Final Report
Interest-aware Information Diffusion in Dynamic Social Network

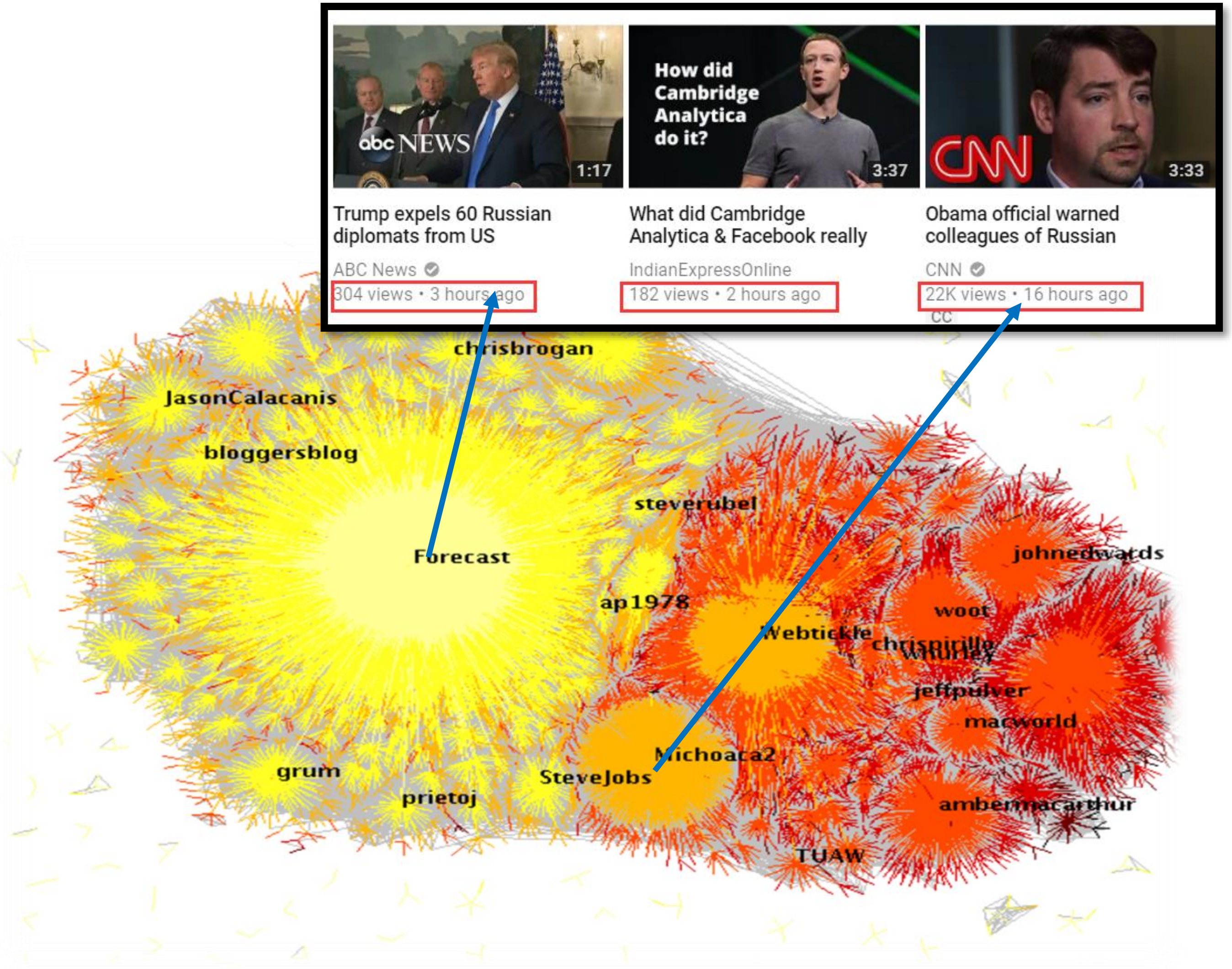
Zhenhao Cao
Ru Wang
Mobile Internet
2018. 6

Outline

- Introduction
- Related Work
- Challenge & Motivation
- Proposed Model
- Experiments
- References

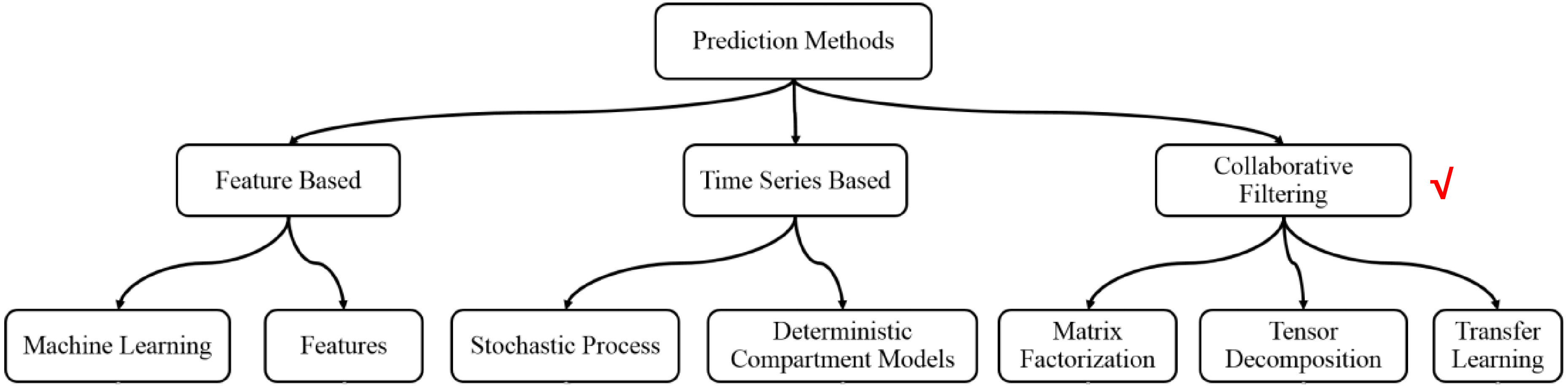
Introduction

- Social Network



Introduction - A Taxonomy

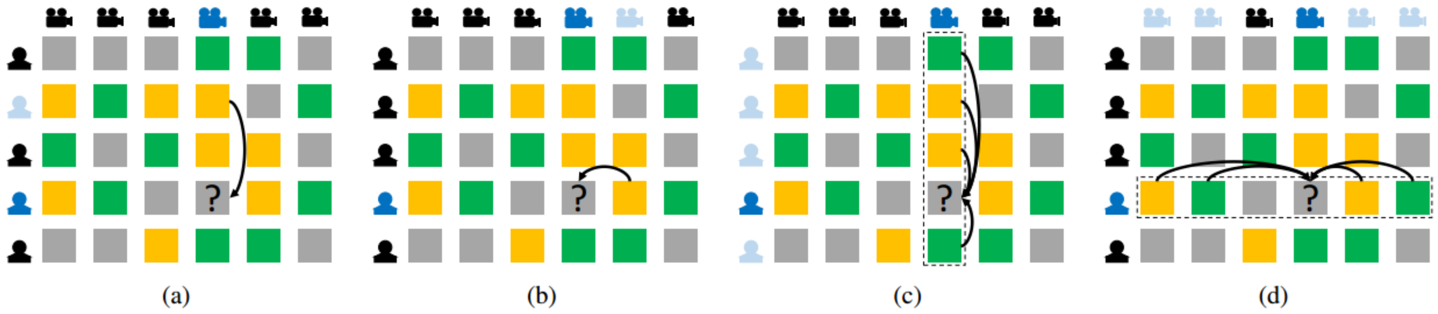
- An earlier survey: a taxonomy for information cascade prediction



