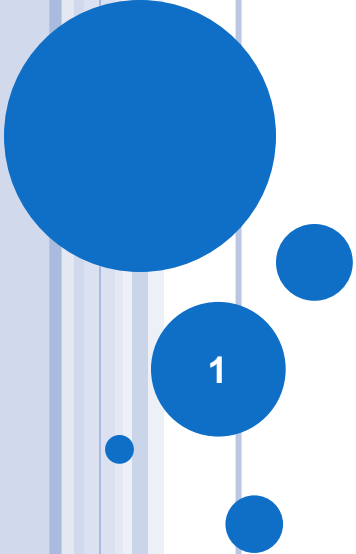


LEARNING TO RANK SOCIAL UPDATE STREAMS



Liangjie Hong*, Lehigh University
Ron Bekkerman, LinkedIn Corp.
Joseph Adler, LinkedIn Corp.
Brian D. Davison, Lehigh University

* Part of this work was done when the author was on an internship at LinkedIn Corp.

OVERVIEW

- Social Update Streams
- Overview of LinkedIn
- Social Stream Ranking & Dataset
- Methods
- Experiments
- Conclusion

ONE SLIDE TAKEAWAY

IMPORTANT



- Task
 - Improve user engagement by re-ranking social updates
- Main results
 - We demonstrate that recommender systems + preference-based learning can be used to re-rank social updates.
 - A linear model can achieve 60% of the performance of latent factor models, on average.
 - A tensor factorization model with regression on explicit features works the best.
 - The cold-start problem makes it impossible to model some kinds of interactions.

SOCIAL UPDATE STREAMS



A screenshot of a social media update stream. At the top, there is a search bar with the text "What's on your mind?" and a "Share" button. Below the search bar, a list of updates is shown. Each update includes a profile picture, a name, and a text post. The updates are: Miko Matsumura: "I am not at #SXSW FAIL"; John Dodge: "NASA's LROC Satellite will offer bird's eye view of moon: A lunar satellite NASA plans to launch in May will sen..."; Austin Hill: "Lithium breakthrough could charge batteries in 10 seconds"; Walt Mossberg: "can feel Spring arriving, in fits and starts."; Siobhan Nash Santinelli: "A new attitude ..."; and a post from Dennis Lewis: "Dennis Lewis just posted a humorous and revealing story entitled 'The Student' on his blog at http://dennislewisblog.com". To the right of the main stream, there are sections for "TOMORROW" (Peter Vesterbacka's birthday), "HIGHLIGHTS" (On the Road - A Clearing in the Fog of Compl...), "Photos from tour", "Fortune Magazine", "ITNEWS.com", and "Looking for...".

SOCIAL UPDATE STREAMS

Problems?



SOCIAL UPDATE STREAMS

Information overload

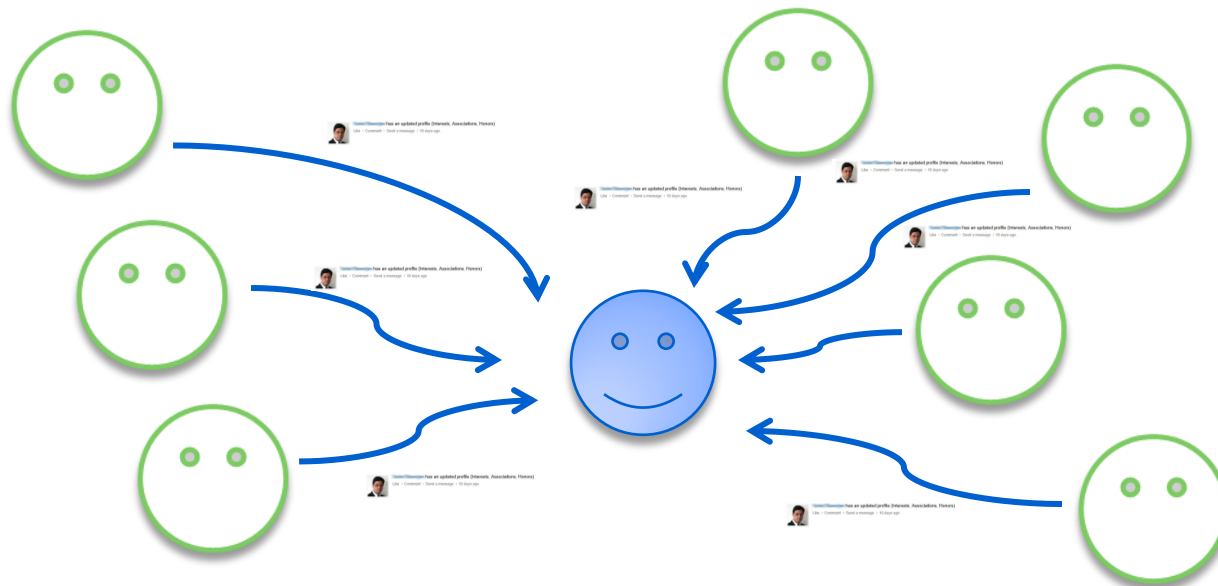


SOCIAL UPDATE STREAMS

Information shortage



SOCIAL UPDATE STREAMS



OVERVIEW OF LINKEDIN



OVERVIEW OF LINKEDIN

- Founded in Dec. 2002, launched in May 2003
- 160M¹ users in 200 countries and territories
- Biggest social network for professionals



¹ As of March 2012

OVERVIEW OF LINKEDIN



LINKEDIN HOMEPAGE

The screenshot displays the LinkedIn homepage for a user named Lianzhi Hong. At the top, there is a navigation bar with links for Home, Profile, Contacts, Groups, Jobs, Inbox, Companies, News, and More. A search bar and a notification bell are also present. Below the navigation bar, a banner for a GIS Master's/Cert Online program is visible. The main content area is divided into several sections: 'LinkedIn Today' with featured articles, 'People You May Know' with profile suggestions, 'Ads by LinkedIn Members', and 'Who's Viewed Your Profile?' showing profile views and search results. A large 'All Updates' section is highlighted with a red border, containing a list of recent activity including profile updates, connection notifications, and posts from users like Eric Leebow and Xieoping Xi. The 'Your LinkedIn Network' section shows 170 connections and 3,856 new people in the network. At the bottom, there are sections for 'Jobs You May Be Interested In' and 'Groups You May Like'.

LINKEDIN HOMEPAGE



has an updated profile (Interests, Associations, Honors)
Like • Comment • Send a message • 16 days ago



has an updated current title: Entrepreneur Platform Builder for Social & Mobile Experience Design Strategy at [Thoroughbranding](#)
Like • Comment • Send a message • 2 days ago



Craig Kessler Question 1: Do you have a budget for Valentine's Day this year? - PLZ RT #bp
7 minutes ago from Twitter • Reply privately • Add comment

OVERVIEW OF LINKEDIN

LINKEDIN TODAY · MAR 17
The most shared news on LinkedIn Take the tour » Powered by [Twitter](#) + [LinkedIn](#)

FRONT PAGE IT ONLINE MEDIA INTERNET COMPUTER... MARKETING &... DESIGN BROWSE ALL SA

Hiring Managers Take Their Time Filling Jobs

[online.wsj.com](#) [Share](#) [Save](#)
Recruiters say they are having trouble finding candidates for many skilled positions, and once candidates are found, hiring managers are taking longer to pull the trigger.

