more](#)

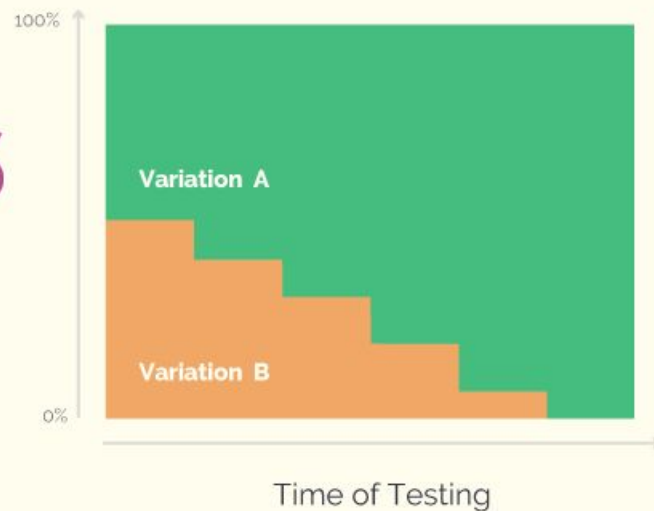
72 %

A/B testovanie

Mnohoruký bandit



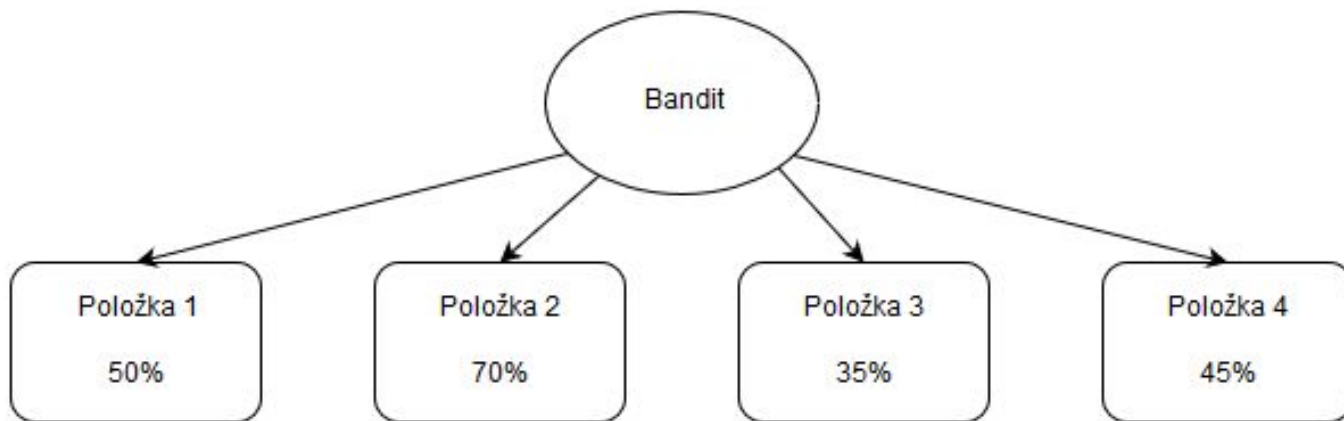
VS



Banditi v odporúčaní

Zohľadňujú

- frekventované pridávanie a odoberanie položiek
- príchod nový používateľa
- zmeny preferencií používateľov v čase

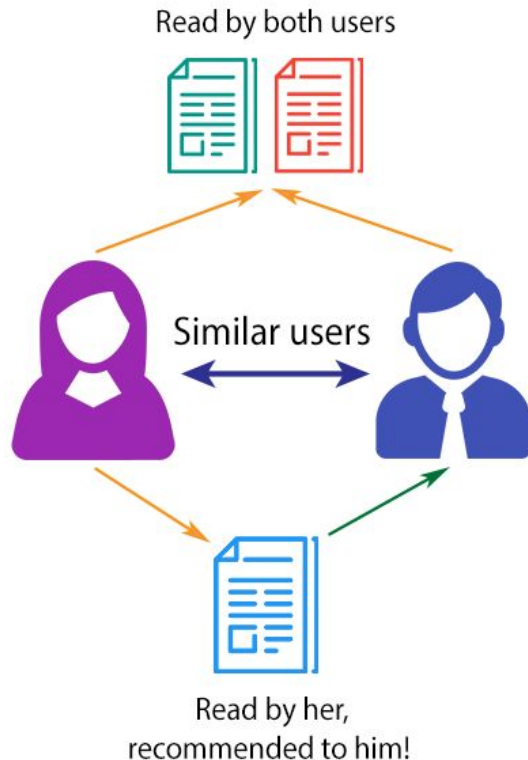


Kolaboratívni banditi

Prideľovanie odmeny - používatelia [2]

