Personalized Retweet Prediction in Twitter

Liangjie Hong*, Lehigh University
Aziz Doumith, Lehigh University
Brian D. Davison, Lehigh University

OVERVIEW

- Motivation
- Related Work
- Our Method
- Experimental Results

Motivations

Social information platforms

MOTIVATIONS

Information overload

MOTIVATIONS

Information shortage

Motivations

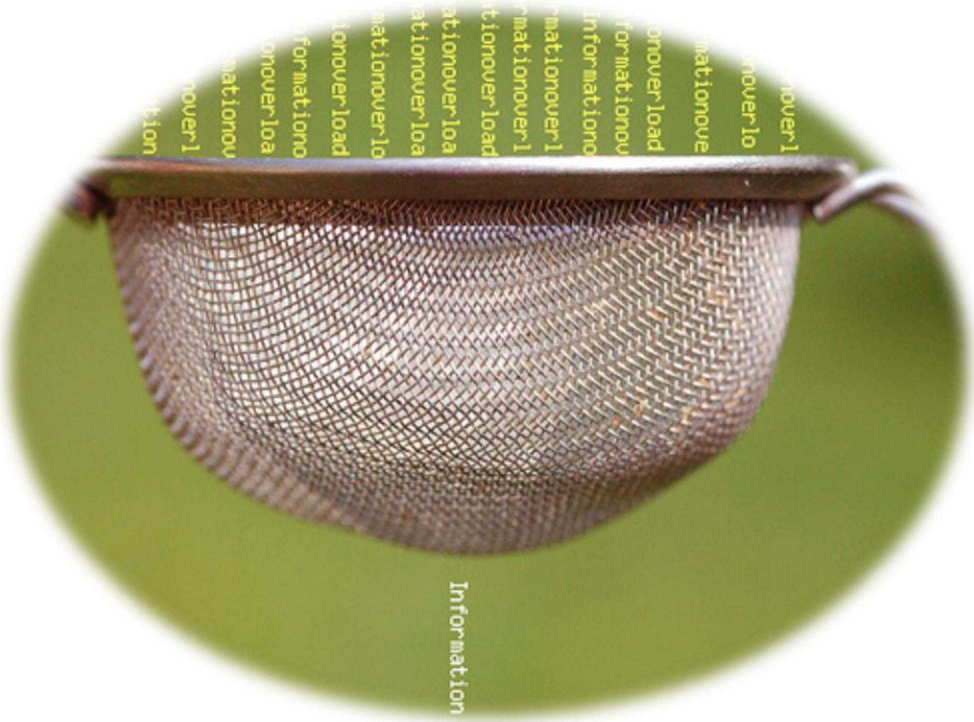

Photo from: http://www.jenful.com/2011/06/google-1-and-the-filter-bubble/

MOTIVATIONS

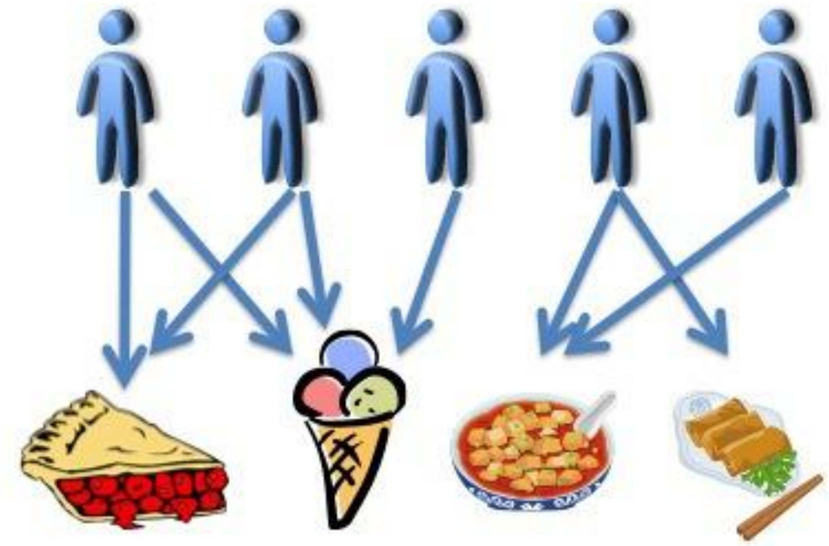

Photo from: http://graphlab.org/

MOTIVATIONS

Photo from: http://kexino.com/

Motivations

Photo from: http://performancemarketingassociation.com/

TASKS

Given a target user and his/her friends, provide a ranked list of tweets from these friends such that the tweets that are potentially retweeted will be ranked higher.

Photo from: http://performancemarketingassociation.com/

Related Work

Generic Popular Tweets Analysis/Prediction

- [Suh et al., SocialCom 2010]
- [Y. Kim and K. Shim, ICDM, 2011]
- [Uysal and W. B. Croft, CIKM 2011]
- [Hong et al., WWW 2011]

Personalized Tweets Prediction

- [Chen et al., SIGIR 2012]
- [Peng et al., ICDM Workshop 2011]

Related Work

Generic Popular Tweets Analysis/Prediction

- [Suh et al., SocialCom 2010]
- [Y. Kim and K. Shim, ICDM, 2011]
- [Uysal and W. B. Croft, CIKM 2011]
- [Hong et al., WWW 2011]

Personalized Tweets Analysis/Prediction

- [Chen et al., SIGIR 2012]
- [Peng et al., ICDM Workshop 2011]

Understanding users’ behaviors & content modeling

Our Method

Design requirements

- Utilize users' historical behaviors
- Collaborative filtering
- Incorporating a rich-set of features
- Coupled modeling with content
- Learning a correct objective function
- Scalability

OUR METHOD

Design requirements

- Utilize users' historical behaviors
- Collaborative filtering

- Latent factor models

OUR METHOD

Design requirements