136

TOP SOURCES

- [online.wsj.com](#)
- [money.cnn.com](#)
- [npr.org](#)
- [bloomberg.com](#)
- [hbr.org](#)

SUGGESTED INDUSTRIES

IT » Based on your professional experience

Online Media » Based on your professional experience

Internet » Popular on LinkedIn

SUGGESTED INDUSTRIES

IT INDUSTRY SEE ALL » Based on your professional experience [Follow](#)

- [Microsoft More 'Ethical' Than...](#) [searchengineland.com](#) 90 shares
- [Sharing Everything I Learned As A...](#) [seomoz.org](#) 107 shares
- [LCD technology torn down and...](#) [engadget.com](#) 77 shares

ONLINE MEDIA INDUSTRY SEE ALL » Based on your professional experience [Follow](#)

- [A Letter to Our Readers](#) [lyfines.com](#) 306 shares
- [New York Times Agrees To Apple Terms...](#) [medianemomatters.com](#) 123 shares
- [Neiman Marcus Launches Nationwide...](#) [marshable.com](#) 304 shares

INTERNET INDUSTRY SEE ALL » Popular on LinkedIn [Follow](#)

- [Whom Should You Hire at a Startup?...](#) [techcrunch.com](#) 702 shares
- [Will You Pay for the New York Times...](#) [marshable.com](#) 323 shares
- [Hiring Managers Take Their Time...](#) [online.wsj.com](#) 136 shares

Download

OVERVIEW OF LINKEDIN

LINKEDIN TODAY · MAR 17
The most shared news on LinkedIn Take the tour » Powered by +

FRONT PAGE IT ONLINE MEDIA INTERNET COMPUTER... MARKETING &... DESIGN BROWSE ALL SA

Hiring Managers Take Their Time Filling Jobs

Recruiters say they are having trouble finding candidates for many skilled positions, and once candidates are found, hiring managers are taking longer to pull the trigger.

online.wsj.com

Share Save

TOP SOURCES

- online.wsj.com »
- money.cnn.com »
- npr.org »
- blomberg.com »
- hbr.org »

SUGGESTED INDUSTRIES

Startup Veterans Launch Imagine K12, A "Y..."

Groupm Is Said to Discuss IPO Valuation

HTC Thunderbolt Now Available

Hiring Managers Take Their Time Filling Jobs

SUGGESTED INDUSTRIES

IT INDUSTRY SEE ALL »

- Microsoft More "Ethical" Than...
searchengineland.com
90 shares
- Sharing Everything I Learned As A...
seomoz.org
107 shares

ONLINE MEDIA INDUSTRY SEE ALL »

- A Letter to Our Readers
lyfiles.com
306 shares
- New York Times Agrees To Apple Terms...
mediamemo.alltagged.com
123 shares

INTERNET INDUSTRY SEE ALL »

- Whom Should You Hire at a Startup? ...
techcrunch.com
702 shares
- Will You Pay for the New York Times...
marketable.com
323 shares

Groups My Groups | Groups Directory | Create a Group | FAQ

User Communications

Overview | Discussions | Updates | Members | Settings | Manage Group Profile

Discussions

Recent Discussions | Recent Comments | Most Comments

- What's the best strategy for promoting your brand to a broad community of users?**
From: Ben Guthrie | 1 hour ago | See 1 comment
- When you travel, do you choose a hotel based on services available, location or ambience? Or do you pass hotels altogether and opt to rent a short term apt instead?**
From: Florencia Pettigrew | 1 hour ago | See 1 comment
- Have you noticed that we only show house ads on this page?**
From: Richard Chen | 1 hour ago | See 1 comment
- What do you think of the new Intuit poll on LinkedIn?**
From: Krista Canfield | 1 hour ago | See 2 comments
- Feel free to share any best practices on user communications?**
From: Mario Sundar | 2 hours ago | See 1 comment

My discussions | My comments | See all discussions »

Start a Discussion

Enter a topic or question to discuss with the members of this group

Add details (Optional)

Submit for Discussion

Connect Your Brand To Over 1 Million Sales Professionals

Advertise On LinkedIn Learn More

Recent Updates

Today

- Mario Sundar commented on "What's the best strategy for promoting your brand to a broad community of users?" from Ben Guthrie. See 1 comment »
- Ben Guthrie commented on "Have you noticed that we only show house ads on this page?" from Richard Chen. See 1 comment »
- Mario Sundar commented on "When you travel, do you choose a hotel based on services available, location or ambience? Or do you pass hotels altogether and opt to rent a short term apt instead?" from Florencia Pettigrew. See 1 comment »
- Krista Canfield and Mario Sundar commented on "What do you think of the new Intuit poll on LinkedIn?" from Krista Canfield. See 2 comments »
- Florencia Pettigrew commented on "Feel free to share any best practices on user communications?" from Mario Sundar. See 1 comment »

See all updates »

OVERVIEW OF LINKEDIN

LINKEDIN TODAY · MAR 17
The most shared news on LinkedIn Take the tour » Powered by **Linked**

FRONT PAGE IT ONLINE MEDIA INTERNET COMPUTER... MARKETING &... DESIGN BROWSE ALL SA

Hiring Managers Take Their Time Filling Jobs

Recruiters say they are having trouble finding candidates for many skilled positions, and once candidates are found, hiring managers are taking longer to pull the trigger.

online.wsj.com

Share Save

TOP SOURCES

- online.wsj.com »
- money.cnn.com »
- npr.org »
- bloomberg.com »
- hbr.org »

SUGGESTED INDUSTRIES

Startup Veterans Launch Imagine K12, A Y...
Groupon Is Said to Discuss IPO Valuation
HTC Thunderbolt Now Available

Groups

SUGGESTED INDUSTRIES

IT INDUSTRY SEE ALL »

- Microsoft More "Ethical" Than...
searchengineland.com
90 shares
- Sharing Everything I Learned As A...
seomoz.org
107 shares

ONLINE MEDIA INDUSTRY SEE ALL »

- A Letter to Our Readers
lyfines.com
306 shares
- New York Times Agrees To Apple Terms...
mediamemo.alltagged.com
123 shares

INTERNET INDUSTRY SEE ALL »

- Whom Should You Hire at a Startup?...
techcrunch.com
702 shares
- Will You Pay for the New York Times...
marketable.com
323 shares

Groups My Groups | Groups Directory | Create a Group | FAQ

User Communications

Overview | Discussions | Updates | Members | Settings | Manage Group Profile

Discussions

Recent Discussions | Recent Comments | Most Comments

What's the best strategy for promoting your brand to a broad community of users?
From: Ben Guthrie | 1 hour ago | See 1 comment

When you travel, do you choose a hotel based on services available, location or ambience? Or do you pass hotels altogether and opt to rent a short term apt instead?
From: Florencia Pettigrew | 1 hour ago | See 1 comment

Have you noticed that we only show house ads on this page?
From: Richard Chen | 1 hour ago | See 1 comment

What do you think of the new Intuit poll on LinkedIn?
From: Krista Canfield | 1 hour ago | See 2 comments

Feel free to share any best practices on user communications?
From: Mario Sundar | 2 hours ago | See 1 comment

My discussions | My comments | See all discussions »

Start a Discussion

Enter a topic or question to discuss with the members of this group

Add details (Optional)

Submit for Discussion

See all updates »

Connect Your Brand To Over 1 Million Sales Professionals
Advertise on LinkedIn Learn More

Recent Updates

Today

Mario Sundar commented on "What's the best strategy for promoting your brand to a broad community of users?" from Ben Guthrie. See 1 comment »

Ben Guthrie commented on "Have you noticed that we only show house ads on this page?" from Richard Chen. See 1 comment »

Mario Sundar commented on "When you travel, do you choose a hotel based on services available, location or ambience? Or do you pass hotels altogether and opt to rent a short term apt instead?" from Florencia Pettigrew. See 1 comment »

Krista Canfield and Mario Sundar commented on "What do you think of the new Intuit poll on LinkedIn?" from Krista Canfield. See 2 comments »

Florencia Pettigrew commented on "Feel free to share any best practices on user communications?" from Mario Sundar. See 1 comment »