- jeden bandit pre všetkých používateľov
- bandit pre každého používateľa
- bandit pre zhľuky používateľov

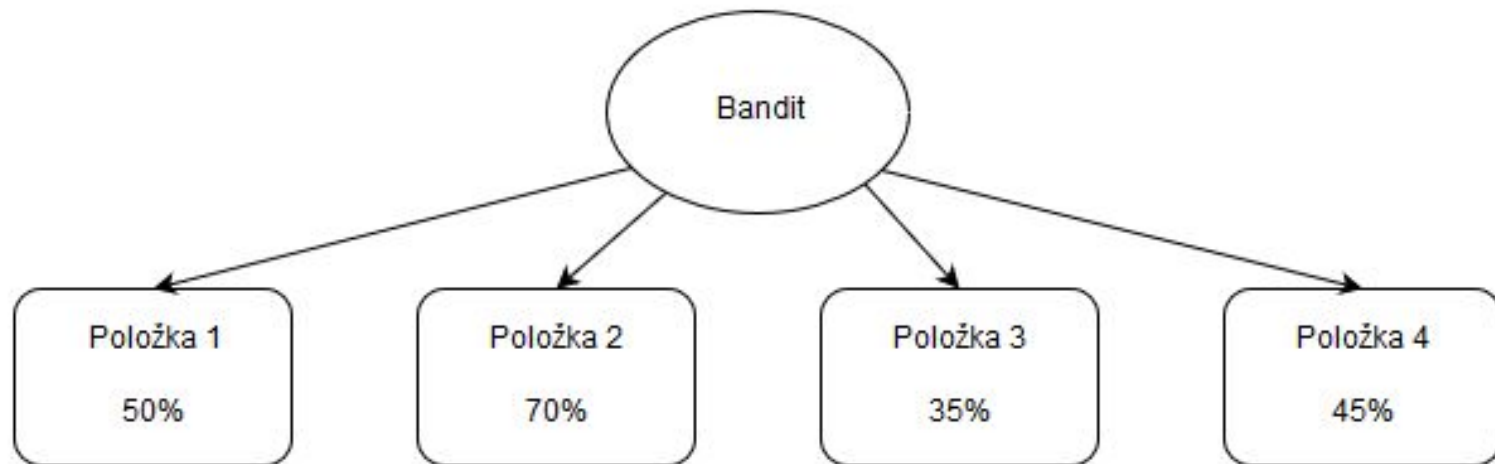
COLLABORATIVE FILTERING



Jeden bandit pre všetkých používateľov

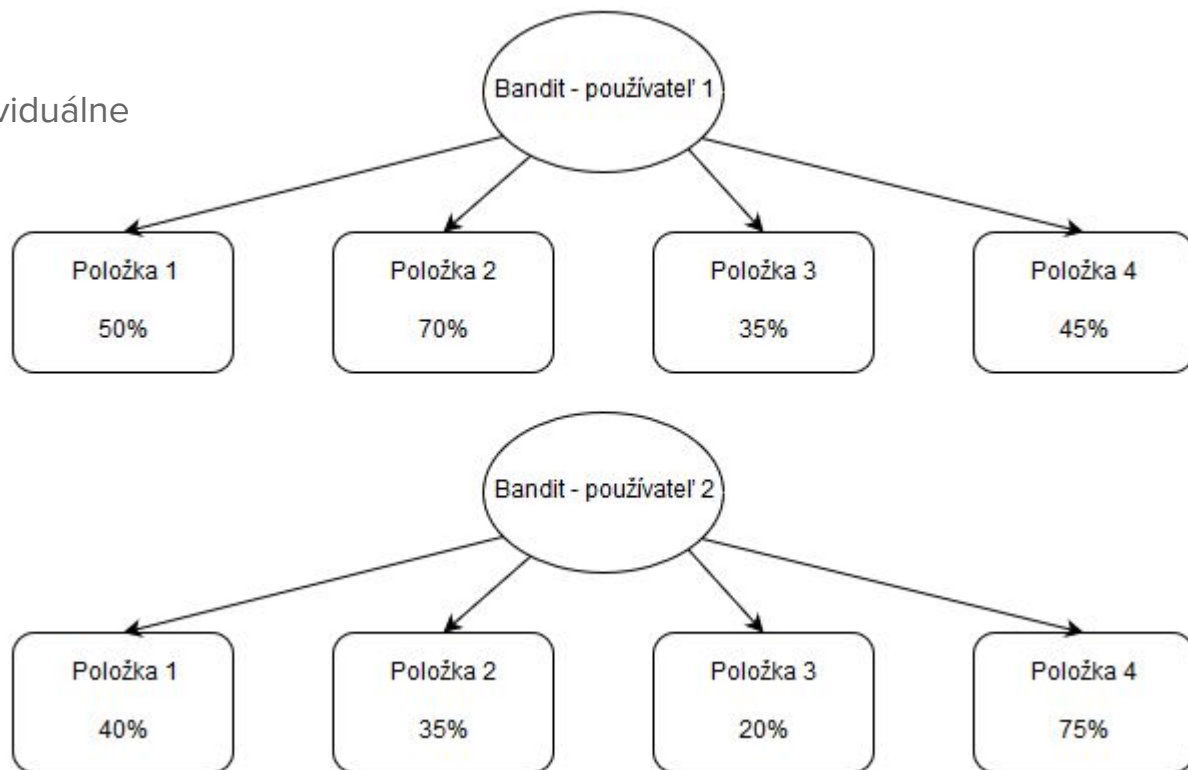
