Preference-Based Rank Elicitation using Statistical Models: The Case of Mallows

Robert Busa-Fekete [1] Eyke Huellermeier [2] Balzs Szrnyi [3]

[1]MTA-SZTE Research Group on Artificial Intelligence, Tisza Lajos krt. 103., H-6720 Szeged, Hungary

[2]Department of Computer Science, University of Paderborn, Warburger Str. 100, 33098 Paderborn, Germany

[3]INRIA Lille - Nord Europe, SequeL project, 40 avenue Halley, 59650 Villeneuve dAscq, France

ICML 2014
June 24, 2014

Outline

1. Preference-based stochastic multi-armed bandit (PB-MAB)
2. PB-MAB Framework with Statistical Models over Ranking
3. Concrete implementations with Mallows model
4. Numerical Experiments

Motivation

- Exploiting revealed preferences to learn a ranking
- Example: Crowdsourcing
  - Amazon Mechanical Turk
  - Widely-used platform in Natural Language Processing (NLP) to annotate training database
  - Machine Translation: for each English sentence, there are given many possible translations
  - Goal: either to find a ranking which reflects to the quality of the translations, or the best translation
  - The annotators might be asked in terms of simple questions: Is translation A better than translation B?

Preference-based stochastic multi-armed bandit setup

- Give M items/arms: $\mathcal{A} = \{a_1, ..., a_M\}$
- Items/arms can only be compared in a pairwise manner
- In a time step t, the (online) learning algorithm selects a pair of items (i^t, j^t) to be compared $\Rightarrow$ feedback
- Pairwise probability for items a_i and a_j :

$$p_{i,j} = \mathbb{P}(a_i \succ a_j) = \mathbb{E}[\mathbb{I}\{a_i \succ a_j\}] \tag{1}$$

  follows a fixed probabilistic distribution
- If $p_{i,j} > 1/2$, then item a_i is preferred to item a_j
- If $p_{i,j} < 1/2$, then item a_j is preferred to item a_i

Goal of the online learner (decision maker/agent)

- Find the best item (with high probability)
- Find a ranking over item (with high probability)
- Minimize the number of pairwise comparisons

Modeling assumptions

- Elicit a ranking based on probabilistic (noisy) feedback
- Establish a connection to statistical models of rank data
- Full ranking
  - $(r_1, \ldots\ldots, r_i, \ldots\ldots, r_j, \ldots\ldots, r_M) \sim \mathbb{P}(.|\wp)$
- Observation
  - $\mathbb{I}\{r_i < r_j\}$
- $\mathbb{P}(.|\wp)$ is a parametric probability distribution over the set of ranking $\mathbb{S}_M$
- Making inference about P based on sampled pairwise comparisons

Modeling assumptions

- Pairwise probabilities can be written as

$$p_{i,j} = \mathbb{P}(a_i \succ a_j) = \sum_{r \in \mathcal{L}(r_j > r_i)} \mathbb{P}(r|\wp) \tag{2}$$

where $\mathcal{L}(r_j > r_i) = \{r \in \mathbb{S}_M | r_j > r_i\}$

- Implicitly, we assume certain regularity properites on P induced by $\mathbb{P}(.|\wp)$
- Pairwise probability cannot be arbitrary

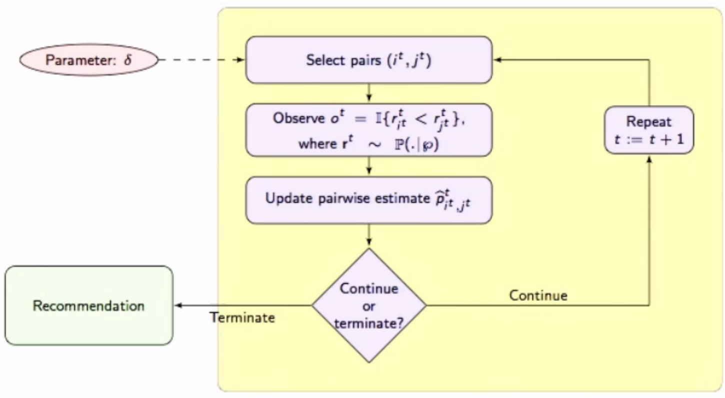
Modeling assumptions

Strong Stochastic Transitivity

For any triplet of arms such that $a_i \succ a_j \succ a_k$,
$p_{i,k} \geq max(p_{i,j}, p_{j,k})$

- If we know $a_i \succ a_j$ and $a_j \succ a_k$ then $a_i \succ a_k$
- Nice regularity property
- Reduce sample complexity