Linked in iPad

Filter by

Mynetwork

- By Me
- 1st Connections
- 2nd Connections
- 3rd + Everyone

Industry

- Information Technology And

143,583 results

Show 2 more recent results

Rachel Griffiths Win 1 of 10 iPads from @LightCMS - RT to enter or read more http://speaklight.com/iPad

10 iPads In 10 Days | LightCMS
Win one of ten iPads from LightCMS! Just tweet to enter.

less than a minute ago

Wrectacular College Lifestyle Festival if you haven't already downloaded the official WREC app, get it today for FREE! http://fb.me/GQyNTYoE

Chico State Wrec Gym for iPhone, iPod touch, and iPad on the iTunes App Store
Learn more, read reviews, and download Chico State Wrec Gym by Business Apps on the iTunes App Store.

less than a minute ago

Petra Belerling E-textbooks expected to grow with iPad on campus - Yahoo! News
http://yhoo.it/CSQsFd #edchat #edtech

E-textbooks expected to grow with iPad on campus - Yahoo! News
Video and color graphics are key to why many people think Apple Inc's iPad and other tablets will usher in the era of electronic textbooks.

less than a minute ago

Zhijie Lai 会备份app附带的数据库。有个小软件可以把备份支持。RT @motodexter: iTunes连上iPad后有时备份的速度是按小时为单位的。怎么会是呢? ! #ipadcn #itunes

L. Michelle Humphress iOS 4.2: The iPad is a Real Computer Now: Apple's iPad might be "magical," but of late it's been feeling a little ...

Trending Links

10 iPads In 10 Days | LightCMS
3997 members

10 iPads In 10 Days | LightCMS
1224 members

Making Future Magic: iPad light painting
617 members

Get the new Newsway iPad application
555 members

RevTwt - Twitter Advertising - Sign Up For Advertiser
347 members

Kaleidoscope Heart by Sara Barelles - Download Kaleidoscope Heart on iTunes
Who shared this link

PROBLEM DEFINITION

For a given recipient and updates from his/her social connections (senders), we want to re-rank these updates to optimize user engagement.



DATASET

Data Summary	April, 2011	September, 2011
Impressions	3M-4M	10M-20M
Updates	30M-40M	100M-200M
Clicked Updates	3M-4M	10M-20M
Non-clicked Updates	27M-36M	90-180M
Distinct Updates	10M-20M	20M-30M
Recipients	1M-2M	4M-5M
Producers	4M-5M	6M-7M

The numbers are obfuscated for commercial reason.

EVALUATION METRIC

- Precision@ k

- $$\frac{\text{\# of clicks in top } k \text{ positions}}{k}$$

- Average Precision (AP) for ranked list i

- $$\frac{\sum_{k=1}^m \text{Precision@ } k \times l_k}{\text{\# of clicks for ranked list of ranked list } i}$$

- l_k : position k is clicked.

- m : total number of positions evaluated.

- Mean Average Precision (MAP)

- average AP across all ranked lists

METHODS

- Linear Models
 - Feature Model
 - Bias Model
 - Hybrid Model
- Latent Factor Models
 - Matrix Factorization
 - Tensor Factorization
 - Regression-based Tensor Factorization

METHODS

- Linear Models
 - Feature Model
 - Bias Model
 - Hybrid Model
- Latent Factor Models
 - Matrix Factorization
 - Tensor Factorization
 - Regression-based Tensor Factorization



From the simplest to
the most complex

METHODS

Linear Models: Feature Model

$$f_i^{(1)} = \beta_{r(i)}^T \phi_{r(i)} + \alpha_{r(i)}^T \phi_i$$

- utilize explicit features.
- f_i represents the estimation of user's click on update i .
- $r(i)$ is the recipient of update i .
- ϕ is a feature vector.
- β and α are coefficients.

METHODS

Linear Models: Latent Bias Model

$$f_i^{(2)} = \mu + b_i + b_{t(i)} + b_{r(i)} + b_{c(i)} + b_{s(i)}$$

- utilize categorical features.
- $t(i)$ is the type of update i .
- $c(i)$ is the type of sender of update i .
- $s(i)$ is the sender of update i .

METHODS

Linear Models: Latent Bias Model

$$f_i^{(2)} = \mu + b_i + b_{t(i)} + b_{r(i)} + b_{c(i)} + b_{s(i)}$$

- utilize categorical features.
- $t(i)$ is the type of update i .
- $c(i)$ is the type of sender of update i .
- $s(i)$ is the sender of update i .

METHODS

Combining Feature and Bias

$$f_i^{(3)} = f_i^{(1)} + f_i^{(2)}$$

Incorporating Temporal Effects

$$f_i^4 = f_i^{(*)} + \zeta \times t_{\text{recency}}$$

METHODS

Learning through L_2 -regularized logistic regression

$$l_1(y_i, f_i^{(*)}) = \log \left[1 + \exp(-y_i f_i^*) \right]$$

METHODS

Linear Model Summary

- Simple
- Fast
- Intuitive

METHODS

Linear Model Summary

- Simple
- Fast
- Intuitive

Does not exploit user-user, user-item interactions at all

METHODS

Latent Factor Model: Matrix Factorization

How to utilize pair-wise interactions?

METHODS

Latent Factor Model: Matrix Factorization



- user-item interaction?

		
	?	?
	?	?
	?	?
	?	?

METHODS

Latent Factor Model: Matrix Factorization

- user-user interaction?

		
	?	?
	?	2
	1	?
	?	4

METHODS

Latent Factor Model: Matrix Factorization

- user-user interaction?
- user-item interaction?

METHODS

Latent Factor Model: Matrix Factorization

$$f_i = \mu + b_i + b_{t(i)} + b_{r(i)} + b_{c(i)} + b_{s(i)} + \boldsymbol{\eta}_{r(i)}^T \boldsymbol{\eta}_{s(i)}$$

METHODS

Latent Factor Model: Matrix Factorization

$$f_i = \boxed{\mu + b_i + b_{t(i)} + b_{r(i)} + b_{c(i)} + b_{s(i)}} + \boxed{\boldsymbol{\eta}_{r(i)}^T \boldsymbol{\eta}_{s(i)}}$$



Latent Bias Model



Matrix Factorization

- very similar to basic MF model used in SVD++

[Koren 2010]

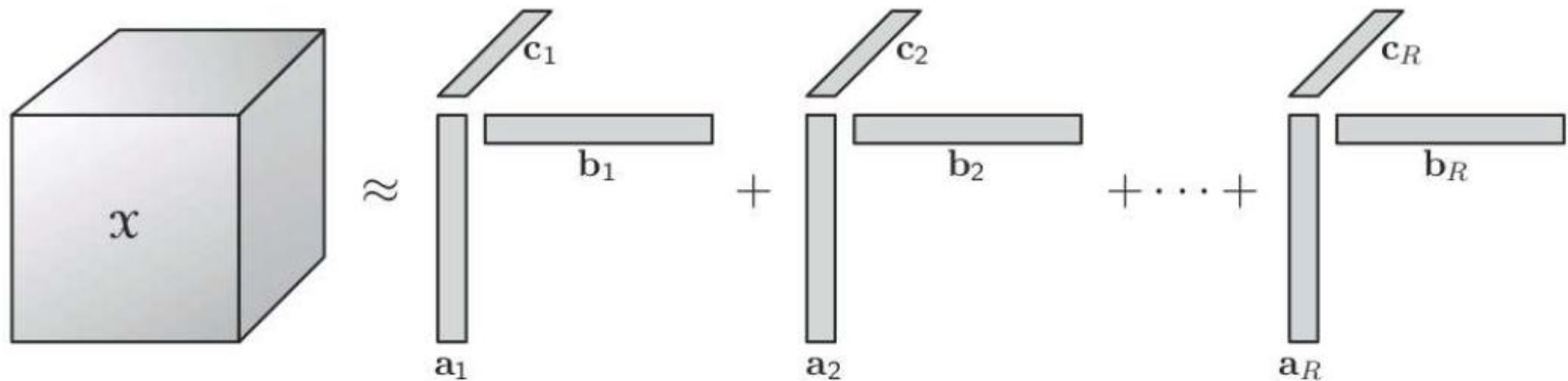
METHODS

Higher-order interactions?