+ veľa dát

- preferencie všetkých používateľov však nie sú rovnaké



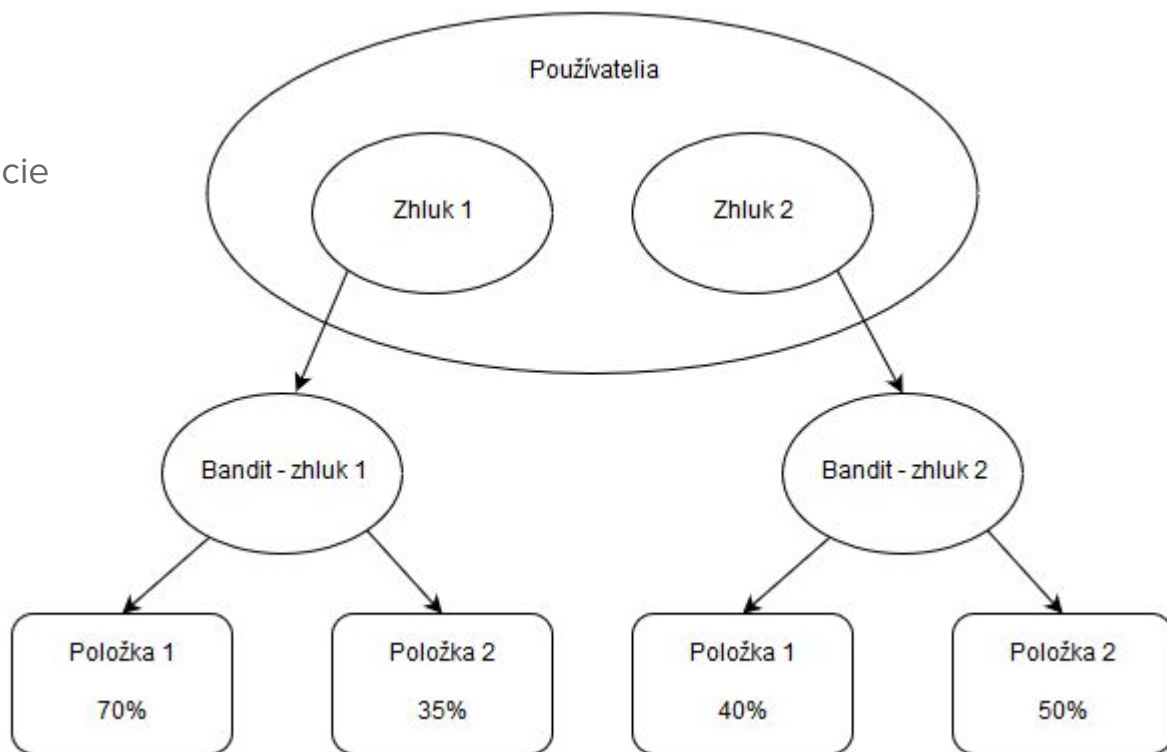
Bandit pre každého používateľa

- + personalizované
- preferencie nie sú až tak individuálne
- málo dát



Bandit pre zhluk používateľov

