Regret-based Online Ranking for a Growing Digital Library

Erick Delage
Department of Management Sciences
HEC Montréal

June 30, 2009
Introduction

The “KL-Rank” Model
Reducing Regret in Ranking
Experiments with a Citation Database

Ranking for Web Search

knowledge discovery conference 2009 - Google Search - Mozilla Firefox

PAKDD 2009 - Front
LNAI 5467 Advances in Knowledge Discovery and Data Mining, 13th Pacific-Asia Conference, PAKDD 2009, Bangkok, Thailand, April 27-30, 2009 Proceedings ...
www.pakdd2009.org - Cached - Similar

Meetings and Conferences in Data Mining, Knowledge Discovery, and ... 7-11 Sep, ECML/PKDD-2009, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Bled, ...
www.kdnuggets.com/meetings/ - Cached - Similar

Knowledge Discovery and Data Mining - KDD2009 Conference in Paris 2 April 2009, Apply for the Google Female Conference and Travel Grant for KDD 2009, ...
Knowledge Extraction Engines - Microsoft AdCenter Labs ...
www.sigkdd.org/kdd2009/ - Cached - Similar

DEXA 2009
20th edition of DEXA, Linz, Austria 31 August - 4 September 2009 - 11th International Conference on Data Warehousing and Knowledge Discovery - DaWaK '09 ...
www.dexa.org - Cached - Similar

Call for Papers DaWaK '09 | DEXA 2009
11th International Conference on Data Warehousing and Knowledge Discovery (DaWaK 2009), Linz, Austria 31 August - 2 September 2009 ...
www.dexa.org/dawa09_cfp - Cached - Similar
More results from www.dexa.org

IC3K 2009 - International Joint Conference on Knowledge ...
International Conference on Knowledge Discovery and Information Retrieval ... A book will be published by Springer in the CCIS Series with IC3K 2009 best ...
www.ic3k.org - Cached - Similar

2009 International Conference on Knowledge Discovery (ICKD 2009)
Types of Data:
- Experts (flexible yet expensive and requires faith in the experts)
- Clickthrough (abundant yet noisy and subject to manipulation)

Types of Learning:
- Offline (uses historical data)
- Online (uses data as it is provided by users)

Our model uses clickthrough data assuming that it is truthful
Online Ranking Model I

User \#t

Library

Query (q_t)

Ranking Mechanism
Online Ranking Model II

Introduction
The “KL-Rank” Model
Reducing Regret in Ranking
Experiments with a Citation Database

User \( #t \)

Library

Rank map \( (\pi_t) \)

Ranked list \( (R_t) \)

Query \( (q_t) \)

Ranking Mechanism
Online Ranking Model III

- Ranked list ($R_t$) → User #t
- Library
- Rank map ($\pi_t$)
- Query ($q_t$) → Clicked item ($I_t$)
- Cost($R_t, I_t$)

6 E. Delage
Regret-based Online Ranking for a Growing Digital Library
We will answer the following question:

“As users are served and the library evolves, does there exist a ranking algorithm which long term performance is comparable to the best ranking chosen with the benefit of hindsight?”
In this Talk

- We will answer the following question:
  
  “As users are served and the library evolves, does there exist a ranking algorithm which long term performance is comparable to the best ranking chosen with the benefit of hindsight?”

- Empirical evidence will show that a greedy policy based on the “KL-Rank model” can outperform many popular ranking algorithms.
Outline

1. Introduction
2. The “KL-Rank” Model
3. Reducing Regret in Ranking
4. Experiments with a Citation Database
Outline

1. Introduction
2. The “KL-Rank” Model
3. Reducing Regret in Ranking
4. Experiments with a Citation Database
Deterministic vs. Non-deterministic Ranking Policies

- **Deterministic Ranking**
  - Rank items according to a score function

  \[
  \text{score}_\theta(l_{\pi(i)}; q) \geq \text{score}_\theta(l_{\pi(i+1)}; q) \quad \forall \ i \in \{1, 2, \ldots, |\mathcal{L}|\},
  \]

  with \( \text{score}_\theta(l; q) = \sum_i \theta_i \cdot \Phi_i(l, q) \).
Deterministic vs. Non-deterministic Ranking Policies

- **Deterministic Ranking**
  - Rank items according to a score function
  
  \[
  \text{score}_\theta(l_{\pi(i)}; q) \geq \text{score}_\theta(l_{\pi(i+1)}; q) \quad \forall \ i \in \{1, 2, ..., |\mathcal{L}|\},
  \]
  
  with \( \text{score}_\theta(l; q) = \sum_i \theta_i \cdot \Phi_i(l, q) \).