METHODS

Latent Factor Model: Tensor Factorization




- Recipient-Type-Sender relationships
- CP decomposition

METHODS

Latent Factor Model: Tensor Factorization

$$f_i = \underbrace{\mu + b_i + b_{t(i)} + b_{r(i)} + b_{c(i)} + b_{s(i)}}_{\text{Latent Bias Model}} + \underbrace{\sum_k \eta_{r(i),k} \eta_{s(i),k} \eta_{t(i),k}}_{\text{Tensor Factorization}}$$



Latent Bias Model

Tensor Factorization

METHODS

How about other explicit features?

METHODS

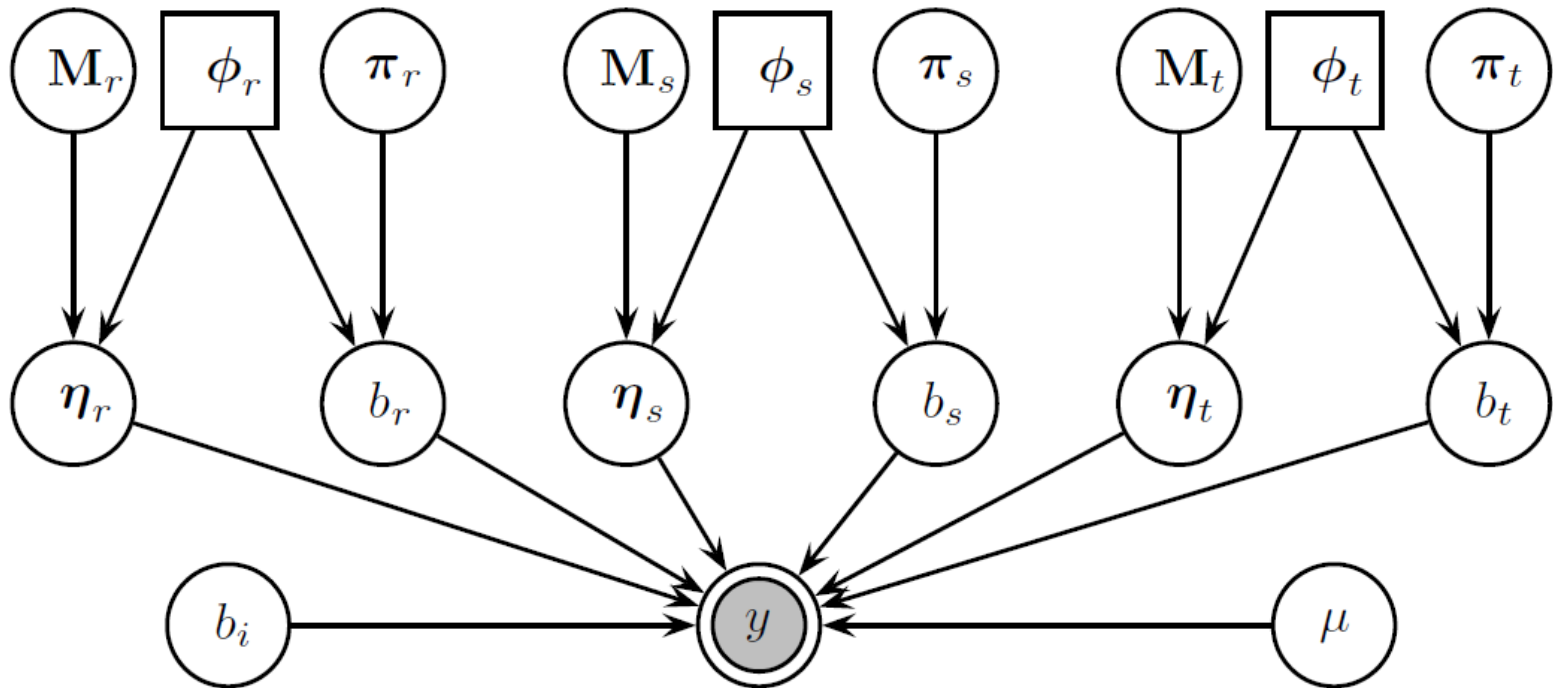
How about other explicit features?

- Regression-based latent factor models
 - another layer of regression
 - replacing zero-mean with regression-based mean

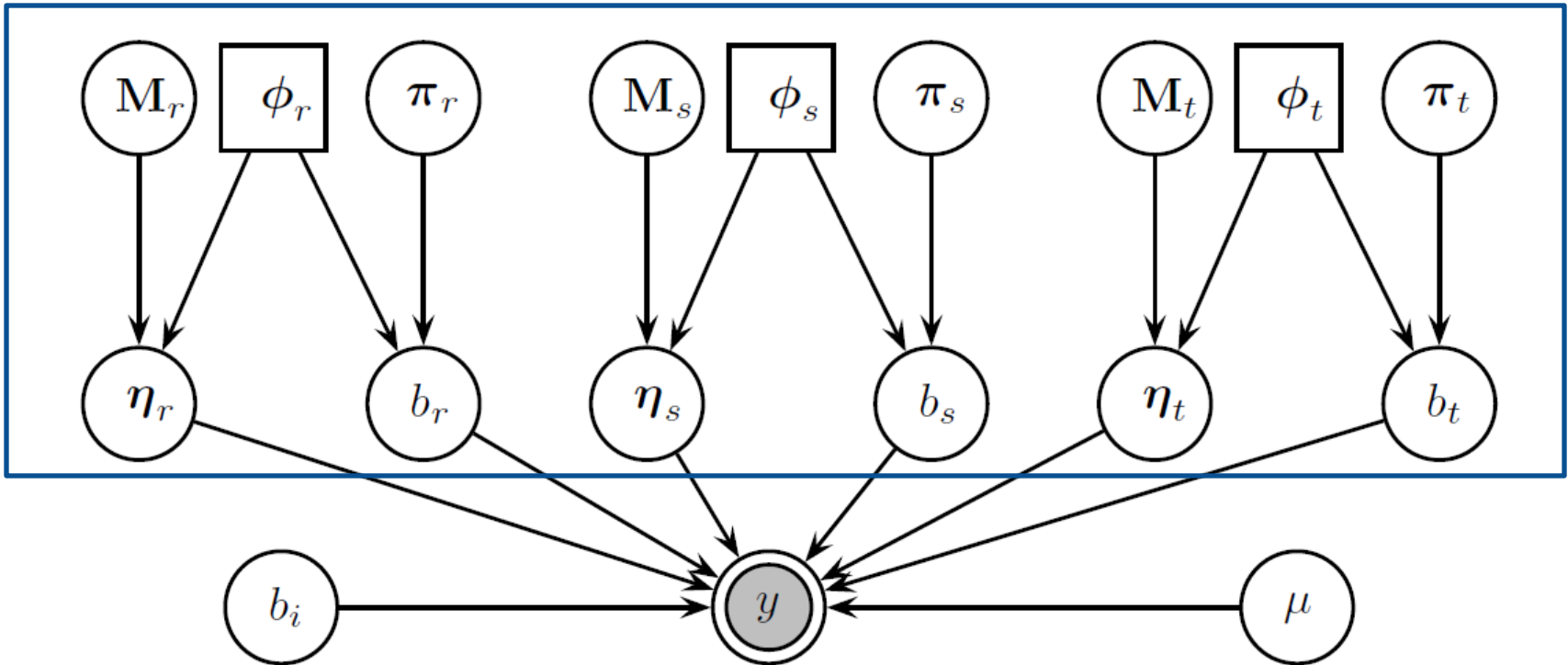
$$\eta_{x(*)} = \mathbf{M}_x \phi_{x(*)} + \epsilon_x \quad x \in \{\mathcal{R}, \mathcal{S}, \mathcal{T}\}$$