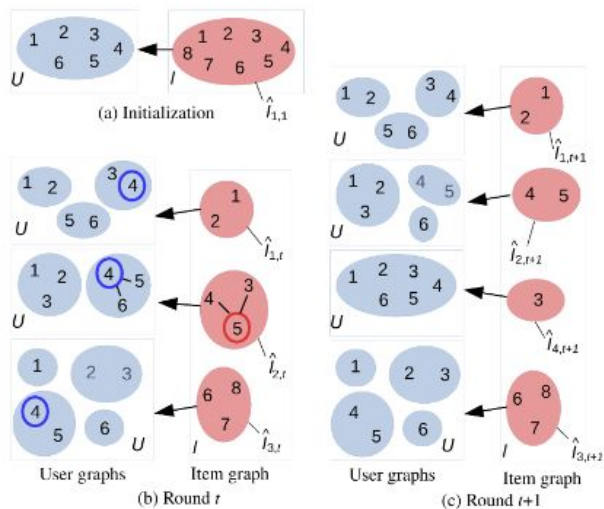
- dostatočné množstvo dát
- dostatočne personalizované
- používatelia zdieľajú preferencie



Kolaboratívni banditi

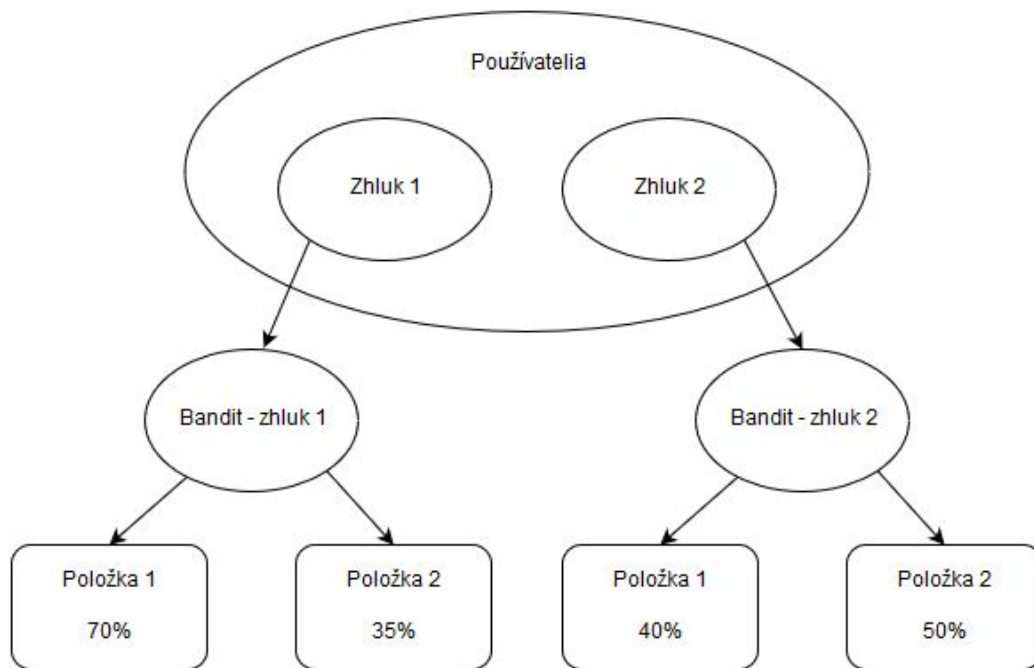
Pridelovanie odmeny - položky [3]