Online learning framework



How to make the setup complete

- **(MPI)** Find the most preferred item (or arm) a_{i^*} defined as

$$i^* = \operatorname*{arg\,max}_{1 \le i \le M} \mathop{\mathbb{E}}_{r \sim \mathbb{P}(.|\wp)} [\mathbb{I}\{r_i = 1\}] \tag{3}$$

- **(MPR)** Find the most probable ranking r^* given the ranking model $\mathbb{P}(.|\wp)$

$$r^* = \operatorname*{arg\,max}_{r \in \mathbb{S}_M} \mathbb{P}(r|\wp) \tag{4}$$

- All goals are meant to be achieved with probability at least $1 - \delta$
- Based as few pairwise comparisons as possible

Mallows ϕ-model

- The probability of observing a ranking r is

$$\mathbb{P}(r|\phi, \tilde{r}) = \frac{1}{Z(\phi)} \phi^{d(r,\tilde{r})} \tag{5}$$

- $\tilde{r} = (\tilde{r}_1, ..., \tilde{r}_M)$ is the center ranking
- $d(.,.)$ is the Kendall's rank distance defined as

$$d(r, \tilde{r}) = \sum_{1 \leq i < j \leq M} \mathbb{I}\{(r_i - r_j)(\tilde{r}_i - \tilde{r}_j) < 0\} \tag{6}$$

- $\phi \in (0, 1]$ is the spread parameter
  - $\phi = 1 \Rightarrow$ uniform distribution
  - $\phi \to 0 \Rightarrow \mathbb{P}(\tilde{r}|\phi, \tilde{r}) \to 1$ (becomes more peaky)
- $Z(\phi)$ is the normalisation factor

Find the most preferred item for Mallows model (**MPI**)

$$i^* = \operatorname*{arg\,max}_{1 \le i \le M} \mathop{\mathbb{E}}_{r \sim \mathbb{P}(.|\phi, \tilde{r})} [\mathbb{I}\{r_i = 1\}] \tag{7}$$

- Most preferred item (MPI) is the one for which $\tilde{r}_{i^*} = 1$
- The center ranking determines a total order on the set of items such that if $\tilde{r}_i < \tilde{r}_j$ then $p_{i,j} > 1/2$, and $\tilde{r}_i > \tilde{r}_j$ then $p_{i,j} < 1/2$

Find the most preferred item for Mallows model (**MPI**)

$$i^* = \arg\max_{1 \le i \le M} \mathop{\mathbb{E}}_{r \sim \mathbb{P}(.|\phi,\tilde{r})} [\mathbb{I}\{r_i = 1\}] \tag{7}$$

- Most preferred item (MPI) is the one for which $\tilde{r}_{i^*} = 1$
- The center ranking determines a total order on the set of items such that if $\tilde{r}_i < \tilde{r}_j$ then $p_{i,j} > 1/2$, and $\tilde{r}_i > \tilde{r}_j$ then $p_{i,j} < 1/2$
- **MALLOWSMPI**(δ)
  1. Pick a random item a_i
  2. Pick another item, say a_j, which has not selected yet, if there is no such, then break
  3. Compare a_i and a_j until $1/2 \notin [\hat{p}_{i,j} - c_{i,j}, \hat{p}_{i,j} + c_{i,j}]$, where

$$c_{i,j} = \sqrt{\frac{1}{2n_{i,j}} log \frac{4n_{i,j}^2 M}{\delta}} \tag{8}$$

  4. if $1/2 < \hat{p}_{i,j} - c_{i,j}$, then keep a_i goto 2, otherwise keep a_j

Find the most probable ranking for Mallows model (**MPR**)

- Most probable ranking is the center ranking

$$\tilde{r} = \underset{r \in \mathbb{S}_M}{\arg\max}\, \mathbb{P}(r|\phi, \tilde{r}) = \underset{r \in \mathbb{S}_M}{\arg\max}\, \frac{1}{Z(\phi)} \phi^{d(r,\tilde{r})} \quad (9)$$

- The center ranking determines a total order on the set of items such that if $\tilde{r}_i < \tilde{r}_j$ then $p_{i,j} > 1/2$, and $\tilde{r}_i > \tilde{r}_j$ then $p_{i,j} < 1/2$

Find the most probable ranking for Mallows model (**MPR**)

- **MALLOWSMPR**(δ)
  - Follow the Merge sort strategy
  - If the sorting algorithm compares two distinct items, say a_i and a_j, then compare them until $1/2 \notin [\hat{p}_{i,j} - c_{i,j}, \hat{p}_{i,j} + c_{i,j}]$, where

$$c_{i,j} = \sqrt{\frac{1}{2n_{i,j}} log \frac{4n_{i,j}^2 C_M}{\delta}} \tag{10}$$

