Activity Ranking in LinkedIn Feed

Deepak Agarwal, Bee-Chung Chen, Rupesh Gupta, Joshua Hartman, Qi He, Anand Iyer, Sumanth Kolar, Yiming Ma, Pannaga Shivaswamy, Ajit Singh, Liang Zhang
LinkedIn Feed

• Professional network
• Heterogeneous updates
  – More than 40 types
  – Shared articles, job changes, connection updates etc.
• Challenges
  – Large scale (313M+ members)
  – Relevance & Personalization
  – Freshness, diversity, user fatigue
• How do we rank activities?
Straw man approaches

• Reverse chronological ranking
  – Fresh but not relevant
• Ranking by social popularity
  – Likes, a useful signal
  – CTR not monotonically related
  – Not all activities have likes
Activities on LinkedIn Feed

- Taxonomy (actor type, verb type, object type)
  - Connection: (member, connect, member)
  - Opinion: (member, like, article)

- What happens if we simply rank by CTR?
Relevance via CTR prediction

• Training data collection
  – Requires randomization

• Personalization features
  – E.g., viewer type affinity, viewer actor affinity

• Large scale logistic regression via ADMM
  – Scalable, distributed algorithm

• Offline evaluation
  – Unbiased estimation via replay
Large Scale Logistic Regression via ADMM

Iteration 1

BIG DATA

Partition 1

Logistic Regression

Partition 2

Logistic Regression

Partition 3

Logistic Regression

Partition K

Logistic Regression

Consensus Computation
Large Scale Logistic Regression via ADMM

Iteration 1
Large Scale Logistic Regression via ADMM

Iteration 2
Logistic regression enough?

• Individual item ranking
• No diversity
  – Imagine seeing ten profile pictures one after the other
• Old relevant items versus new relevant items
  – Freshness
• User fatigue via repetition of the same item
  – Impression discounting
Effect of diversity

• Users prefer a more diverse feed
  – Repetition of actorId or verbType causes lower CTR

– Diversity reranker to ensure the feed is diverse
Effect of freshness

• By age

• User fatigue by repeated impressions
Desktop A/B Test Results

- Personalization
  - Viewer, type affinity
  - Time feature
- Exponential Impression discounting
Conclusions

• Multiple challenges for good user experience
  – Large scale learning problem
  – Heterogeneous updates
  – Relevance, personalization
  – Freshness
  – Fatigue reduction
  – Diversity