- jedna položka
- podľa podobnosti položiek

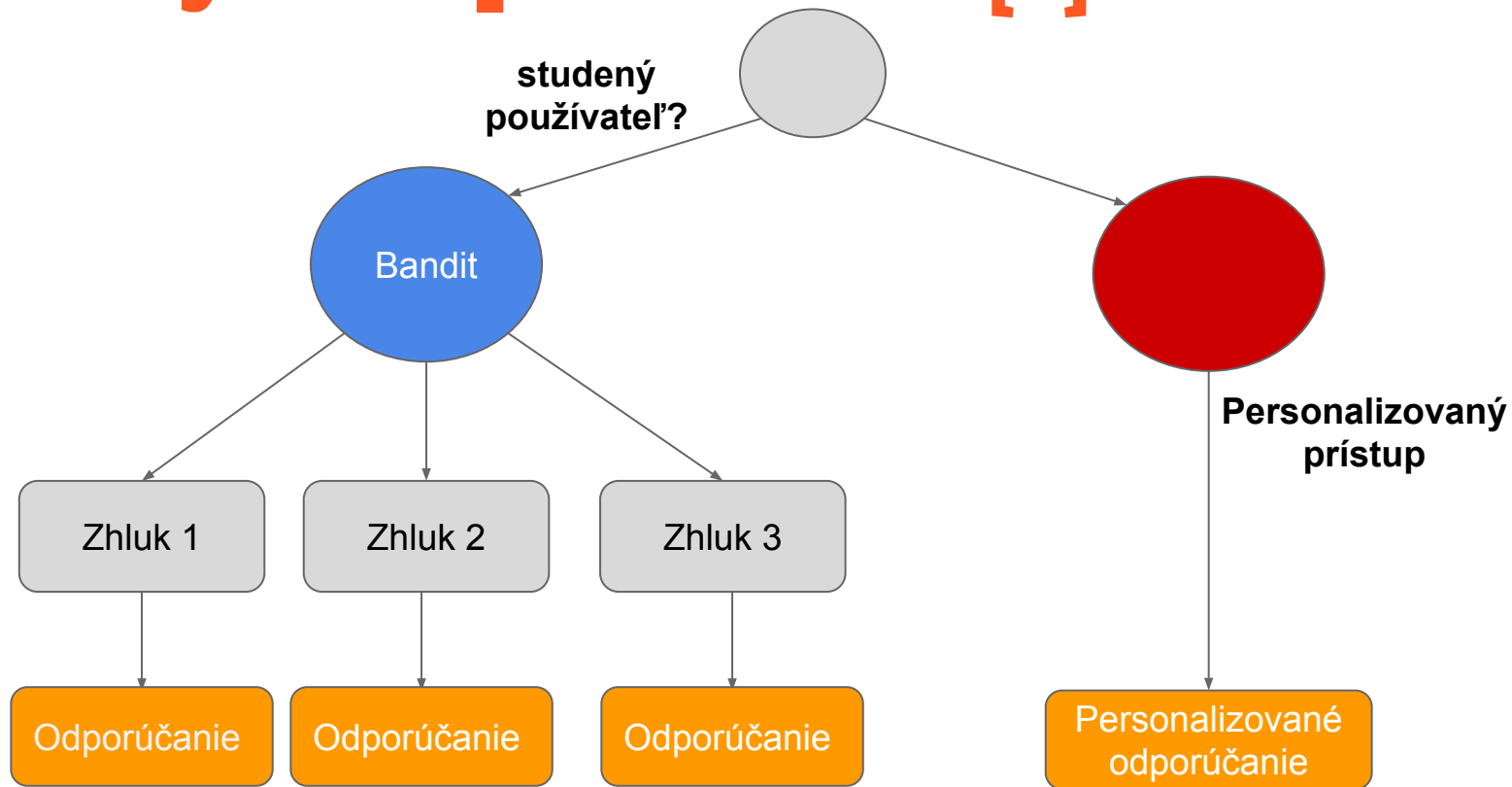


My Recipes		
Summer Recipes		
	Crunchy Croissants Buttery, flaky pastry named for its crescent shape.	01
	Grilled Eggplant Tender with a rich, complex flavor when cooked.	02
	Tangerine Salad Peak tangerine season lasts from autumn to spring.	03
	Pomegranate Juice A superfood for good skin, hair, and health	04
	Chili Salsa Heat up any taco night with minimal preparation	05
	Flour from scratch Grind grains, nuts, or beans in your coffeegrinder.	06