- **Non-deterministic Ranking**
  - Draw a ranking randomly from a distribution
  
  \[
  \mathbb{P}(\tilde{\pi}(i+1) = j | \tilde{\pi}(1), ..., \tilde{\pi}(i)) \propto \begin{cases} 
  0 & \text{if } l_j \in \{l_{\pi(1)}, ..., l_{\pi(i)}\} \\
  e^{\text{score}_\theta(l_j; q)} & \text{o.w.}
  \end{cases}
  \]
Deterministic vs. Non-deterministic Ranking Policies

- **Deterministic Ranking**
  - Rank items according to a score function
  \[
  \text{score}_\theta(l_{\pi(i)}; q) \geq \text{score}_\theta(l_{\pi(i+1)}; q) \quad \forall \ i \in \{1, 2, \ldots, |L|\},
  \]
  with \( \text{score}_\theta(l; q) = \sum_i \theta_i \cdot \Phi_i(l, q). \)

- **Non-deterministic Ranking**
  - Draw a ranking randomly from a distribution
  \[
  \mathbb{P}(\tilde{\pi}(i+1) = j | \tilde{\pi}(1), \ldots, \tilde{\pi}(i)) \propto \begin{cases} 
  0 & l_j \in \{l_{\pi(1)}, \ldots, l_{\pi(i)}\} \\
  e^{\text{score}_\theta(l_j; q)} & \text{o.w.}
  \end{cases}
  \]
  - Non-deterministic can closely replicate any determ. ranking
  - In both cases, generating a ranking costs \( O(n \log(n)) \)
Choosing the Right Cost Model for Ranking

- **Common cost models:**
  - Fixed cost for not ranking the clicked item first
  - Relative distance of the “clicked” items
  - Neg. normalized discounted cumulated gain (see [Li et al., 2008])

- Finding the best deterministic ranking mechanism is hard (non-continuous & non-convex)
Choosing the Right Cost Model for Ranking

- Common cost models:
  - Fixed cost for not ranking the clicked item first
  - Relative distance of the “clicked” items
  - Neg. normalized discounted cumulated gain (see [Li et al., 2008])

- Finding the best deterministic ranking mechanism is hard
  (non-continuous & non-convex)

- The problem is continuous for non-deterministic rankings
Choosing the Right Cost Model for Ranking

- **Common cost models:**
  - Fixed cost for not ranking the clicked item first
  - Relative distance of the “clicked” items
  - Neg. normalized discounted cumulated gain (see [Li et al., 2008])

- Finding the best deterministic ranking mechanism is hard (non-continuous & non-convex)

- The problem is continuous for non-deterministic rankings

- **The problem becomes convex when using KL-Cost:**

\[
\text{KL-Cost}(\tilde{\pi}_\theta, l) = \min_{\nu_{\tilde{\pi}} \in \mathcal{D}(l)} D_{KL}(\nu_{\tilde{\pi}} \| \mu_{\tilde{\pi}_\theta})
\]

(measures the K-L divergence between the ranking distrib. and the set of ranking distrib. that put item \( l \) first w.p. one)
Comparing a Heuristic with Best Policy in Hindsight

After $\tilde{t}$ users have been served, a good heuristic is to use for user $\tilde{t} + 1$ a random ranking that is optimal w.r.t. the convex problem:

$$\tilde{\pi}_{\theta_{\tilde{t}+1}} : \theta_{\tilde{t}+1} = \arg\min_{\theta \in \Theta} \sum_{t=1}^{\tilde{t}} \text{KL-Cost}(\tilde{\pi}_{\theta}, l_t)$$
Comparing a Heuristic with Best Policy in Hindsight

