THE 18TH ACM SIGKDD CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING

KDD 2012 BEIJING

Beijing, China August 12-16, 2012

Online Learning to Diversify from Implicit Feedback

Karthik Raman
Cornell University
karthik@cs.cornell.edu

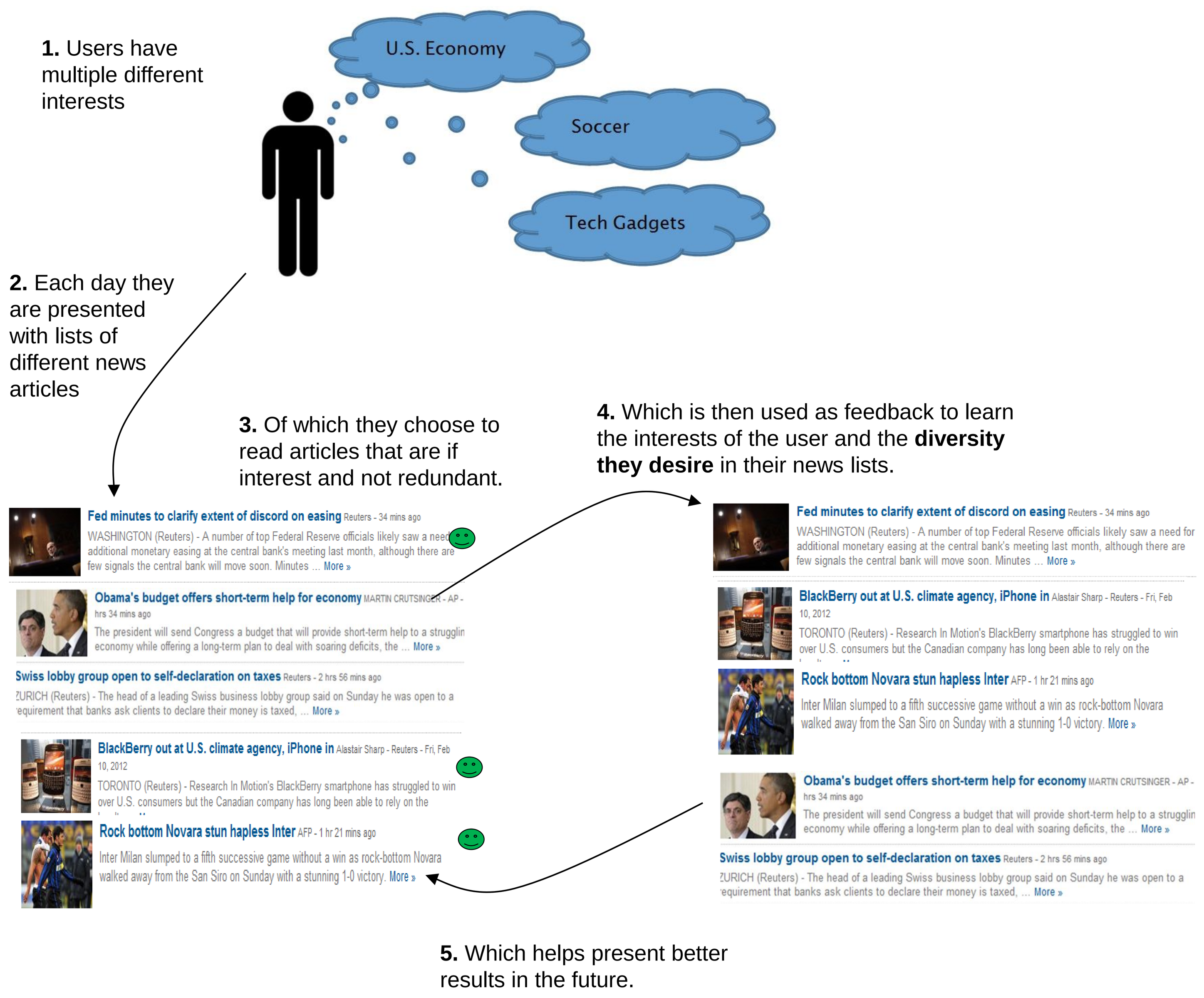
Pannaga Shivaswamy
Cornell University
pannaga@cs.cornell.edu

Thorsten Joachims
Cornell University
tj@cs.cornell.edu

Overview

Present simple, efficient **online learning** algorithms for learning both relevance and **diversity** in rankings, which are shown to be theoretically and empirically robust.

Example: News Recommendation



Experiments:

- Used 2 news datasets: RCV1 and 20NG
- Simulated diverse users with 5 interests.
- TFIDF values used for document features,

Findings:

- Highly robust to noisy user feedback and quality of feedback.
- Able to learn desired amount of diversity.