$$b_{x(*)} = \boldsymbol{\pi}_x^T \phi_{x(*)} + \epsilon_{b_x}$$

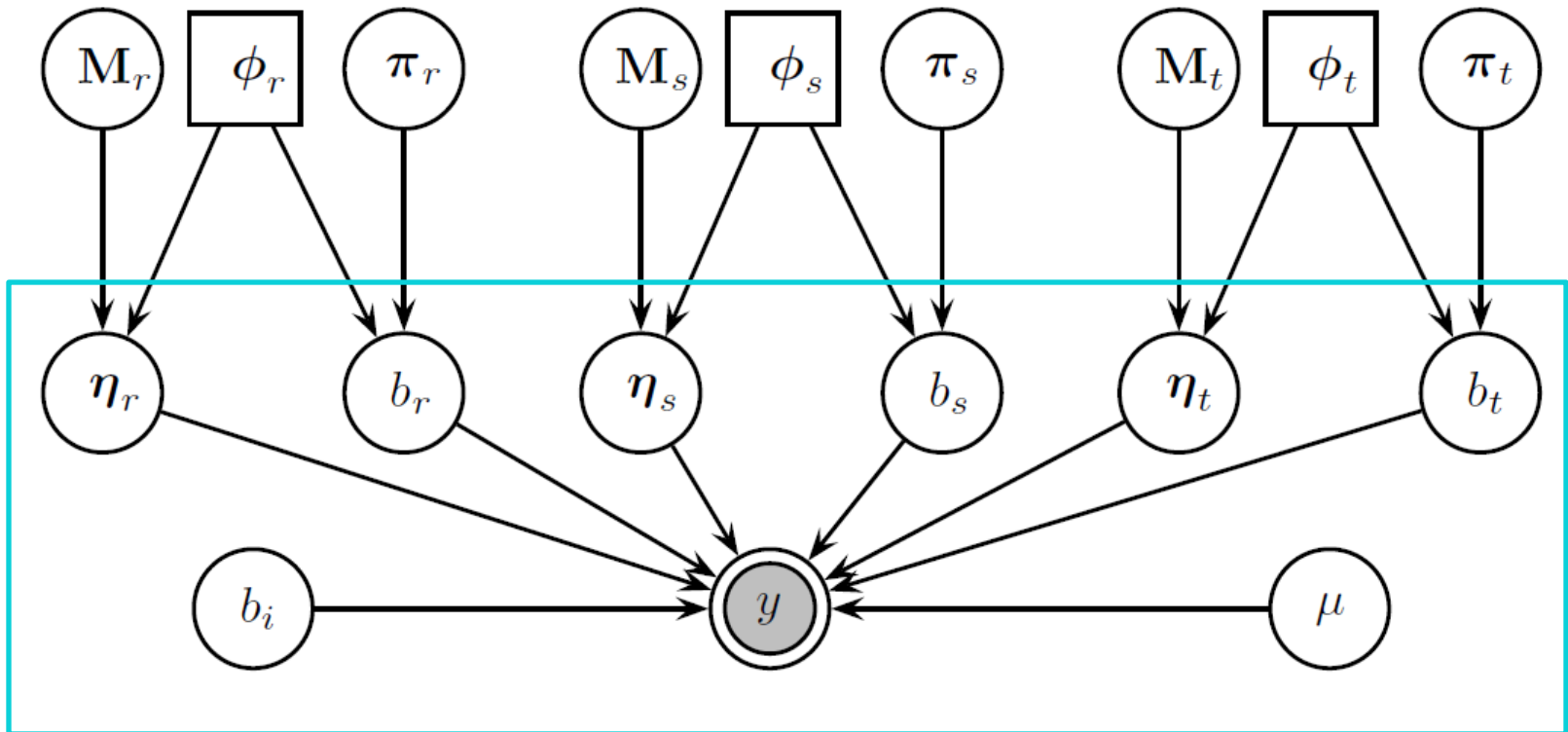
METHODS



METHODS



METHODS



METHODS

Learning procedure

- Maximum A Posterior (MAP)
- Stochastic Gradient Descent

METHODS

Going beyond pointwise learning

- Optimizing Bayesian Personalized Ranking (BPR)

$$\sum_{m \in \mathcal{O}_{i,+}} \sum_{n \in \mathcal{O}_{i,-}} \sigma(f_m - f_n)$$

[Rendle et al. 2009]

EXPERIMENTS

Models

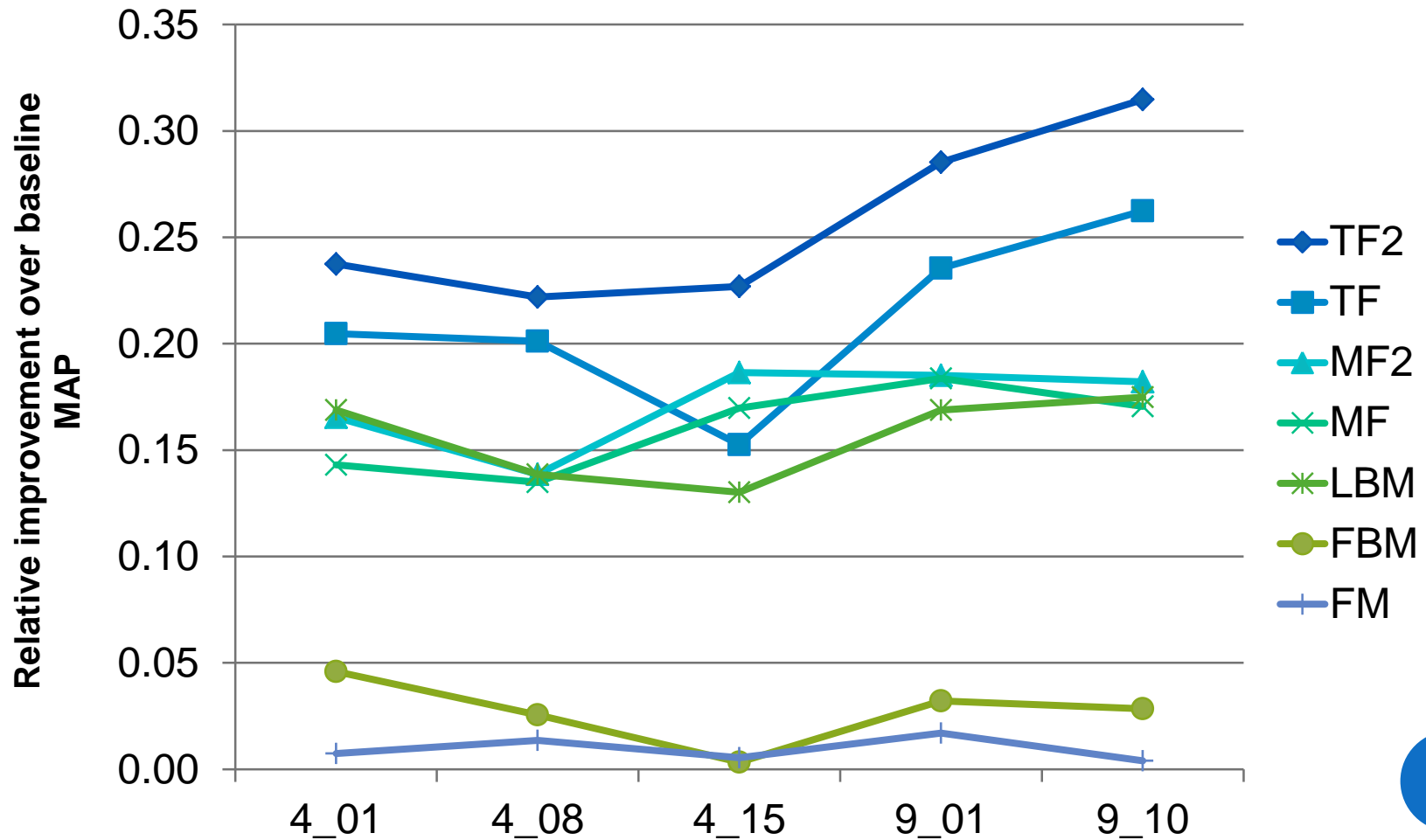
Methods
Baseline (BL)
Feature Model (FM)
Latent Bias Model (LFM)
Feature Bias Model (FBM)
Matrix Factorization (MF)
Tensor Factorization (TF)
Matrix Factorization with Features (MF2)
Tensor Factorization with Features (TF2)

