Seeker: Real Time Interactive Search

Ari Biswas, Thai T Pham, Michael Vogelsong, Benjamin Snyder, and Houssam Nassif

amazon

Problem Statement

Seeker: allows users to adaptively refine search rankings in real time, through feedbacks in the form of likes and dislikes.

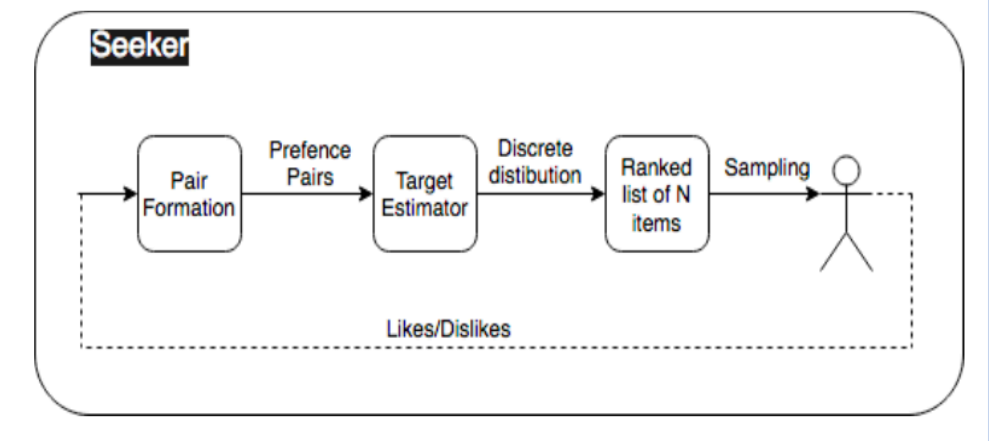
User has a mental picture of desired item and can answer ordinal questions like: *"Is this item similar to what you have in mind?"*

Problem formulation:

- User has in mind target item t.
- Catalog has N items.
- Display $M << N$ items at a time.
- User interacts with items through likes and dislikes.
- User can interact with same item repeatedly.

Aim: Surface item t in M using the fewest interactions possible.

Logical Components



Pair Formation

Definitions/assumptions:

- Represent item i as embedding feature vector x_i.
- Assume high correlation between human similarity and distance metric in embedding space.
- Preference pair s_{ij} consists of liked item a_i and disliked item b_i.
- User prefers i over j given t: $||x_i - x_t||^2 < ||x_j - x_t||^2$

Probability of choosing i over j:
$P(s_{ij}|t) = logistic(||x_j-x_t||^2-||x_i-x_t||^2)$

$P(s_{ij}|t)=0.5$ if i & j equidistant from t.
$P(s_{ij}|t)=1$ if $i==t$ and j infinitely far.

Target Estimation

Assume all preference pairs are independent from each other.

Likelihood of preference set S:

$$P(S|t) = \prod_{s_{ij}} P(s_{ij}|t)$$

Posterior of an item being the target:

$$g_t = P(t) \prod_{s_{ij}} P(s_{ij}|t)$$

Item Ranking

Generate discrete distribution using softmax (Boltzmann) equation:

$$p_j = \frac{e^{g_j}}{\sum_i e^{g_i}}$$

Sample by Gumbel-softmax trick:

$$argmax_j\{g_j + Gumbel(0,1)\} \sim \frac{e^{g_j}}{\sum_i e^{g_i}} = p_j$$

Anneal using inverse temperature η_k to balance explore/exploit:

$$argmax_j\{\eta_k g_j + Gumbel(0,1)\} = argmax_j\left\{g_j + \frac{Gumbel(0,1)}{\eta_k}\right\} \sim \frac{e^{\eta_k g_j}}{\sum_i e^{\eta_k g_i}} = p_j$$

This is similar to Thompson Sampling bandit, but doesn't account for estimates uncertainty. We include estimate uncertainty thru annealing:

$$argmax_j\left\{g_j + \sqrt{\frac{C^2}{n_j}} Gumbel(0,1)\right\}$$

where n_j is the number of times a user interacts with item j. C depends on the variance factor of the reward subgaussian distribution. In our case, reward is binary and $C^2 = \frac{1}{8}$.

Algorithm Summary

Input: $M \leq N, n_i \forall i \in N$
for all $x_i, i = 1, \cdots, N$ **do**
$\gamma_i \sim \text{Gumbel}(0, 1)$
$z_i = g_i + \frac{C\gamma_i}{\sqrt{n_i}}$
results = sort($\{ z_i \mid i = 1, \cdots, N \}$) in descending order
return top M results

Experimental Results

- Compare Seeker to pure exploitation, random, and epsilon-greedy bandit.
- Interact with system up to K interactions or until $t \in M$.
- Set K=15, M=12, N=2228.
- Monitor target normalized rank ρ.
- Recall@ρ_j: Percentage of sessions with $\rho_i \leq \rho_j$.