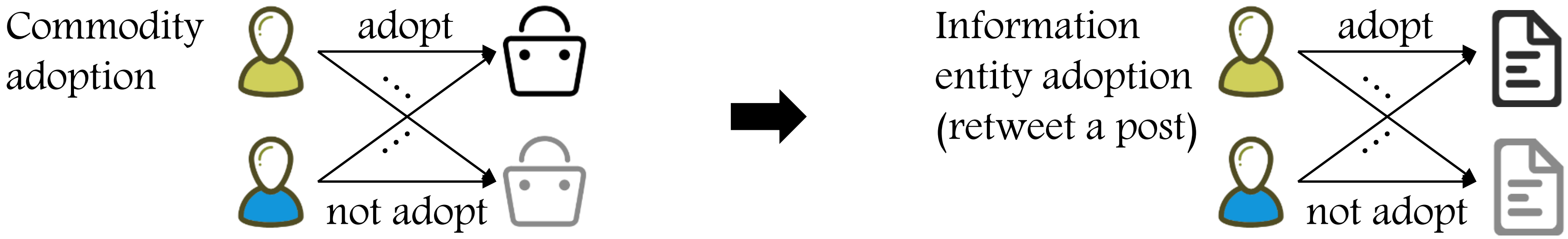
- Collaborative Filtering methods
  - Leverage homophily: insightful
  - Get rid of troublesome feature engineering

Introduction - Why CF?

- Key idea behind CF: Homophily

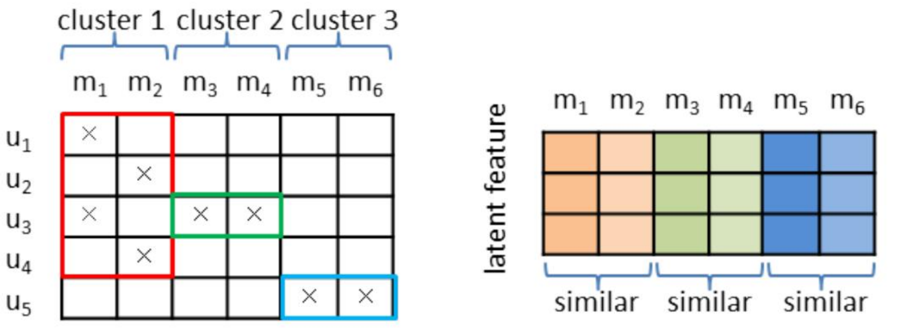


- Transplantable to information diffusion modeling



Related Work - Extant CF-based Studies

- CRPM & IRPM [1] (CIKM2015)

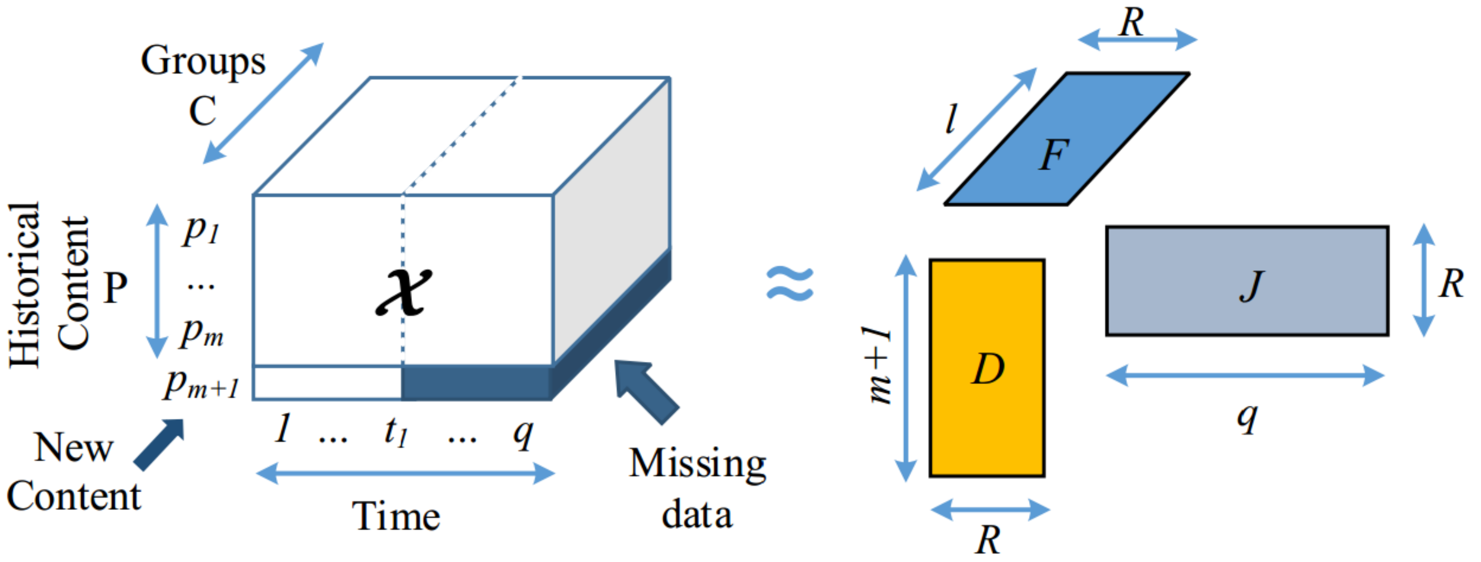


(a) user-message matrix (R) (b) latent feature-message matrix (V)

$$\min_{U,V} \mathcal{J}(R,U,V) = \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}(R_{ij} - U_iV_j)^2 + \frac{\alpha_1}{2}\sum_{j=1}^{N}\sum_{k=1}^{K} T_{jk}\|V_j - \frac{1}{|C_k|}\sum_{m_x \in C_k} V_x\|_F^2 + \frac{\beta}{2}\|U\|_F^2 + \frac{\gamma}{2}\|V\|_F^2$$

Related Work - Extant CF-based Studies

- GPOP [2] (WWW2017)



$$\mathcal{L} = \frac{1}{2}\|\mathcal{M} * (\mathcal{X} - [[D, J, F]])\|_F^2 + \frac{\lambda}{2}(\|D\|_F^2 + \|J\|_F^2 + \|F\|_F^2)$$

Related Work - Extant CF-based Studies

- A Collaborative Filtering Model for Personalized Retweeting Prediction [3] (DASFAA2015)