EXPERIMENTS

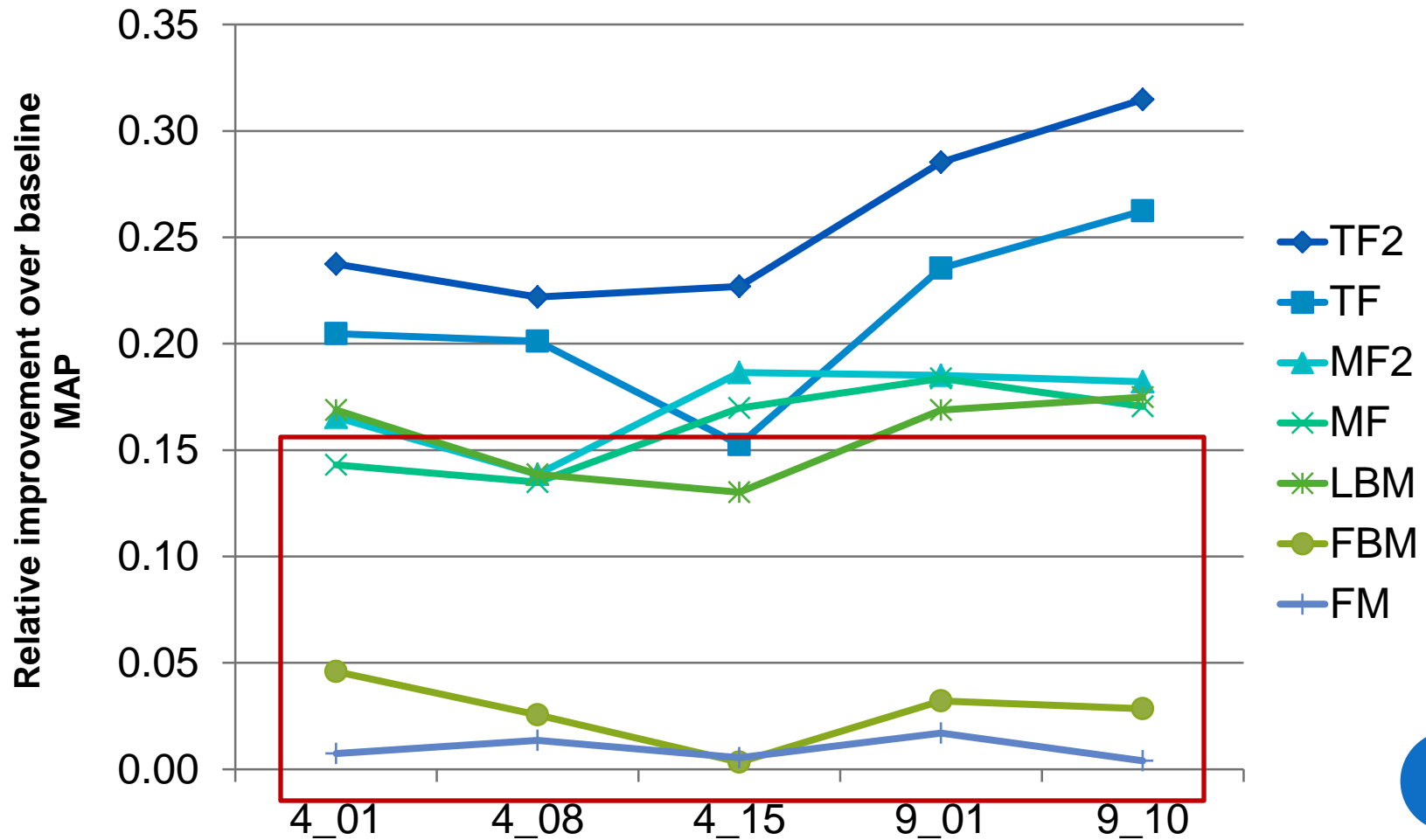
Models

Features	Comments
Seniority	the seniority level of a user
Visiting	how frequently a user visits LinkedIn
PageRank	discretized PageRank scores
Connectedness	how well a user is connected to others
Social strength	social strength between recipient and sender
Professionalism	how professional an update's language is
Recency	the freshness of an update

