Context-aware Image Tweet Modelling and Recommendation

Tao Chen, Xiangnan He and Min-Yen Kan
School of Computing, National University of Singapore

Introduction

Research Questions

- How to represent the semantics of an image tweet?
  - Are visual objects sufficient?
- How to utilize such representations for personalized image tweet recommendation task?

Our Contributions

- We propose a CITING framework to model image tweet by its contextual text
- We develop a feature-aware matrix factorization model to capture user's personal interest
- We have released code and datasets: https://github.com/kite1988/famf

CITING: Context-aware Image Tweet Modelling

1. Hashtag enhanced text

Word Breaker

14.3%

- Conflate the variants of hashtags
- #icebucket, #ALSIceBucketChallenge

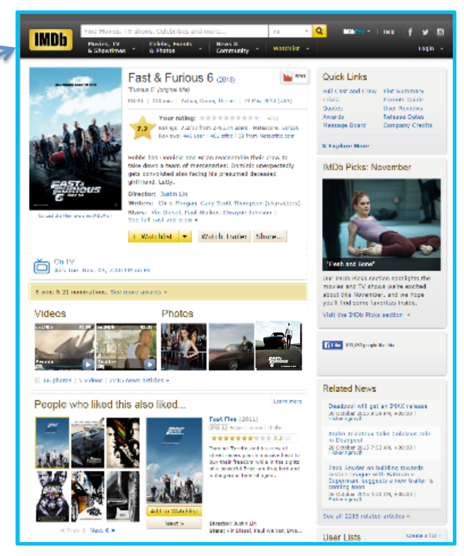
ice bucket ALS ice bucket challenge

Coming soon!!!
imdb.to/IGxE9f

URLs

22.7%

3. Extended context



82% contain the image

OCR Tool

24.6%

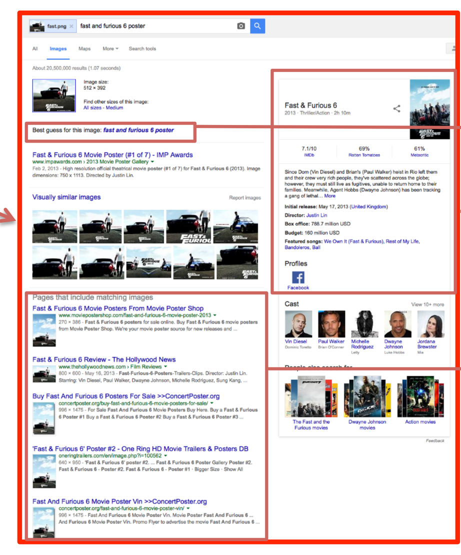
2. Text in Image

- Overlaid text and screenshot of pure text

Google Image Search

70.6%

4. Similar contexts



Best guess

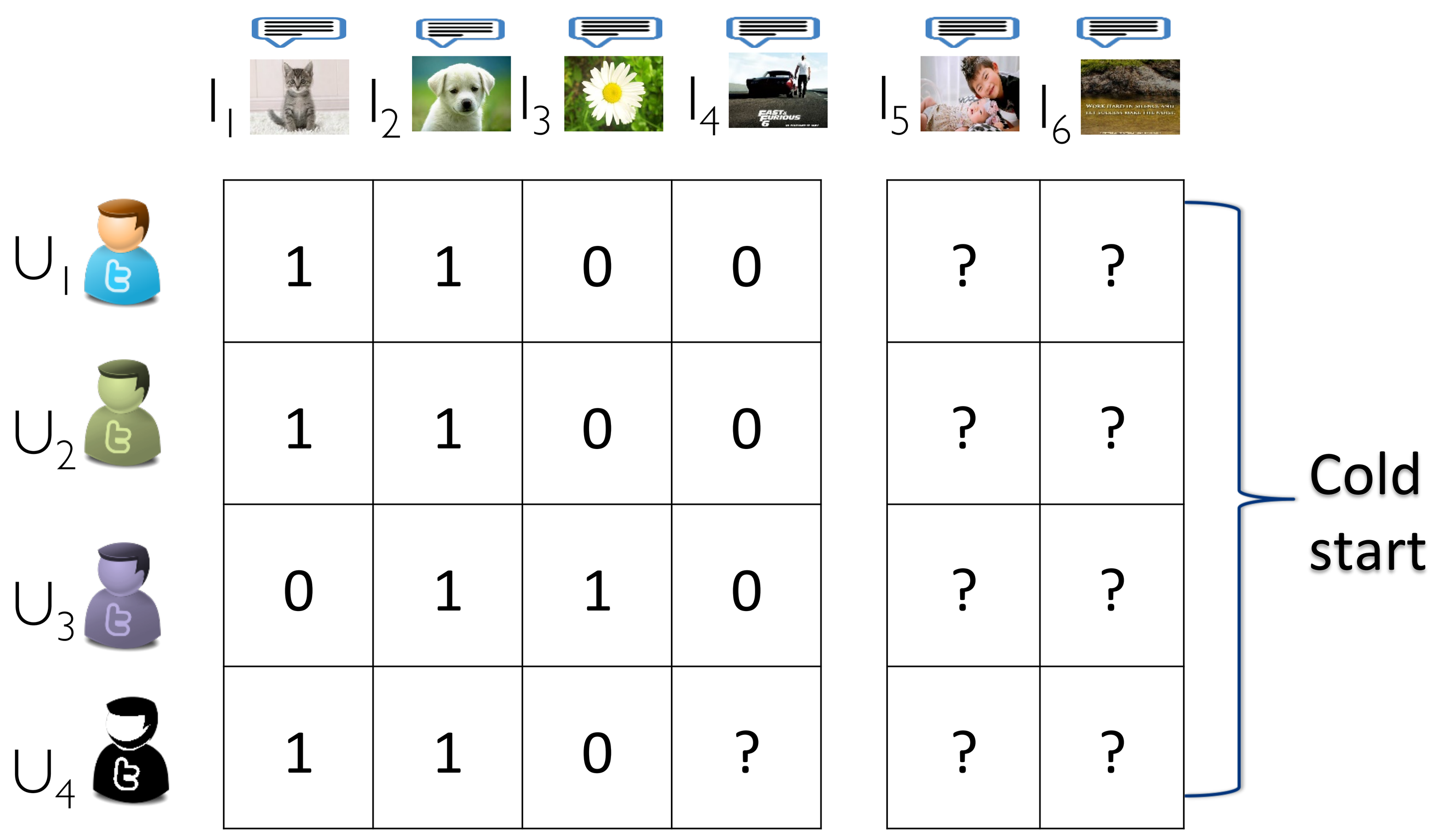
Named entity

Pages that contain the image

Propose filtered rules to fuse the four contextual text

- Text quality: 1>3>2>4
- Save acquisition cost by 18% and improve representation quality

Personalized Image Tweet Recommendation



- The traditional Matrix Factorization model does not work well due to serious cold start problems
- To alleviate this, we propose feature-aware Matrix Factorization (**FAMF**) to model user's interest
  - Decompose user-item interaction to user-feature interaction