	T1	T2
U1	R11	R12
U2	R21	R22

$=$

	f1	f2
U1	P11	P12
U2	P21	P22

$\times$

	T1	T2
f1	Q11	Q12
f2	Q21	Q22

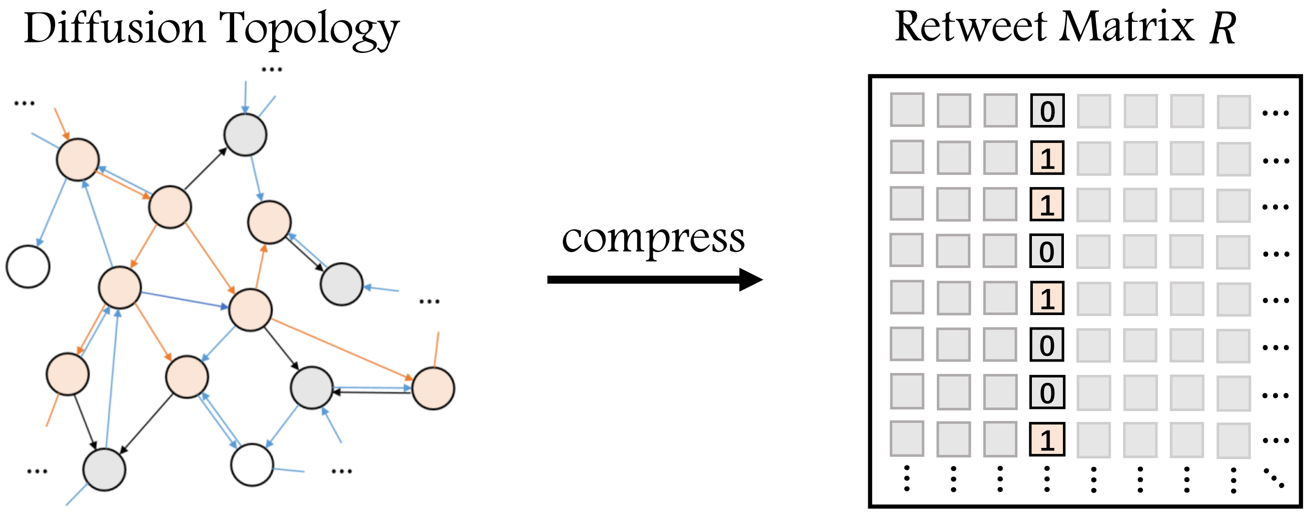
$$y_{u,i} = {p_u}^T \left(\frac{1}{Z} \sum_{w \in T_i} q_w + \alpha d_{p(i)}\right) + \frac{1}{Z} \sum_{w \in T_i} {q_w}^T \beta d_{p(i)}.$$

Challenge & Motivation

- More sufficient utility of social network information
- Better adapted for Information Diffusion modeling
- Novel insights into user retweet behavior

Challenge & Motivation

- More sufficient utility of social network information
  - A flat “snapshot” of users’ historical behaviors
  - Information loss: Permutation? Sequence? Diffusion topologies?



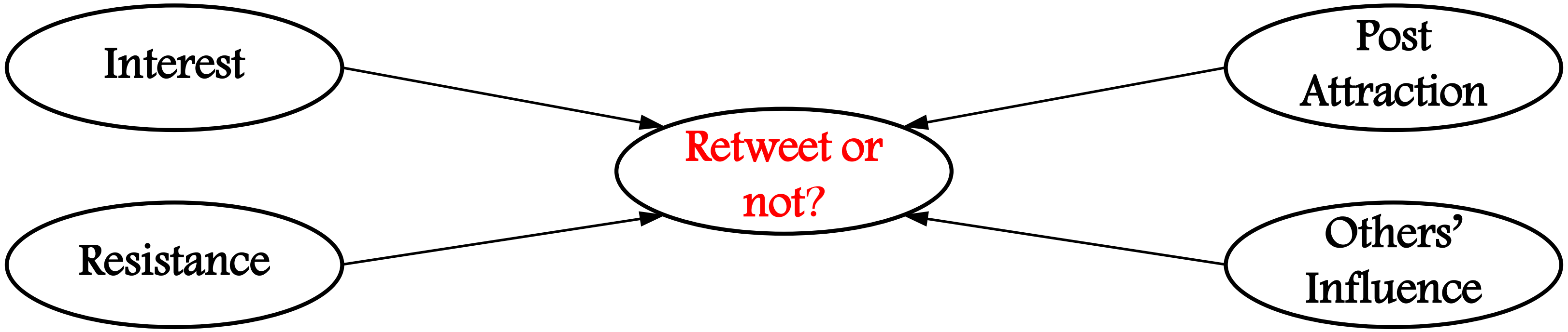
Challenge & Motivation

- More sufficient utility of social network information

- Better adapted for Information Diffusion modeling
  - Leverage diffusion topologies
    - Essence of information diffusion
    - A main difference from recommendation system problems

Challenge & Motivation

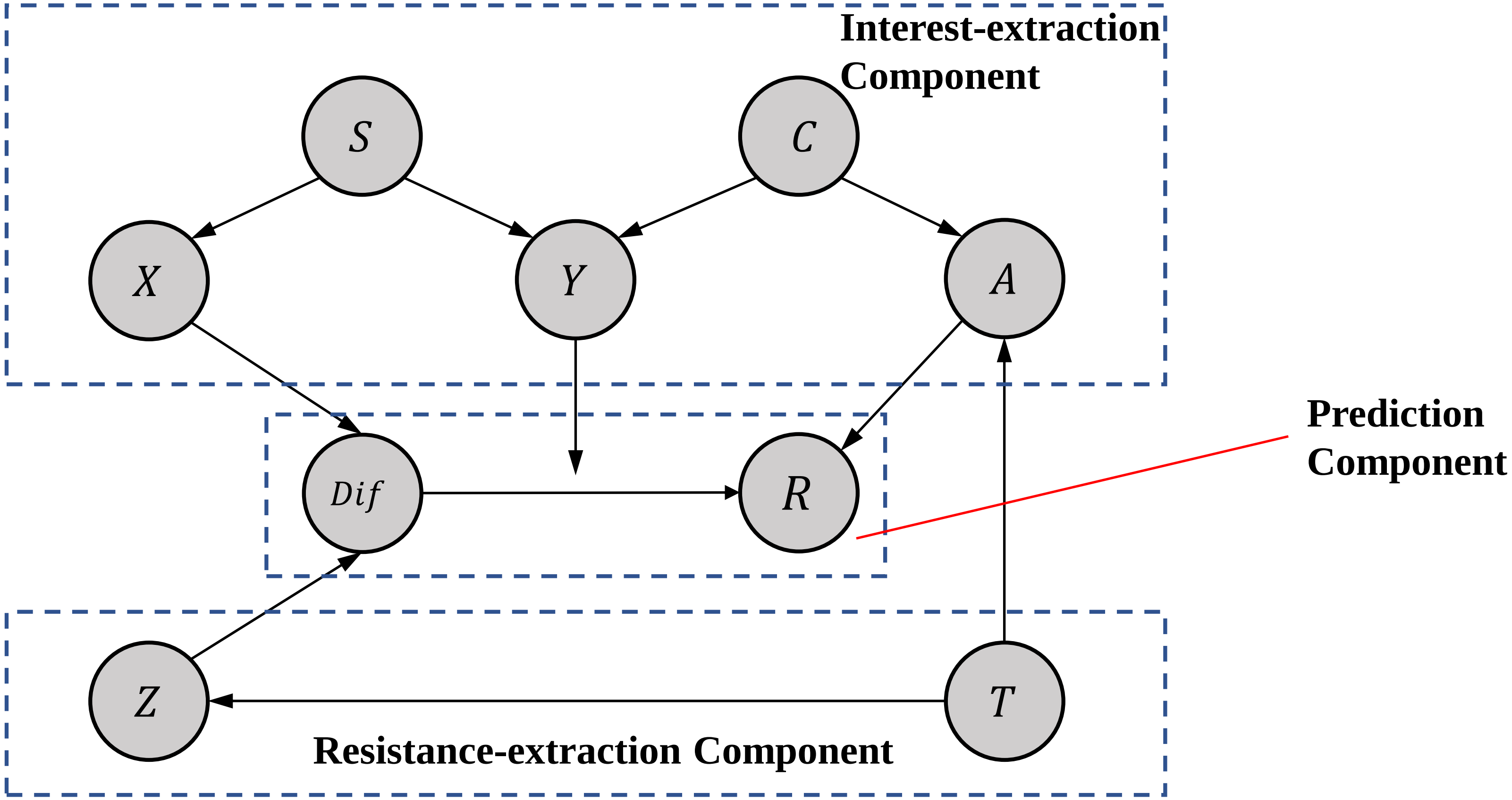
- More sufficient utility of social network information
- Better adaption to Information Diffusion modeling
- Novel insights into user retweet behavior