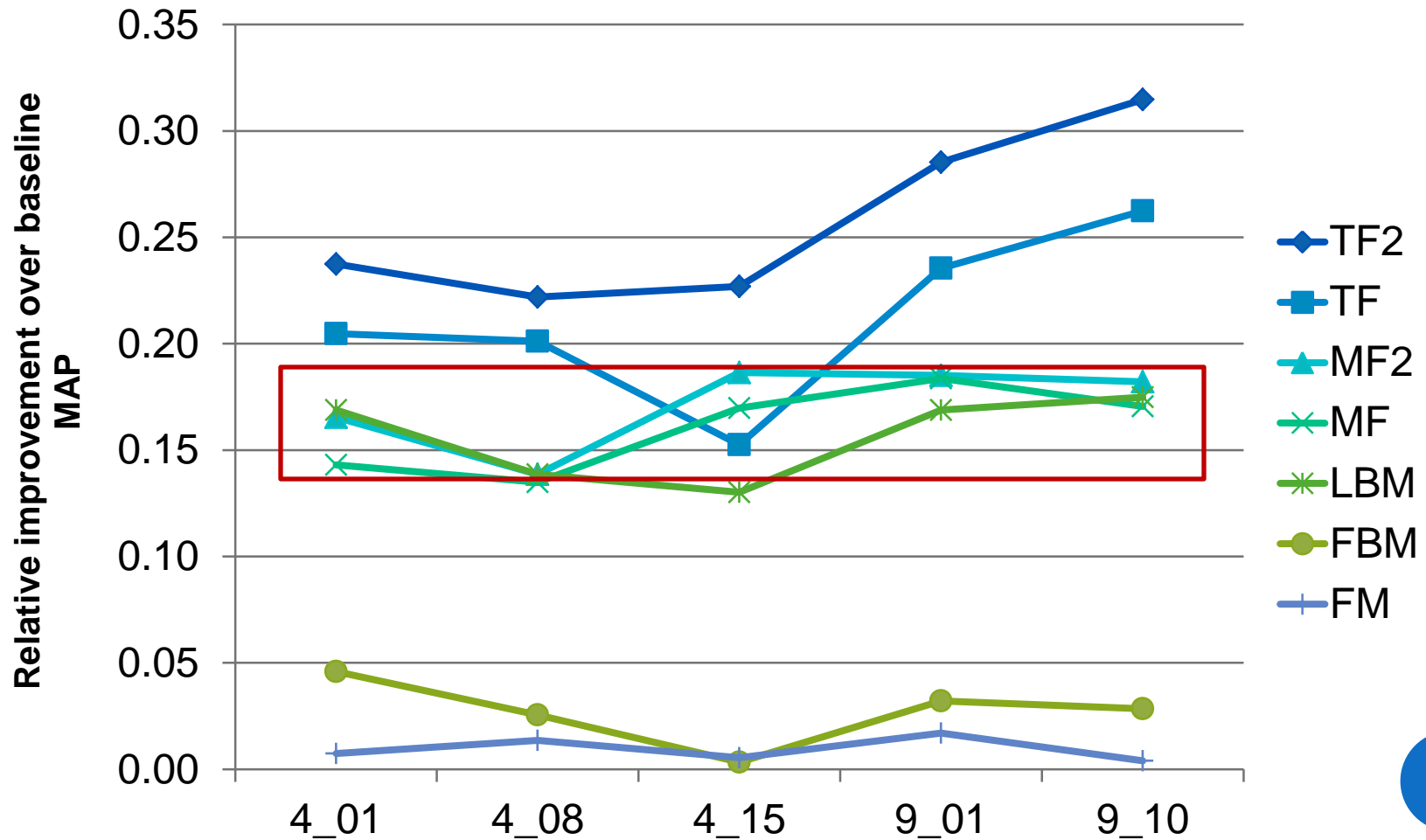
EXPERIMENTS: PAIRWISE LEARNING



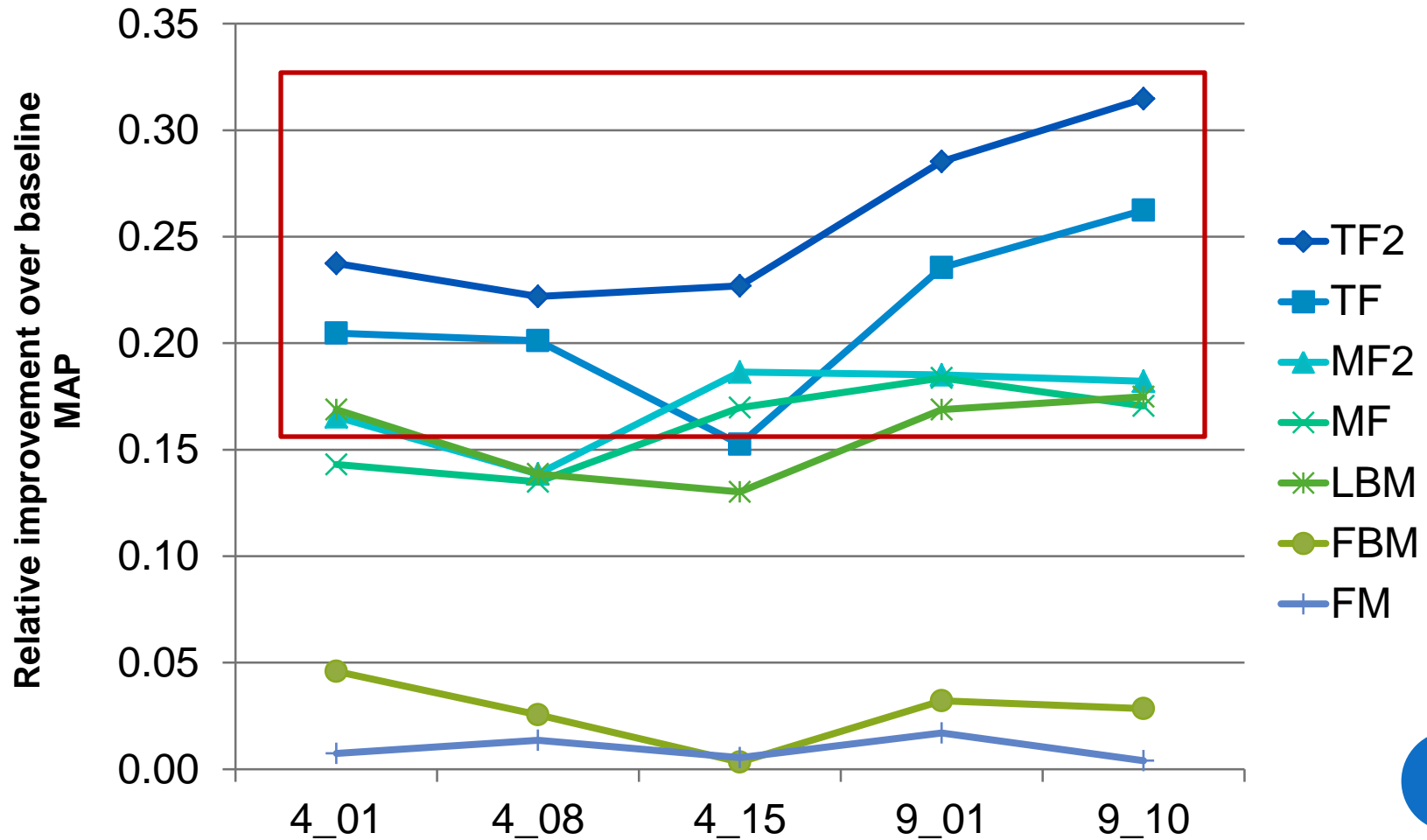
EXPERIMENTS: PAIRWISE LEARNING



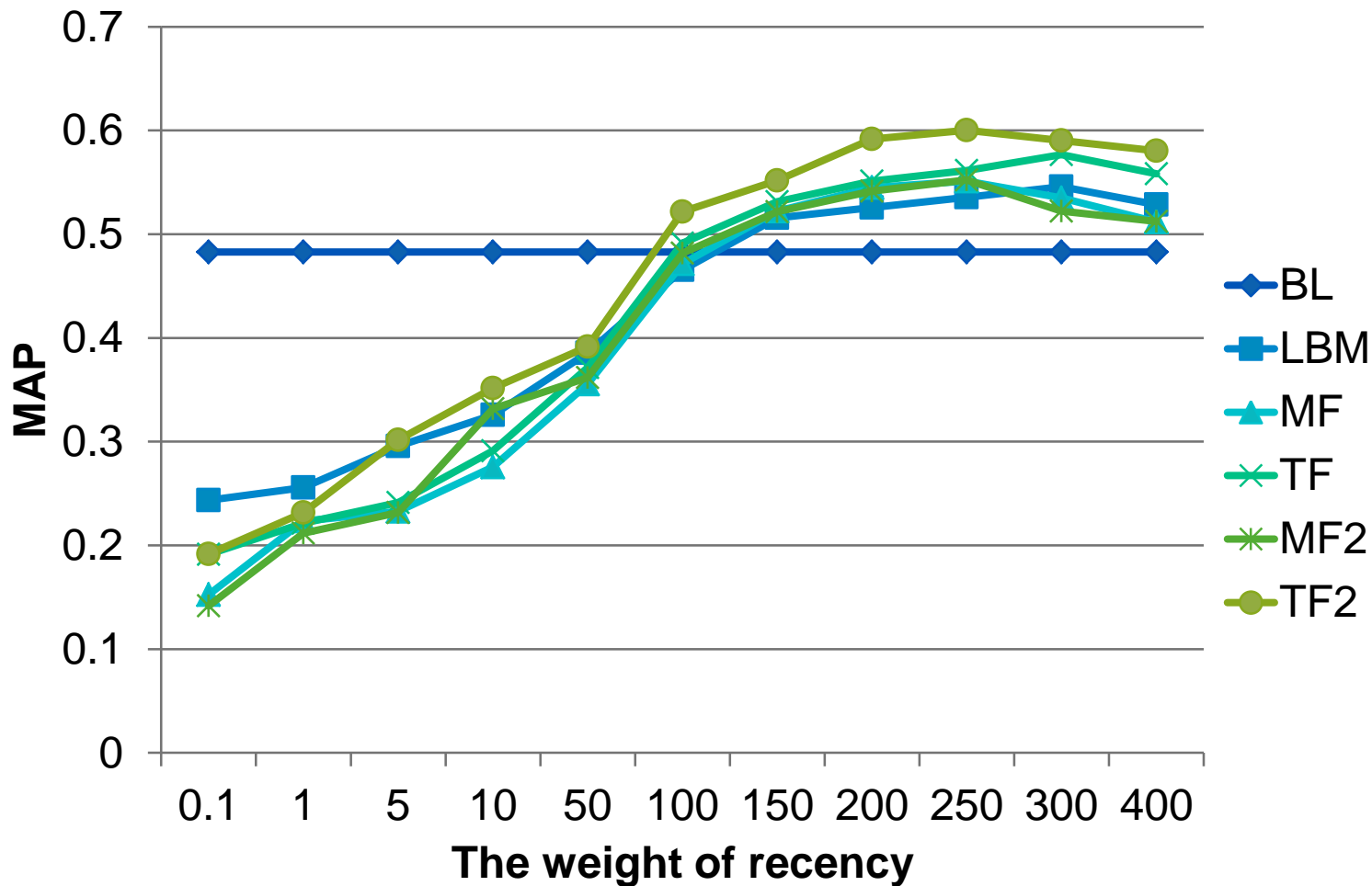
EXPERIMENTS: PAIRWISE LEARNING



EXPERIMENTS: PAIRWISE LEARNING



EXPERIMENTS: PARAMETER SENSITIVITY



EXPERIMENTS

Example of highly ranked types of updates

Type Description	Bias b_t
Job Seeker Product Update	0.5765
Joining Sub-Group	0.5407
Company News	0.4592
Joining Group	0.2625
Profile Picture Update	0.2516
Initiating Direct Ads Campaign	0.2253
Profile Update	0.1394

CONCLUSIONS

- We demonstrate that recommender systems + preference-based learning can be used to re-rank social updates.
- A linear model can achieve 60% of the performance of latent factor models, on average.
- A tensor factorization model with regression on explicit features works the best.
- The cold-start problem makes it impossible to model some kinds of interactions.

THANK YOU.



Liangjie Hong

PhD candidate

WUME Lab

Lehigh University

lih307@cse.lehigh.edu

EXPERIMENTS: COMPARISON