After $\bar{t}$ users have been served, a good heuristic is to use for user $\bar{t} + 1$ a random ranking that is optimal w.r.t. the convex problem:

$$
\hat{\pi}_{\theta_{\bar{t}+1}} : \theta_{\bar{t}+1} = \arg\min_{\theta \in \Theta} \sum_{t=1}^{\bar{t}} \text{KL-Cost}(\hat{\pi}_\theta, I_t)
$$

Unfortunately, this approach has no guarantee that in the long run the performance is “close” to optimal (i.e., generates low regret).

$$
R(T) = \sum_{t=1}^{T} \text{KL-Cost}(\hat{\pi}_{\theta_t}, I_t) - \min_{\theta \in \Theta} \sum_{t=1}^{T} \text{KL-Cost}(\hat{\pi}_{\theta}, I_t) = ???
$$
Outline

1. Introduction
2. The “KL-Rank” Model
3. Reducing Regret in Ranking
4. Experiments with a Citation Database
One can choose a feature extracting function $\Phi(\cdot, \cdot)$ and a set of competing ranking policies parameterized by $\theta \in \Theta$.

If $\max_{q, l} \|\Phi(l, q)\|$ and $\max_{\theta_1, \theta_2 \in \Theta} \|\theta_1 - \theta_2\|$ are bounded, then cumulated regret is $O(T^{1/2})$ with “greedy projection”:

$$\theta_{t+1} = \mathcal{P}_\Theta \left( \theta_t - \frac{\alpha}{\sqrt{t}} \nabla_{\theta} \text{KL-Cost}(\tilde{\pi}_\theta, I_t) \right).$$

Average regret is assured to go to zero at the rate of $O(T^{-1/2})$ for any sequence of $T$ user queries. (based on [Zinkevich, 2003])
Regret in a Growing Library

Difficulties:

- As library grows, new features are needed to distinguish among the larger set of items.
- Regret needs to be measured w.r.t. the features that were available when each user visited:

\[
R(T) = \sum_{t=1}^{T} \text{KL-Cost}(\tilde{\pi}_{\theta_t}, l_t) - \min_{\theta} \sum_{t=1}^{T} \text{KL-Cost}(\tilde{\pi}_{P_{\Theta_t}(\theta)}, l_t).
\]
Regret in a Growing Library

Difficulties:

- As library grows, new features are needed to distinguish among the larger set of items.
- Regret needs to be measured w.r.t. the features that were available when each user visited:

\[
\mathcal{R}(T) = \sum_{t=1}^{T} \text{KL-Cost}(\tilde{\pi}_{\theta_t}, l_t) - \min_{\theta} \sum_{t=1}^{T} \text{KL-Cost}(\tilde{\pi}_{P_{\Theta_t}(\theta)}, l_t) .
\]

Solution:

- Use “greedy projection with exponential restart”.
- Regret cumulated bounded by \(O(T^{3/4})\) if, for all \(\bar{t} \leq T\),

\[
\max_{\theta_1, \theta_2 \in \Theta_{\bar{t}}} \|\theta_1 - \theta_2\| \cdot \max_{t < \bar{t}, q, l \in L_t} \|\Phi_t(l, q)\| = O(\bar{t}^{1/4}) .
\]
- Avg. regret drops at rate \(O(T^{-1/4})\) for any sequence of users.
Outline

1. Introduction
2. The “KL-Rank” Model
3. Reducing Regret in Ranking
4. Experiments with a Citation Database
The Cora Dataset

- Cora is an internet portal developed by A. McCallum et al.
- The dataset contains information about 10000 papers from 1913 to 1999: e.g., year of publication, topic, citations.
- We constructed 22000 queries based on each paper's citations:

  ![Diagram]

  - **In Cora**: Paper A cites Paper B on Topic C
  - **Query**: The author of A queries Topic C and chooses B as most relevant
The Cora Dataset

- Cora is an internet portal developed by A. McCallum et al.
- The dataset contains information about 10000 papers from 1913 to 1999: e.g., year of publication, topic, citations.
- We constructed 22000 queries based on each paper's citations:

  In Cora
  
  Paper A cites Paper B on Topic C

  Query
  
  The author of A queries Topic C and chooses B as most relevant

- In implementation, each paper has a dedicated fitness parameter $\theta_i$.
- Our theory states that if $\Theta_{\bar{t}} \subseteq \text{Ball}(0, \alpha^{\bar{t}^{1/4}})$, for all $\bar{t} \leq T$, then average regret is $O(T^{-1/4})$. 
Competing Algorithms