Our Work

Our Work - ReTrend

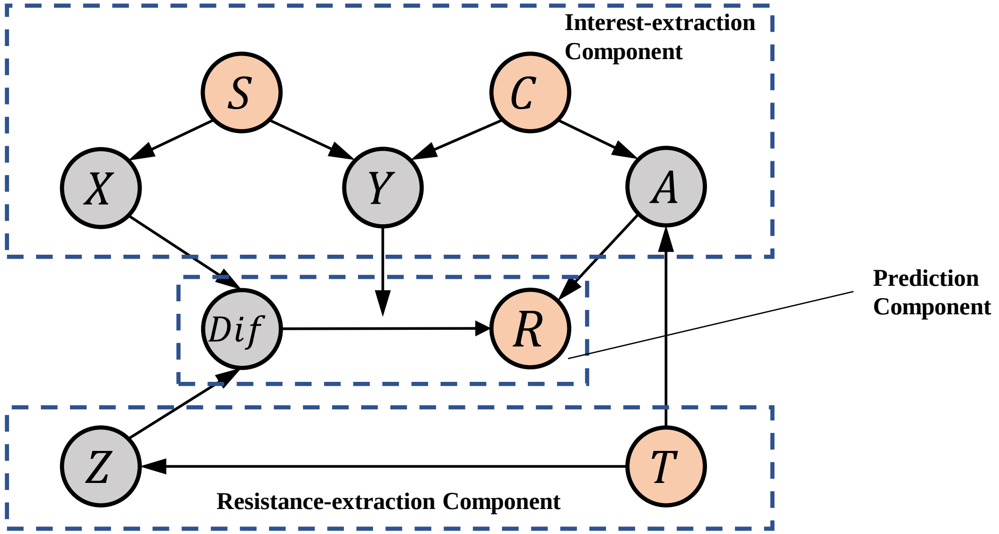
- A novel framework for information diffusion



ReTrend - Observable Data

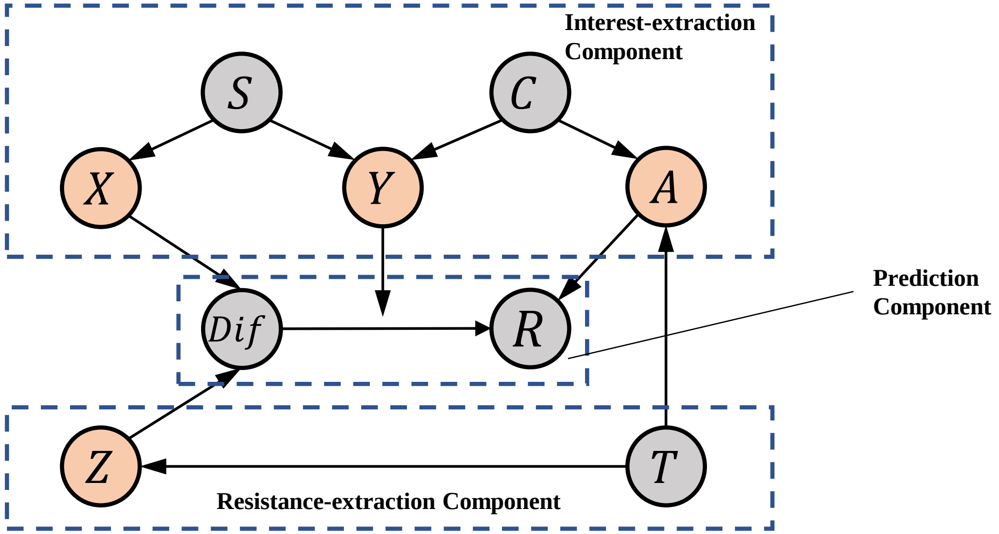
- Four matrices carrying observable data

  - Subscription Matrix (S)
  - Contagion Matrix (C)
  - Resistance Matrix (T)
  - Retweet Matrix (R)



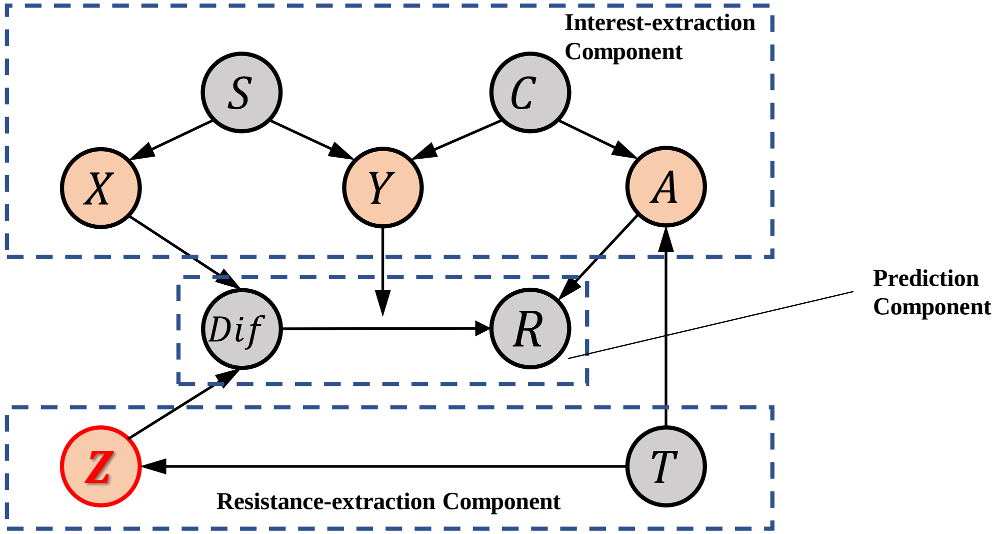
ReTrend - Learning Latent Feature

- Four factor matrices carrying latent feature vectors
  - User Interest Matrix (X)
  - User Influence Matrix (Y)
  - User Resistance Matrix (Z)
  - Item Attraction Matrix (A)