Training/Testing	BL	FM	LBM	FBM
4_01(Tr.)/4_08(Te.)	0.5278	0.5317	0.5943	0.5520
4_08(Tr.)/4_15(Te.)	0.5435	0.5509	0.6040	0.5574
4_15(Tr.)/4_22(Te.)	0.5218	0.5246	0.5823	0.5235
9_01(Tr.)/9_10(Te.)	0.4829	0.4911	0.5457	0.4984
9_10(Tr.)/9_18(Te.)	0.4779	0.4798	0.5432	0.4915
9_18(Tr.)/9_25(Te.)	0.4768	0.4803	0.5329	0.4886

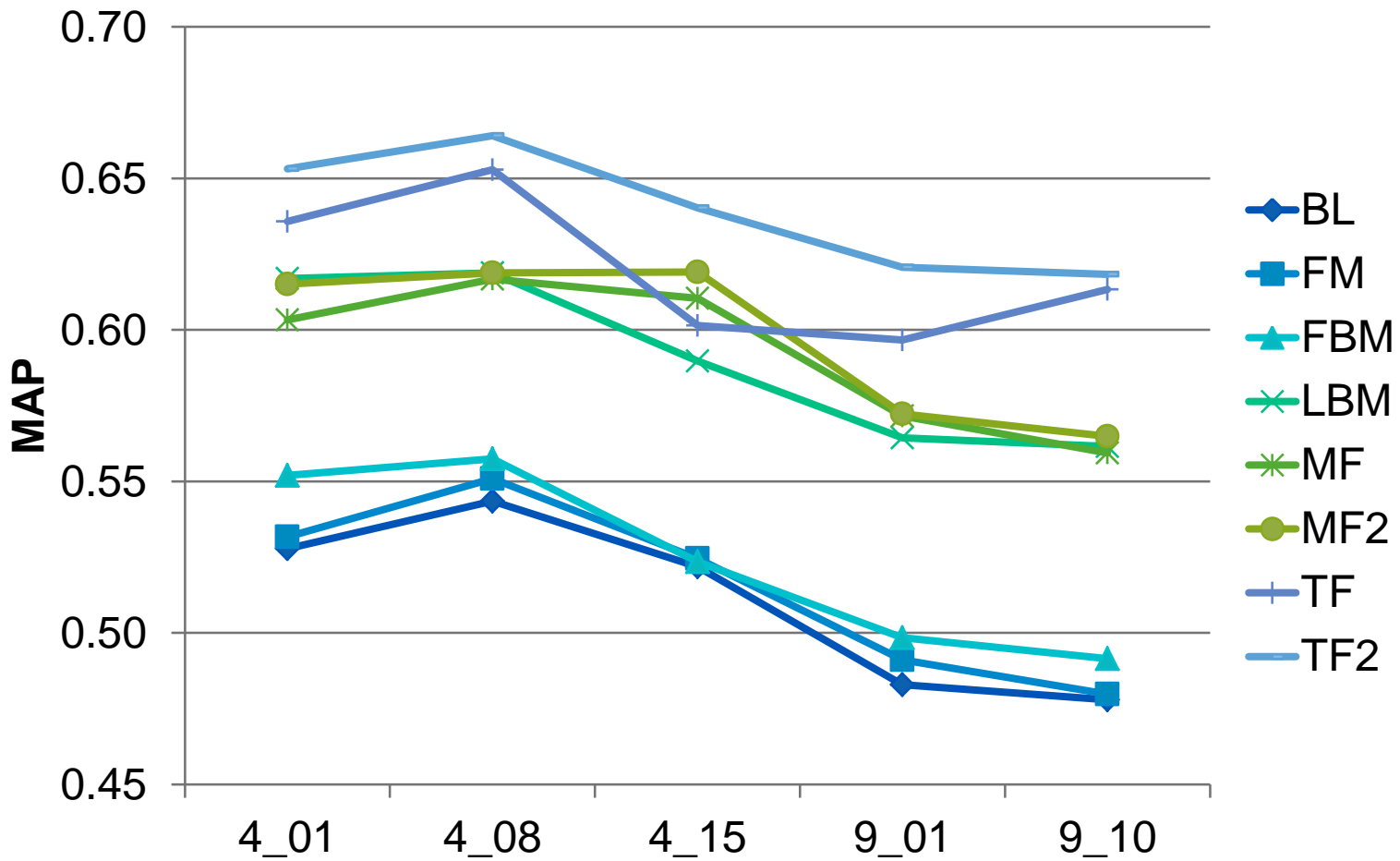
Training/Testing	MF	TF	MF2	TF2
4_01(Tr.)/4_08(Te.)	0.5955	0.6258	0.5951	0.6336
4_08(Tr.)/4_15(Te.)	0.6079	0.6228	0.6088	0.6535
4_15(Tr.)/4_22(Te.)	0.5962	0.6014	0.5991	0.6312
9_01(Tr.)/9_10(Te.)	0.5511	0.5766	0.5523	0.6003
9_10(Tr.)/9_18(Te.)	0.5412	0.5833	0.5449	0.6109
9_18(Tr.)/9_25(Te.)	0.5359	0.5799	0.5362	0.5992

EXPERIMENTS

The effects of pairwise learning

Training/Testing	LBM	MF	MF2	TF	TF2
4_01(Tr.)/4_08(Te.)	0.6169	0.6033	0.6151	0.6358	0.6532
4_08(Tr.)/4_15(Te.)	0.6188	0.6168	0.6188	0.6528	0.6641
4_15(Tr.)/4_22(Te.)	0.5897	0.6104	0.6191	0.6014	0.6402
9_01(Tr.)/9_10(Te.)	0.5644	0.5716	0.5723	0.5966	0.6207
9_10(Tr.)/9_18(Te.)	0.5593	0.5621	0.5607	0.5999	0.6183

EXPERIMENTS



EXPERIMENTS

Parameter Sensitivity