Pridel'ovanie odmeny - kombinácia [4]



Studený štart používateľa [5]



Mnohorukí banditi

výber z N dostupných možností (automatov, pivovarov, ...)

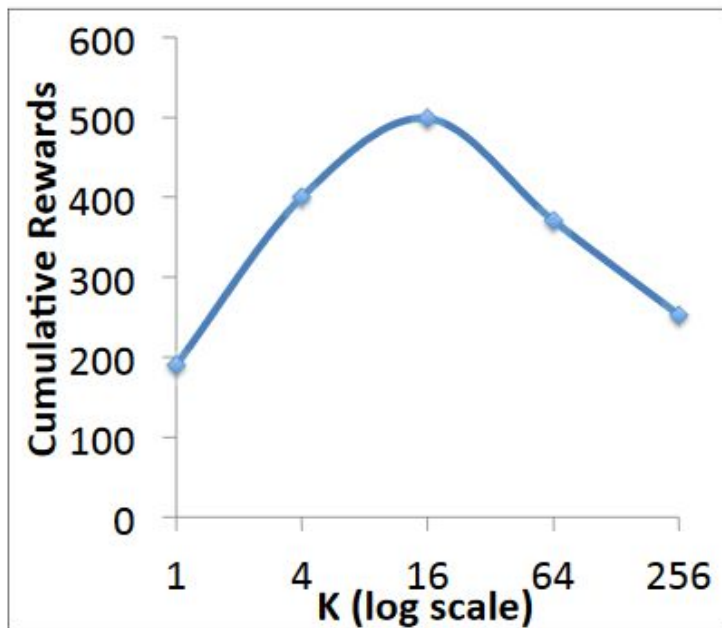


maximalizovanie výhry (spokojnosti)

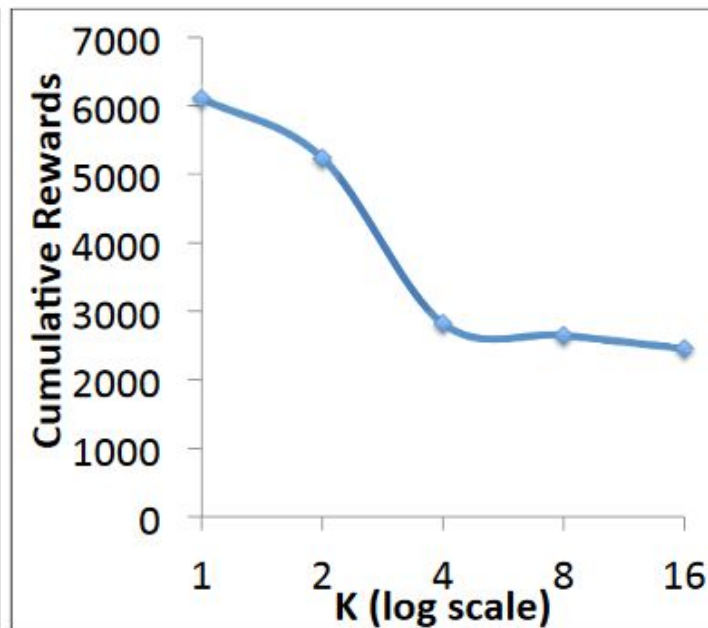


Diskusia

Závislé od domény [2]



a. Delicious



b. LastFM

Zdroje

- [1] <https://dl.acm.org/citation.cfm?id=1772758>
- [2] <https://dl.acm.org/citation.cfm?id=2662063>
- [3] <https://arxiv.org/abs/1502.03473>
- [4] <https://epubs.siam.org/doi/abs/10.1137/1.9781611975321.69>
- [5] <https://dl.acm.org/citation.cfm?id=3079681>