ReTrend - Learning Latent Feature

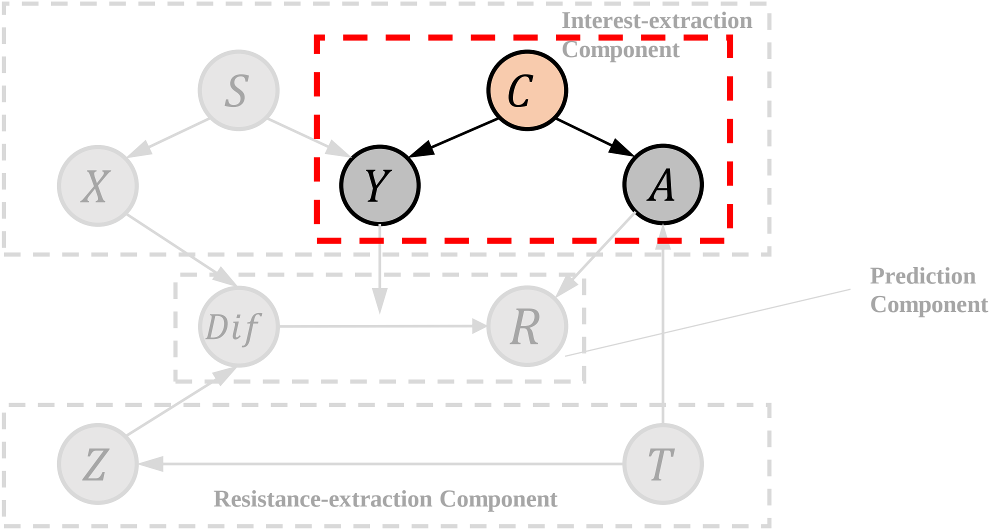
- Four factor matrices carrying latent feature vectors
  - User Interest Matrix (X)
  - User Influence Matrix (Y)
  - **User Resistance Matrix (Z)**
  - Item Attraction Matrix (A)



- We deem this inherent attribute 'resistance' varies over latent space but remains fixed for a fixed user

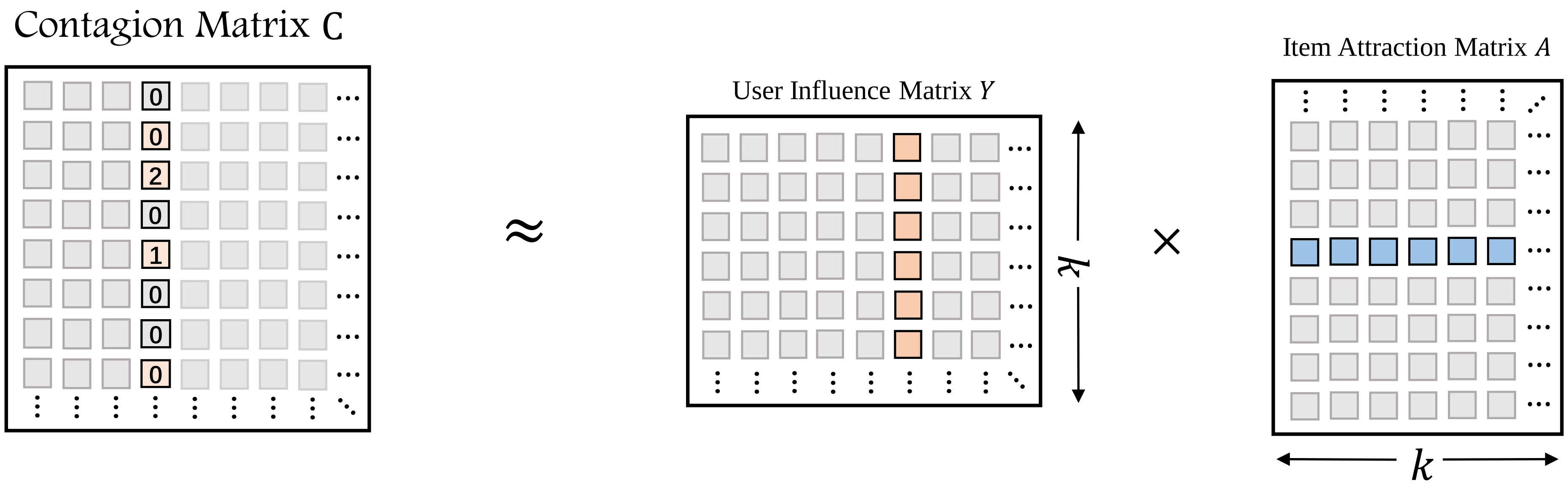
ReTrend - Logic Explanation

- Take Contagion Matrix for example
- Contagion Matrix: |user| × |post|
- Entry C_{ui}: count of retweet behaviors triggered by user u w.r.t. post i
- C_{ui} reflects two facts:
  - to what degree a user can trigger his friends to retweet the post
  - how attractive the post is



ReTrend - Logic Explanation

- Take Contagion Matrix for example

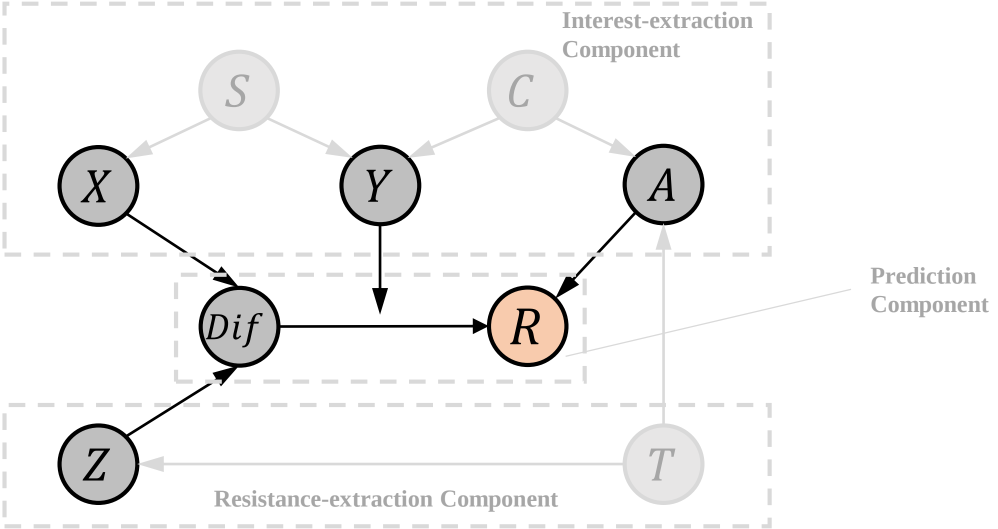


- Assume a Gaussian observation noise

$$C_{u,i} \sim \mathcal{N}(Y_u^T A_i, \sigma_C^2)$$

ReTrend - Logic Explanation

- For Retweet Matrix
- Retweet behavior can be determined by user interest, resistance, parent influence and post attraction



