Learning to Top-K Search Using Pairwise Comparisons

Brian Eriksson
brian.eriksson@technicolor.com

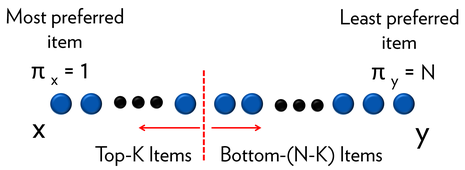
technicolor

Learning to Rank Problem Setup

N Unordered Items

$\pi : \{1,2,\ldots,N\} \rightarrow \{1,2,\ldots,N\}$

Unknown underlying preference mapping



Goal: Find the Top-K preferred items

Potential Applications:

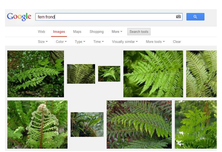
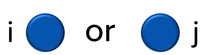


Image-Based Search

Recommender Systems

Papers Submitted to a Conference

How? We focus on pairwise comparisons :



Which item is more preferred?

$$c_{i,j} = \begin{cases} 1 & \pi_i < \pi_j \\ 0 & \text{otherwise} \end{cases}$$

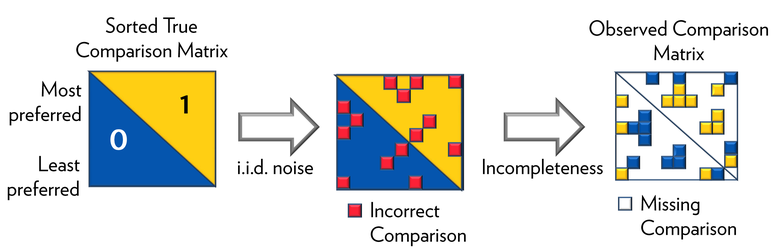
Pairwise comparison matrix, C

Limitations:

1. Inconsistency : Not all the pairwise comparisons will be correct.

   Noise Model: $P\left(c_{i,j} \neq 1\left(\pi_i < \pi_j\right)\right) \leq q$ Where : $q \geq 0$

2. Incompleteness : Not all $O(N^2)$ pairwise comparisons will be available.
   - User query limitations – e.g., paper reviews are time consuming
   - Computation time limitations – e.g., SIFT Feature Comparisons

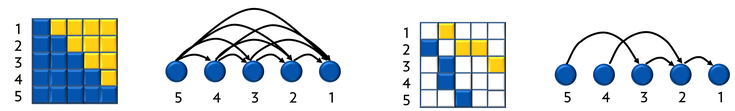


Top-K Search using Randomly Chosen Comparisons

Consider, the noise-free case (q=0) where the pairwise comparisons are observed at-random with:

$$P(c_{i,j} \text{ is observed}) = p \qquad \text{Where : } p > 0$$

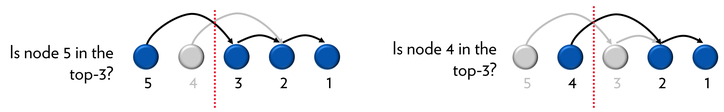
We construct the directed *Sampling Comparison Graph*, where each item is a vertex and there is an edge between two vertices if we observe the pairwise comparison (i.e., $e_{i,j}=1$ if we observe $c_{i,j} = 0$)



Complete Comparison Matrix — Incomplete Comparison Matrix

Methodology: PathRank - To find the Top-K items, eliminate all items that begin observed paths of length > K.

Example: Top-3 Search using an incomplete comparison matrix

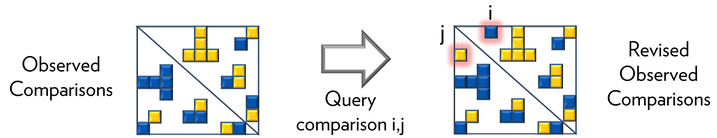


No, there exists a path of length 4. — Possibly, as there exists no path of length >3.

Theorem 1: Given N items with unknown underlying ordering, π, and i.i.d. pairwise comparison observation probability, p>0, then with probability ≥ (1-α) (where α >0), the PathRank algorithm will only return items from the top-(2K(1-p)/p + 2log(N/α))

Top-K Search using Adaptively Chosen Comparisons

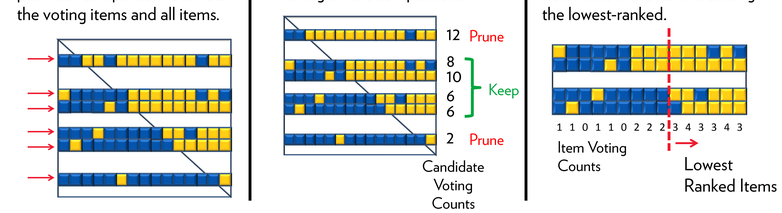
Now consider *potentially erroneous* pairwise comparisons that can be adaptively queried.



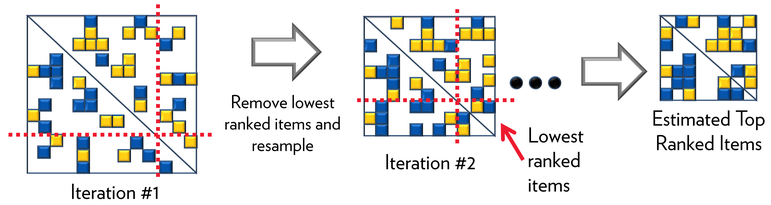
Known results : In the noise-free case (q=0), the Top-K search problem can be solved using a bisection search.

Then, the Top-K items can be found *exactly* using ≤ Nlog(k+1) pairwise comparisons.
-But, comparison errors defeat this methodology.

Using adaptively-chosen comparisons and a two-stage voting methodology, we can eliminate at least N/8 of the lowest ranked items using only O(N log(N)) pairwise comparisons.



This procedure is repeatedly performed to reduce to a top-ranked subset.



Theorem 2: Given N items with unknown underlying ordering, π, and i.i.d. comparison error probability, q ≥ 0, is not too large, then with probability ≥ (1-α) (where α > 0), the iterative robust algorithm will only return items from the top-O(log(N)) using only $O(N \log^2(N))$ adaptively-chosen pairwise comparisons.

Experiments:

The iterative adaptive procedure is run until a subset of <50 top-ranked items are returned.

Number of items	Prob. of Incorrect Comparisons (q)	Fraction of Total Comparisons Used	Lowest Ranked Item Returned
10,000	0.10	3.66×10^{-2}	36.31
1,000,000	0.10	5.53×10^{-4}	36.14
10,000	0.40	3.69×10^{-2}	117.21
1,000,000	0.40	5.56×10^{-4}	101.26

Results averaged across 100 realizations