User's latent factor

Item's latent factor

$$\hat{y}_{u,i} = v_u^T \Big(\sum_{n=1}^{N} \frac{1}{Z_{n,i}} \sum_{f \in F_{n,i}} q_f \Big)$$

N types of features (e.g., CITING text)

A feature's latent factor

- Pair-wise Learning to Rank
  - Positive tweets (retweets) should be ranked better than negative tweets (non-retweets)
  - Bayesian Personalized Ranking [Rendle et al. 2009]
  - Infer the parameters via stochastic gradient descent (SGD)

Experiments

Dataset from Twitter

	Twitter Users	Retweets	All Tweets	Ratings
Training	926	174,765	1,316,645	1,592,837
Test		9,021	77,061	82,743

Evaluation

- For each user, keep the recent 10 retweets as test set and the rest as training set
- Report mean average precision (MAP) and precision at top positions

	Method	Features	P@1	P@3	P@5	MAP
1	Random		0.114**	0.115	0.115	0.156**
2	Length	Post's text	0.176**	0.158	0.150	0.173**
3	Profiling	Post's text	0.336**	0.227	0.197	0.202**
4	FAMF	Visual objects	0.211**	0.205	0.192	0.211**
5	FAMF	Post's text	0.359*	0.325	0.287	0.275**
6	FAMF	Non-filtered context	0.413	0.352	0.319	0.296
7	FAMF	CITING	0.419	0.355	0.319	0.298
8	FAMF	CITING + Visual objects	0.425	0.350	0.313	0.298

**: p<0.01, *: p<0.05

- 4-8 vs. 1-3: FAMF is effective in modelling user interest
- 4 vs. 5: Visual objects are not sufficient to model Twitter image's semantics
- 7 vs. other: Our CITING text significantly outperforms the others
- 7 vs. 6: The filtered fusion rules improve text quality
- 8 vs. 7: The further incorporation of visual objects does consistently improve the performance

Tao Chen: chentaokite@gmail.com

ACM MM 2016, Amsterdam, The Netherlands