$$R_{u,i} \sim \mathcal{N}(g(X_u, Z_u, Y_{par(u)})^T A_i, \sigma_R^2)$$

where

$$g_{uk}(X_u, Z_u, Y_{par(u)}) = \begin{cases} \max\{X_{uk} - Z_{uk}, \max_{v \in par(u)}\{Y_{vk}\}\}, & par(u) \neq \varnothing \\ 0, & par(u) = \varnothing \end{cases}$$

ReTrend - Entire Model

- Conditional distribution over all observed data as

$$p(S,C,T,R|X,Y,Z,A,\boldsymbol{\sigma}) = \prod_{u=1}^{N}\prod_{v=1}^{N}\left[\mathcal{N}(S_{u,v}|X_u^T Y_v, \sigma_S^2)\right]^{I_{u,v}^S}$$

$$* \prod_{u=1}^{N}\prod_{i=1}^{M}\left[\mathcal{N}(R_{u,i}|g(X_u, Z_u, Y_{par(u)})^T A_i, \sigma_R^2)\right]^{I_{u,i}^R}\left[\mathcal{N}(C_{u,i}|Y_u^T A_i, \sigma_C^2)\right]^{I_{u,i}^C}\left[\mathcal{N}(T_{u,i}|Z_u^T A_i, \sigma_T^2)\right]^{I_{u,i}^T}$$

- Place zero-mean spherical Gaussian priors on latent feature vectors

$$p(X|\sigma_X^2) = \prod_{u=1}^{N}\mathcal{N}(X_u|0, \sigma_X^2\boldsymbol{I}) \qquad p(Z|\sigma_Z^2) = \prod_{u=1}^{N}\mathcal{N}(Z_u|0, \sigma_Z^2\boldsymbol{I})$$

$$p(Y|\sigma_Y^2) = \prod_{u=1}^{N}\mathcal{N}(Y_u|0, \sigma_Y^2\boldsymbol{I}) \qquad p(A|\sigma_A^2) = \prod_{i=1}^{M}\mathcal{N}(A_i|0, \sigma_A^2\boldsymbol{I})$$

ReTrend - Entire Model

- By modifying the log-likelihood, we obtain the loss function as

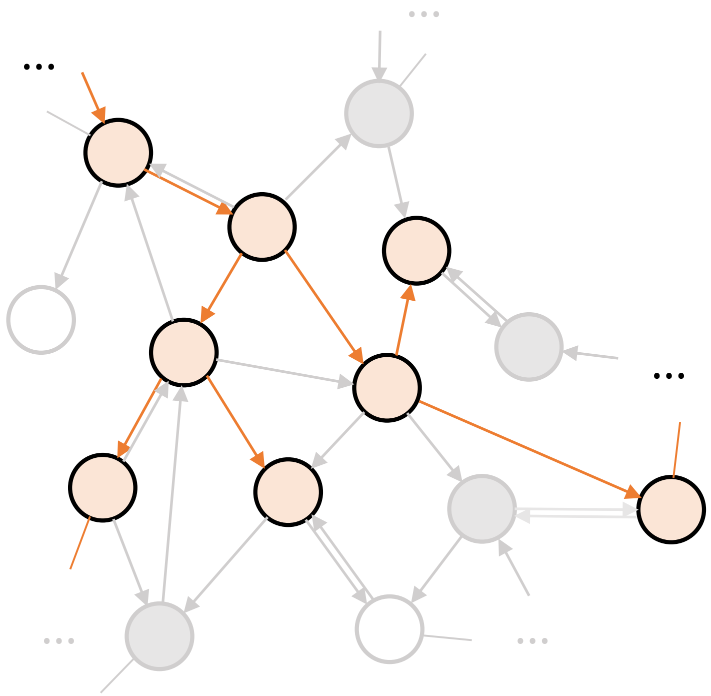
$$\min_{X,Y,Z,A} J(S,C,T,R,X,Y,Z,A) = \|R - g(X,Z,Y)^T A\|^2 + \lambda_S \|S - X^T Y\|^2 + \lambda_C \|C - Y^T A\|^2 + \lambda_T \|T - Z^T A\|^2 + \lambda_X \|X\|^2 + \lambda_Y \|Y\|^2 + \lambda_Z \|Z\|^2 + \lambda_A \|A\|^2$$

where $\|\cdot\|^2$ represents the Frobenius norm, and $\lambda_{Mtx} = \frac{\sigma_R^2}{\sigma_{Mtx}^2}$, $Mtx = S, C, T \cdots$.