  - Worst case performance of merge sort algorithm: $C_M = \lceil M \log_2 M - 0.91392 \cdot M + 1 \rceil$ (Theorem 1, Flajolet & Golin (1994))

Sample complexity

- **MallowsMPI**(δ)

$$\mathcal{O}(\frac{M}{\rho^2}\log\frac{M}{\delta\rho}), \tag{11}$$

where $\rho = \frac{1-\phi}{1+\phi}$

  - $\lim_{\phi\to 0} 1/\rho^2 = 1$ (more peaky distribution $\Rightarrow$ easier task)
  - $\lim_{\phi\to 1} 1/\rho^2 = \infty$ (more uniform $\Rightarrow$ harder task)

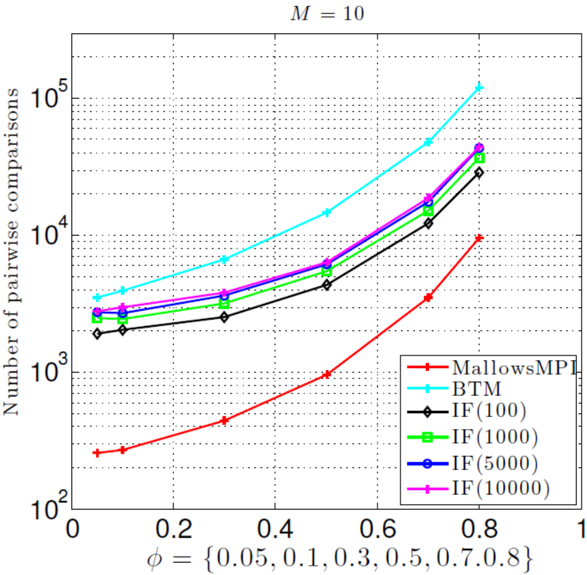
- **MALLOWSMPR**(δ)

$$\mathcal{O}(\frac{M\log_2 M}{\rho^2}\log\frac{M\log_2 M}{\delta\rho}) \tag{12}$$

Numerical experiments for identifying **Most Preferred Item (MPI)**

- Verify that if the model assumptions are valid, the purposed algorithm is efficient.
- If Mallows model is assumed $\Rightarrow$ best arm $=$ most preferred item
  - Beat the mean (in PAC setting) [Yue and Joachims, 2011]
  - Interleaved Filter [Yue et al., 2012]
- Compare the algorithms in terms of sample complexity
  - Sample complexity: number of pairwise comparison take prior to the termination
- 100 repetitions
- The confidence parameter δ was set to 0.05, and thus, the accuracy was significantly higher than 0.95 in every case.

Numerical experiments for identifying MPI



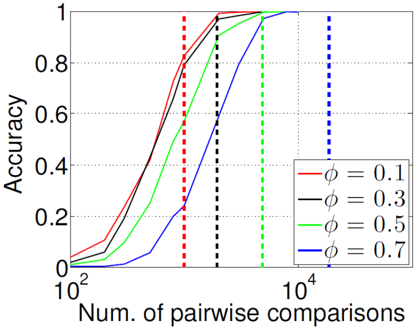
Numerical experiments for identifying **Most Probable Ranking (MPR)**

- Most probable ranking = center ranking

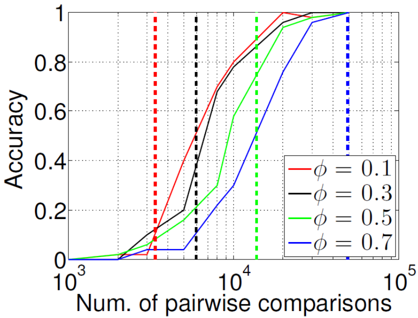
$$\tilde{r} = \arg\max_{r \in \mathbb{S}_M} \mathbb{P}(r | \phi, \tilde{r}) \tag{13}$$

- Parameter estimation method for Mallows which can handle incomplete ranking [Cheng et al., 2009]
- Validated on datasets that consist of pairwise comparisons
- Assessed the accuracy of the estimator for center ranking on datasets with various size.

Numerical experiments for identifying **MPR**



(a) $M = 10$.



(b) $M = 20$.

- Solid Line: Parameter estimation method by Cheng et al. [2009]
- Dashed vertical lines: Merge sort-based MPR algorithm

Conclusion

- The purposed algorithms are efficient, if the modeling assumption holds
- Minimize sample complexity
- Guarantee a certain level of confidence

The End