- Utilize users' historical behaviors
- Collaborative filtering
- Incorporating a rich-set of features

↓

- Latent factor models
  - Factorization Machines [Rendle, ACM TIST 2012]

Our Method

Factorization Machines

- Generic enough
  - matrix factorization
  - pairwise interaction tensor factorization
  - SVD++
  - neighborhood models
  - …
- Technically mature
  - [Rendle, ICDM 2010]
  - [Rendle et al., SIGIR 2011]
  - [Freudenthaler et al., NIPS Workshop 2011]
  - [Rendle et al., WSDM 2012]
  - [Rendle, ACM TIST 2012]

Our Method

Extending Factorization Machines

- Non-negative decomposition of term-tweet matrix
  - Compatible to standard topic models
- Co-Factorization Machines
  - Multiple aspects of the dataset
    - Shared feature paradigm
    - Shared latent space paradigm
    - Regularized latent space paradigm

Our Method

Design requirements

- Utilize users' historical behaviors
- Collaborative filtering
- Incorporating a rich-set of features
- Coupled modeling with content
- Learning a correct objective function
- Scalability

Our Method

Design requirements

- Learning objective functions for different aspects
  - User decisions
    - Ranking-based loss
      Weighted Approximately Rank Pairwise loss (WARP)
  - Content modeling
    - Log-Poisson loss
    - Logistic loss

Our Method

WARP loss

- Proposed by [Usunier et al., ICML 2009]
- Image retrieval tasks and IR tasks
  [Weston et al., Machine Learning 2010]
  [Weston et al., ICML 2012]
  [Weston et al., UAI 2012]
  [Bordes et al, AISTATS 2012]
- Can mimic many ranking measures
  - NDCG, MAP, Precision@k

- Applied to collaborative filtering

Our Method

Design requirements

- Utilize users' historical behaviors
- Collaborative filtering
- Incorporating a rich-set of features
- Coupled modeling with content
- Learning a correct objective function
- Scalability (Stochastic Gradient Descent)