- SGD for optimization

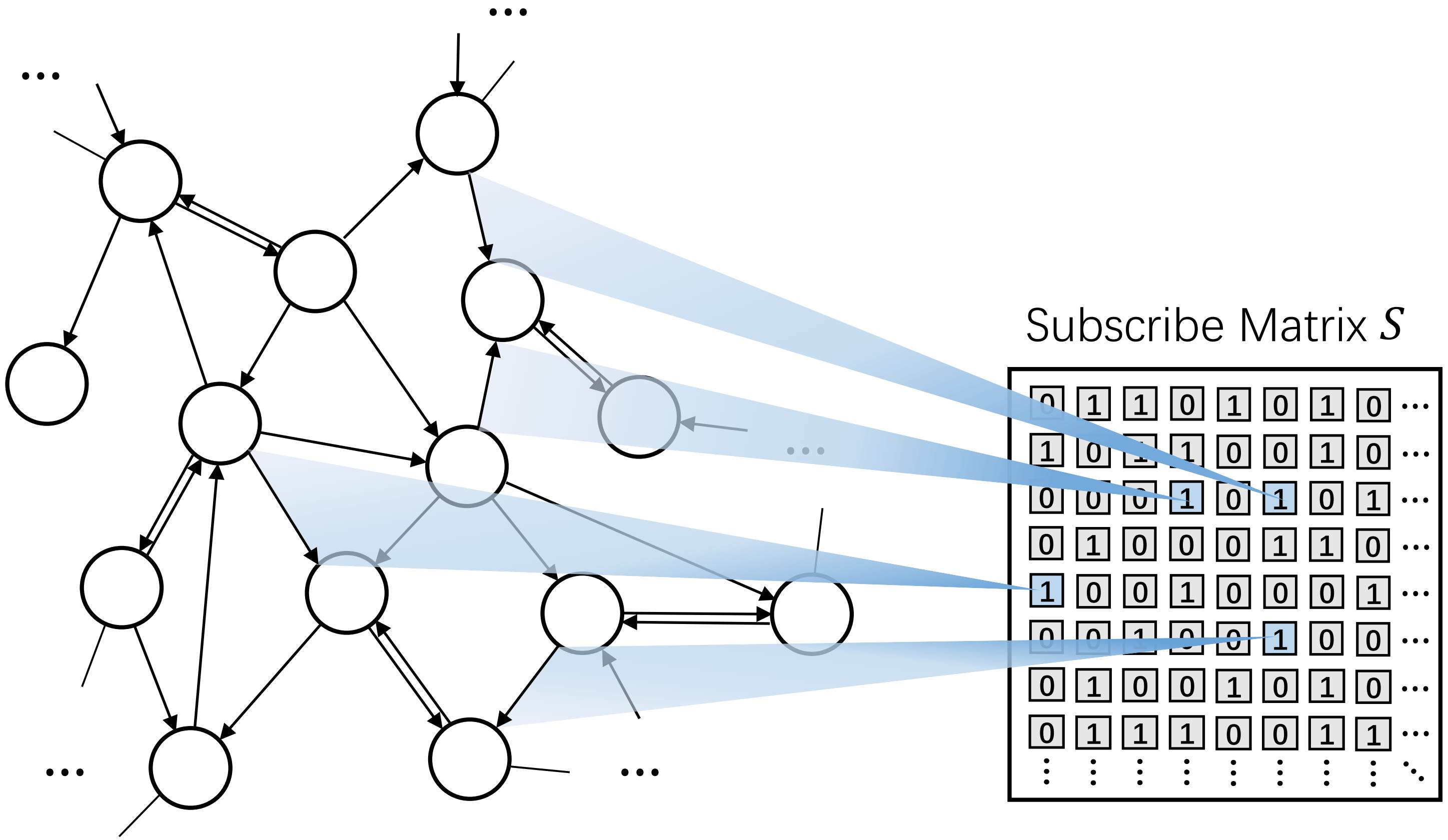
ReTrend - Retweet-tree Encoding

- How ReTrend leverage information better?
- Tree-structured essence of information cascade – Retweet-tree