- **NoRegret KL-Rank**: Rank according to KL-Rank’s random policy (guaranteed to lead to no average regret)
- **Greedy KL-Rank**: Rank directly according to score function obtained using KL-Rank
Competing Algorithms

- **NoRegret KL-Rank**: Rank according to KL-Rank’s random policy (guaranteed to lead to no average regret)
- **Greedy KL-Rank**: Rank directly according to score function obtained using KL-Rank
- **Online Ranknet**: Online version of RankNet by Burges et al.
- **Category Ranking**: Projection based algorithm proposed by Crammer & Singer
- **HITS & Pagerank**: Graph based ranking algorithms
- **2-class McRank**: Li et al.’s classification method
Introduction
The “KL-Rank” Model
Reducing Regret in Ranking
Experiments with a Citation Database

Statistics obtained in set of 22000 queries

<table>
<thead>
<tr>
<th>Method</th>
<th>Avg Rel. Click Dist.</th>
<th>Clicked 1st Rate</th>
<th>Avg NDCG score</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoRegret KL-Rank</td>
<td>25.7% (4)</td>
<td>14.3% (5)</td>
<td>34.4% (5)</td>
</tr>
<tr>
<td>Greedy KL-Rank</td>
<td>19.2% (1)</td>
<td>17.8% (1)</td>
<td>39.1% (1)</td>
</tr>
<tr>
<td>Online RankNet</td>
<td>19.8% (2)</td>
<td>16.1% (4)</td>
<td>37.4% (2)</td>
</tr>
<tr>
<td>Category Ranking</td>
<td>23.2% (3)</td>
<td>16.5% (2)</td>
<td>37.3% (2)</td>
</tr>
<tr>
<td>HITS / PageRank</td>
<td>26.5% (5)</td>
<td>16.6% (2)</td>
<td>36.2% (4)</td>
</tr>
<tr>
<td>McRank</td>
<td>26.3% (5)</td>
<td>14.0% (5)</td>
<td>34.1% (5)</td>
</tr>
</tbody>
</table>

Although Greedy KL-Rank has no performance guarantee (is more opportunistic), it outperforms all other methods.
Conclusion

- Clickthrough data is most natural form of feedback from users and fits well the KL-Rank model.
- NoRegret KL-Rank is the first to provide guarantees w.r.t. long term performance (average regret) of policies on arbitrary sequence of users.
- Analysis provides guidance for adding features as the library grows while avoiding overfitting.
- In practice, Greedy KL-Rank can be less conservative and seems to outperform many popular methods.
Conclusion

- Clickthrough data is most natural form of feedback from users and fits well the KL-Rank model
- NoRegret KL-Rank is the first to provide guarantees w.r.t. long term performance (average regret) of policies on arbitrary sequence of users
- Analysis provides guidance for adding features as the library grows while avoiding overfitting
- In practice, Greedy KL-Rank can be less conservative and seems to outperform many popular methods
- Future work:
  - Compare methods on other datasets and real-time
  - Study worst case performance guarantees for Greedy KL-Rank
  - Study effects of clickthrough biases and malignant usage
Questions & Comments ...

... Thank you!
Greedy Projection with Exponential Restart Algorithm

Decompose timeline in periods of size that grows exponentially

![Diagram showing exponential restarts](Diagram)

Let $t = 0$, for cycles $m = 1, 2, \ldots$:

- Keep $\Phi_m(\cdot, \cdot)$ and $\Theta_m$ constant over period $t \in [2^{m-1}, 2^m - 1]$
- Initialize $\theta_{t+1} \in \Theta_m$
- For $k = 1, 2, \ldots, 2^{m-1}$:
  1. At time $t = 2^{m-1} + k - 1$, receive query $q_t$
  2. Generate a random ranking $\tilde{R}_t$ using $\tilde{\pi}_{\theta_t}$
  3. Learn about the preferred item $l_t$
  4. Update $\theta_{t+1} = \mathcal{P}_{\Theta_m} \left( \theta_t - \frac{\alpha}{\sqrt{k}} \nabla_{\theta} \text{KL-Cost}(\tilde{\pi}_{\theta_t}, l_t) \right)$