EXPERIMENTS

Twitter data

- 0.7M target users with 11M tweets
- 4.3M neighbor users with 27M tweets
- "Complete" sample for each target user
- Mean Average Precision (MAP) as measure
- Train/test on consecutive time periods

EXPERIMENTS

Comparisons

- Matrix factorization (MF)
- Matrix factorization with attributes (MFA)
- CPTR [Chen et al, SIGIR 2012]
- Factorization machines with attributes (FMA)
- CoFM with shared features (CoFM-SF)
- CoFM with shared latent spaces (CoFM-SL)
- CoFM with latent space regularization (CoFM-REG)

EXPERIMENTS

Comparisons

- Matrix factorization (MF)
- Matrix factorization with attributes (MFA)
- CPTR [Chen et al, SIGIR 2012]
- Factorization machines with attributes (FMA)
- CoFM with shared features (CoFM-SF)
- CoFM with shared latent spaces (CoFM-SL)
- CoFM with latent space regularization (CoFM-REG)

EXPERIMENTS

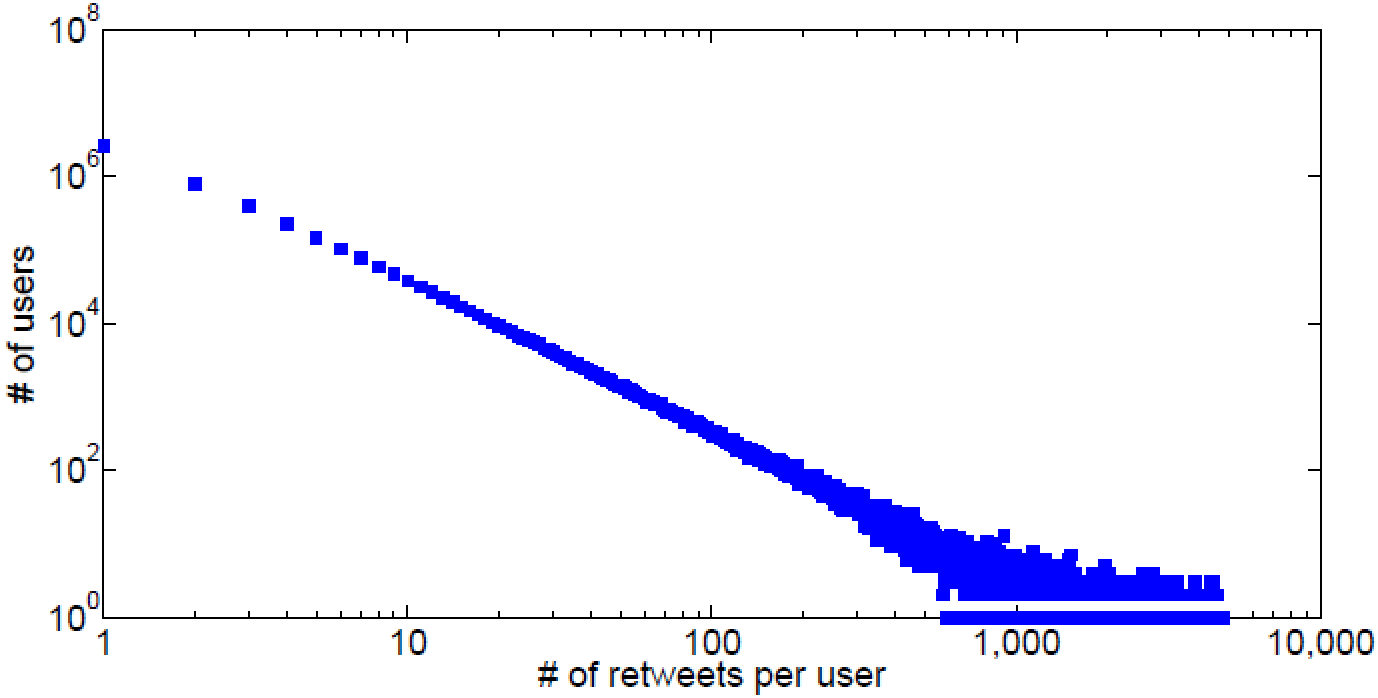

Figure 1: The sparsity of retweets per user.