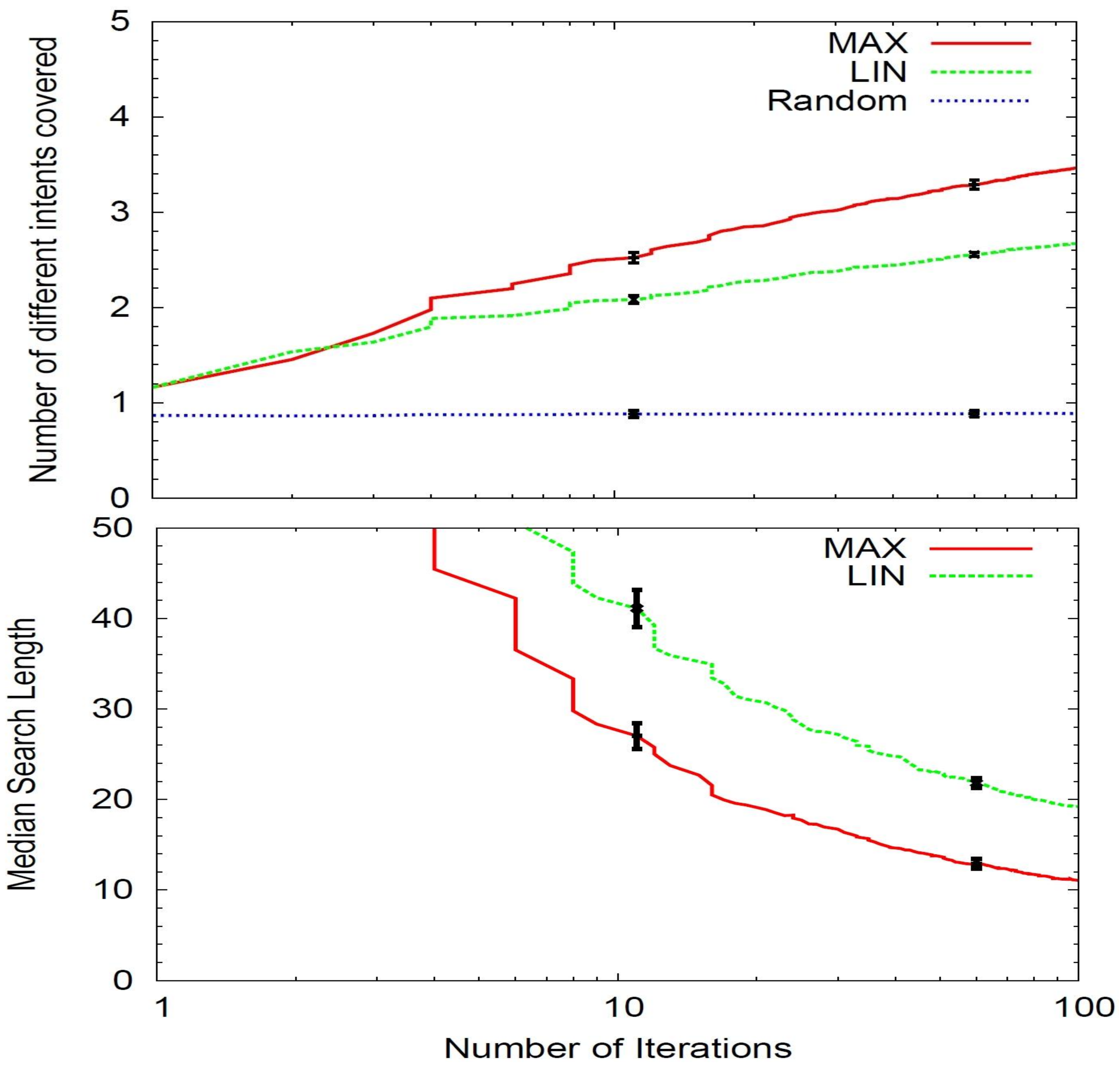


Fig1: Change in the number of intents covered (above) and search length (below) over time.

Motivation:

- Most research on focused on **extrinsic** diversity.
- **Intrinsic Diversity:** Diversity in the aspects of a single information need.
- No previous method for learning required amount of diversity from user feedback/click data.