EXPERIMENTS

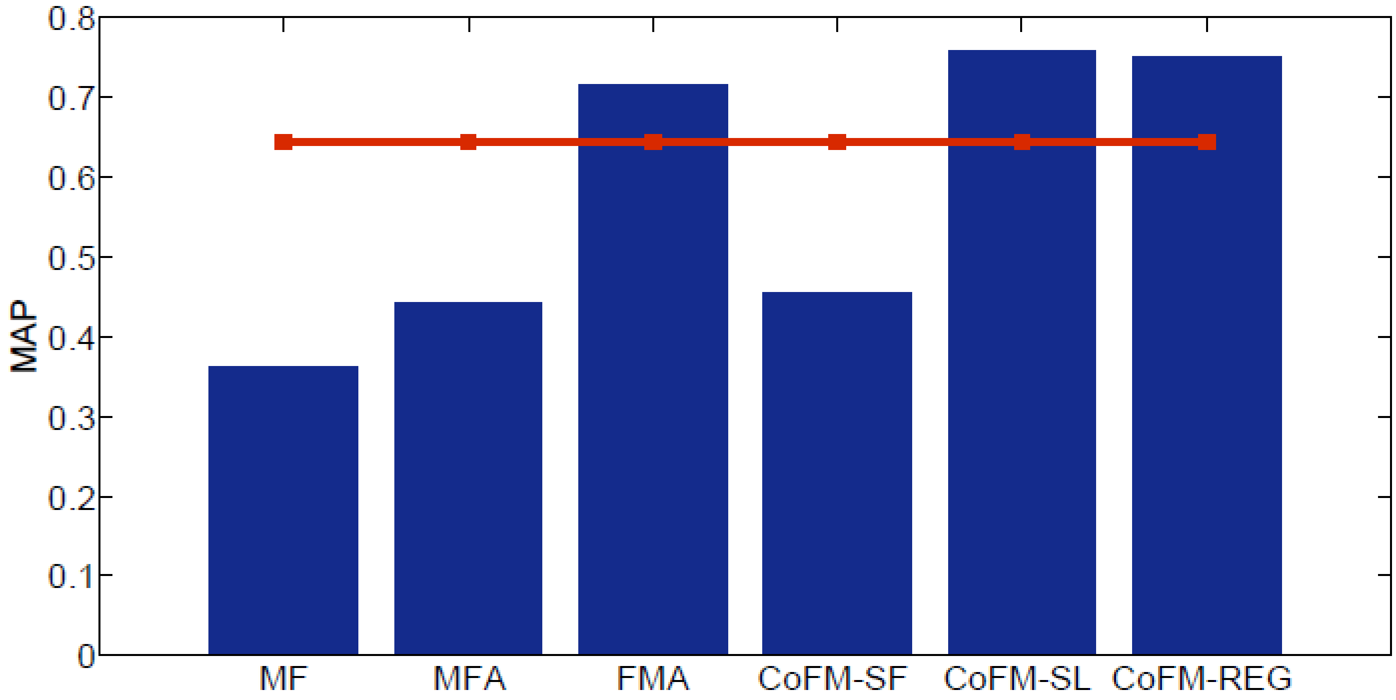

Figure 2: The results on retweet prediction. The red line is the baseline CPTR**.**

EXPERIMENTS

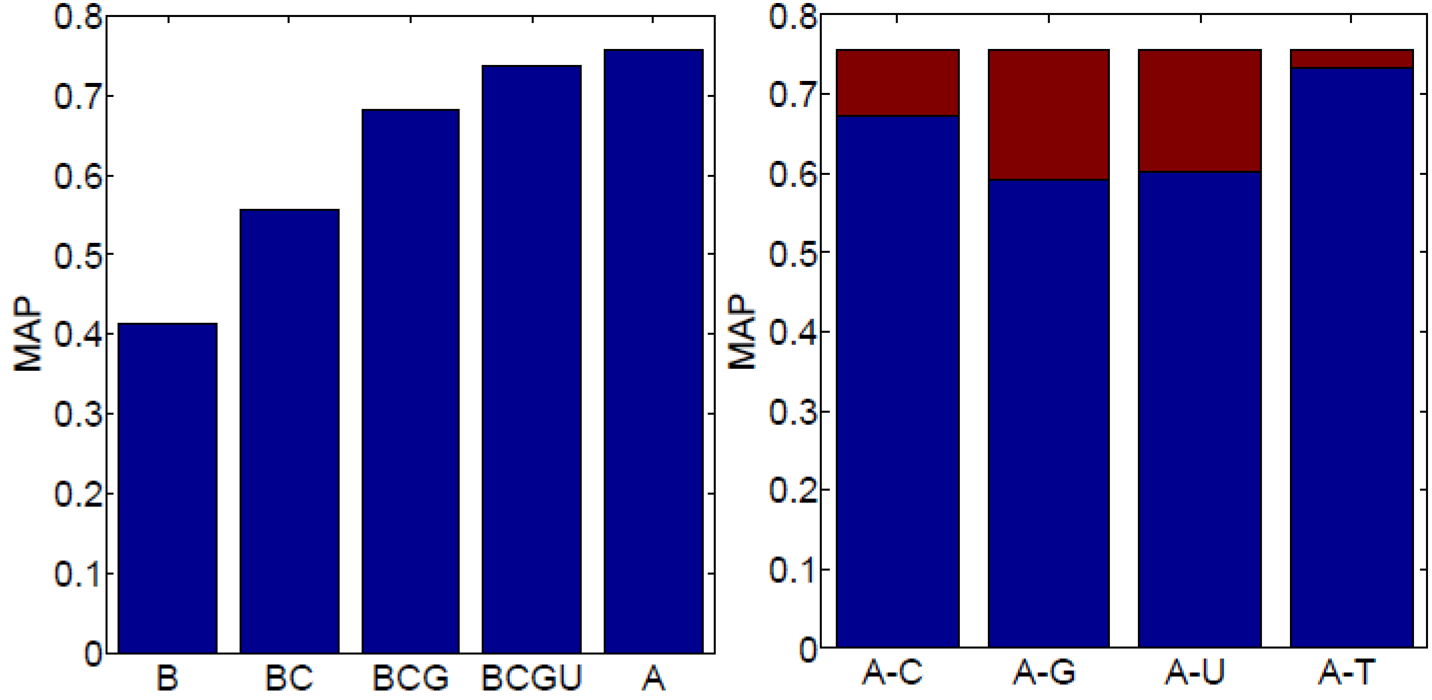

Figure 3: The impact of different groups of features. The effect of "add on" is shown on the left and the effect of "take out" is on the right. For both figures, "A", "B", "C", "G", "U" and "T" stand for "All", "Base model", "Content feature", "Graph feature", "User feature" and "Temporal feature" respectively.

EXPERIMENTS

Examples of topics are shown. The terms are top ranked terms in each topic. The topic names in bold are given by the authors.
Entertainment
album music lady artist video listen itunes apple produced movies #bieber bieber new songs
Finance
percent billion bank financial debt banks euro crisis rates greece bailout spain economy
Politics
party election budget tax president million obama money pay bill federal increase cuts

CONCLUSIONS

- Main contributions
  - Propose Co-Factorization Machines (CoFM) to handle two (multiple) aspects of the dataset.
  - Apply FM to text data with constraints to mimic topic models
  - Introduce WARP loss into collaborative filtering/recsys models
  - Explore a wide range of features and demonstrate the effectiveness of feature sets with significant improvement over several non-trival baselines.

THANK YOU.

Liangjie Hong

PhD candidate

WUME Lab

Lehigh University

lih307@cse.lehigh.edu