Learning Diversity:

- Capture diversity via **non-linear combination** of document feature vectors:

$$\phi^j(\mathbf{x},\mathbf{y}) = F(\{\phi^j(d_{i_1}), \phi^j(d_{i_2}), ..., \phi^j(d_{i_k})\})$$

- Choice of F determines how much redundancy in predicted rankings. Couple of examples:

Name	$F(A)$
LIN	$F(A) = \sum_{a \in A} a$
MAX	$F(A) = \max_{a \in A} a$

Algorithm:

1. Start with **w = 0**
2. Present ranking as per current **w**
3. Observe user feedback.
4. Perceptron Update:

   w += *Feedback FeatVec* – *Presented FeatVec*
5. Repeat from step 2 for next user session

- Simple and efficient
- Theoretically guaranteed to converge to optimal
- Ranking in step 2 can be easily computed via simple **greedy** algorithm for any **submodular** *F*.

- Outperforms supervised learning within few iterations despite not receiving true labels.

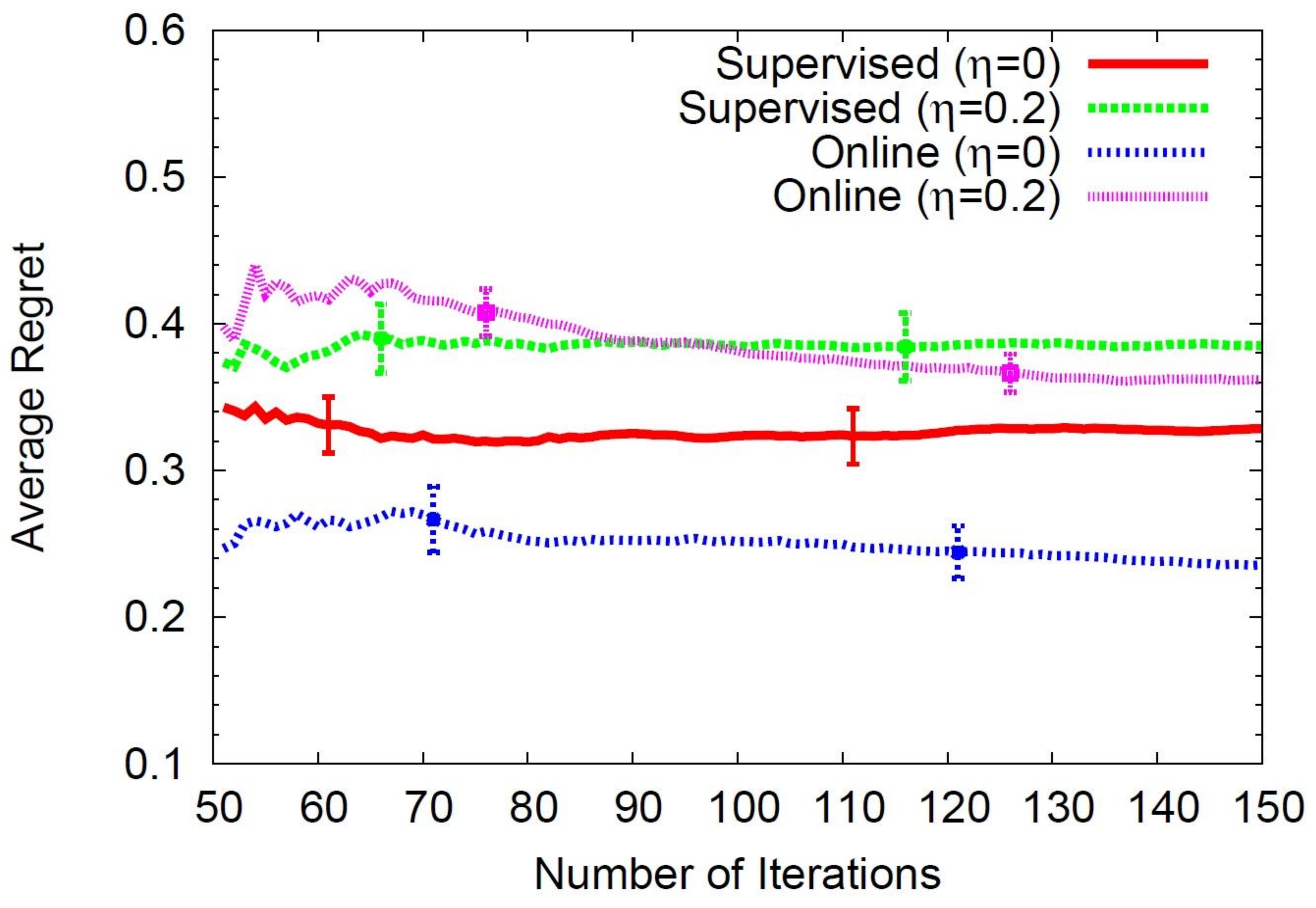


Fig2: Comparison with supervised learning.

Future Directions:

- User study of model in recommendation system.
- Extending to extrinsic diversity.

Paper ID: rt833