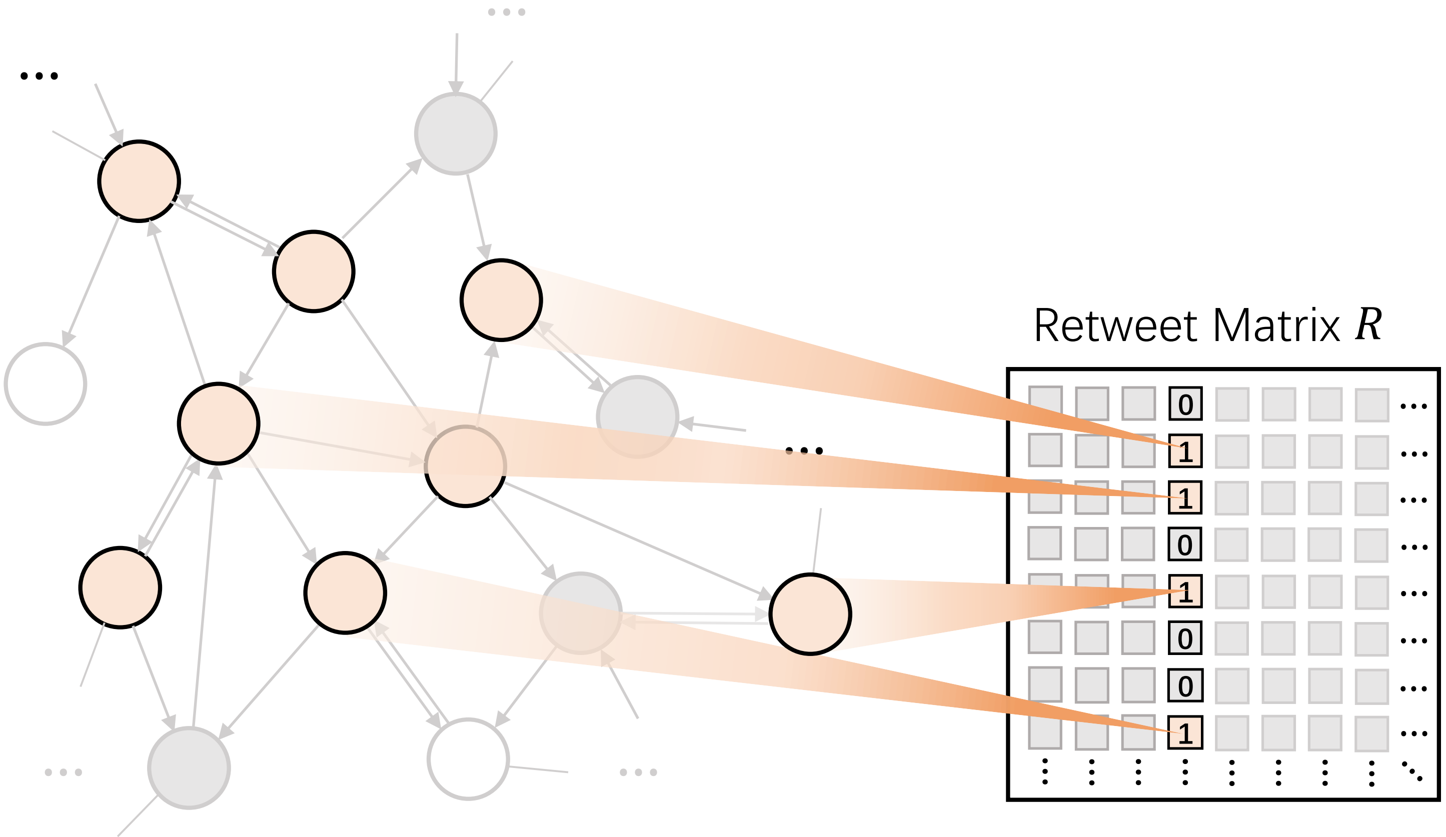
ReTrend - Retweet-tree Encoding

- Subscription Matrix



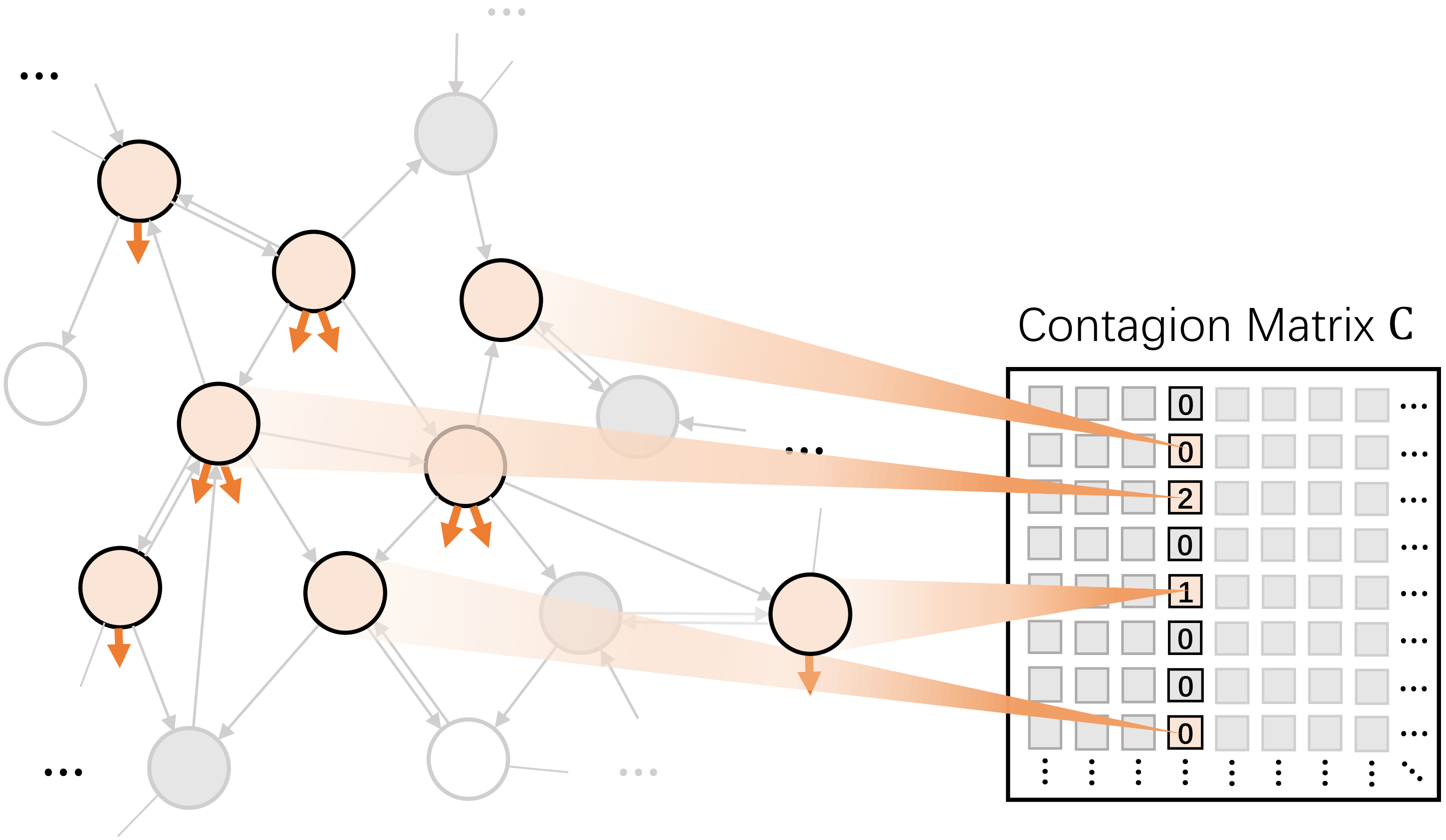
ReTrend - Retweet-tree Encoding

- Retweet Matrix



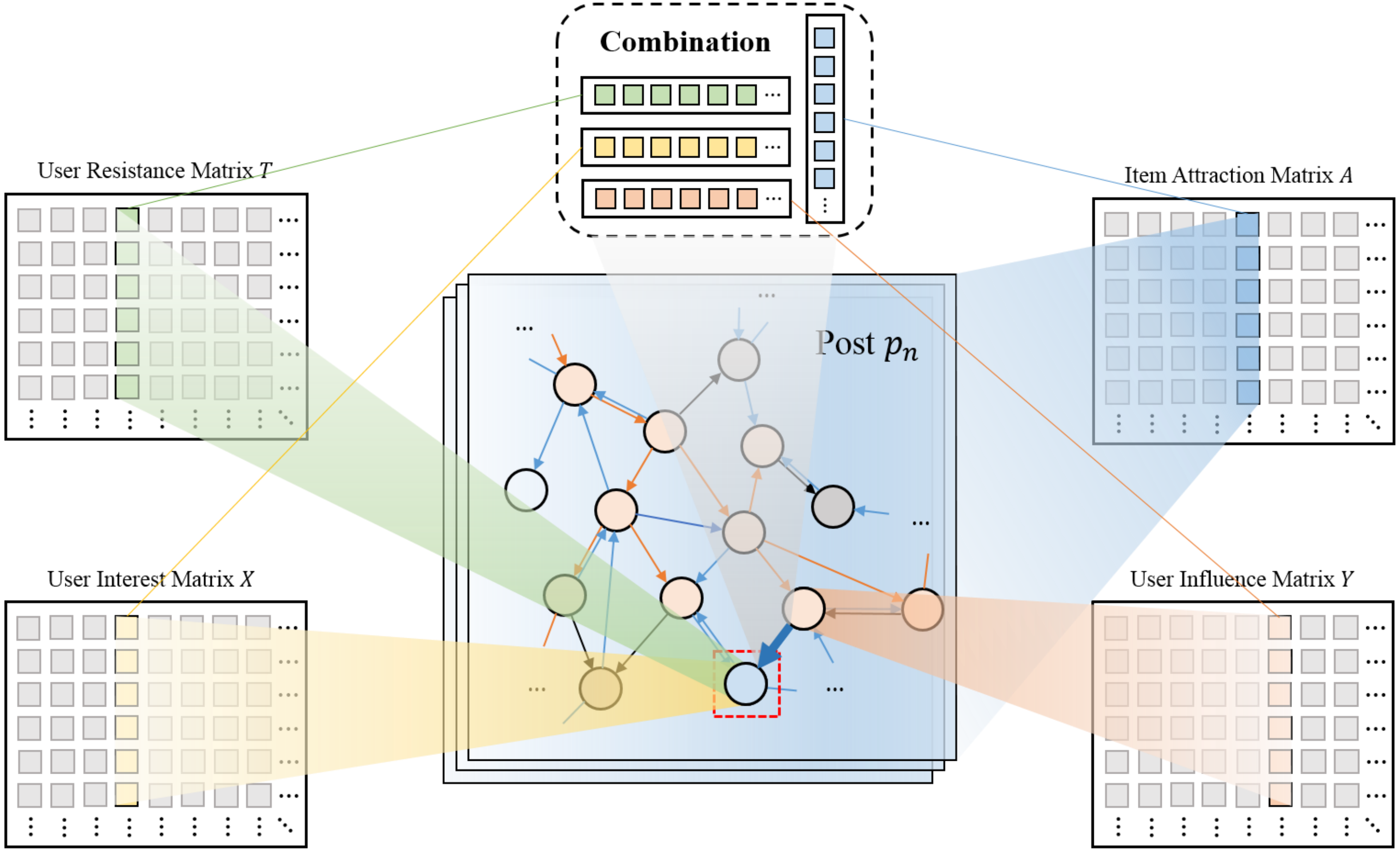
ReTrend - Retweet-tree Encoding

- Contagion Matrix



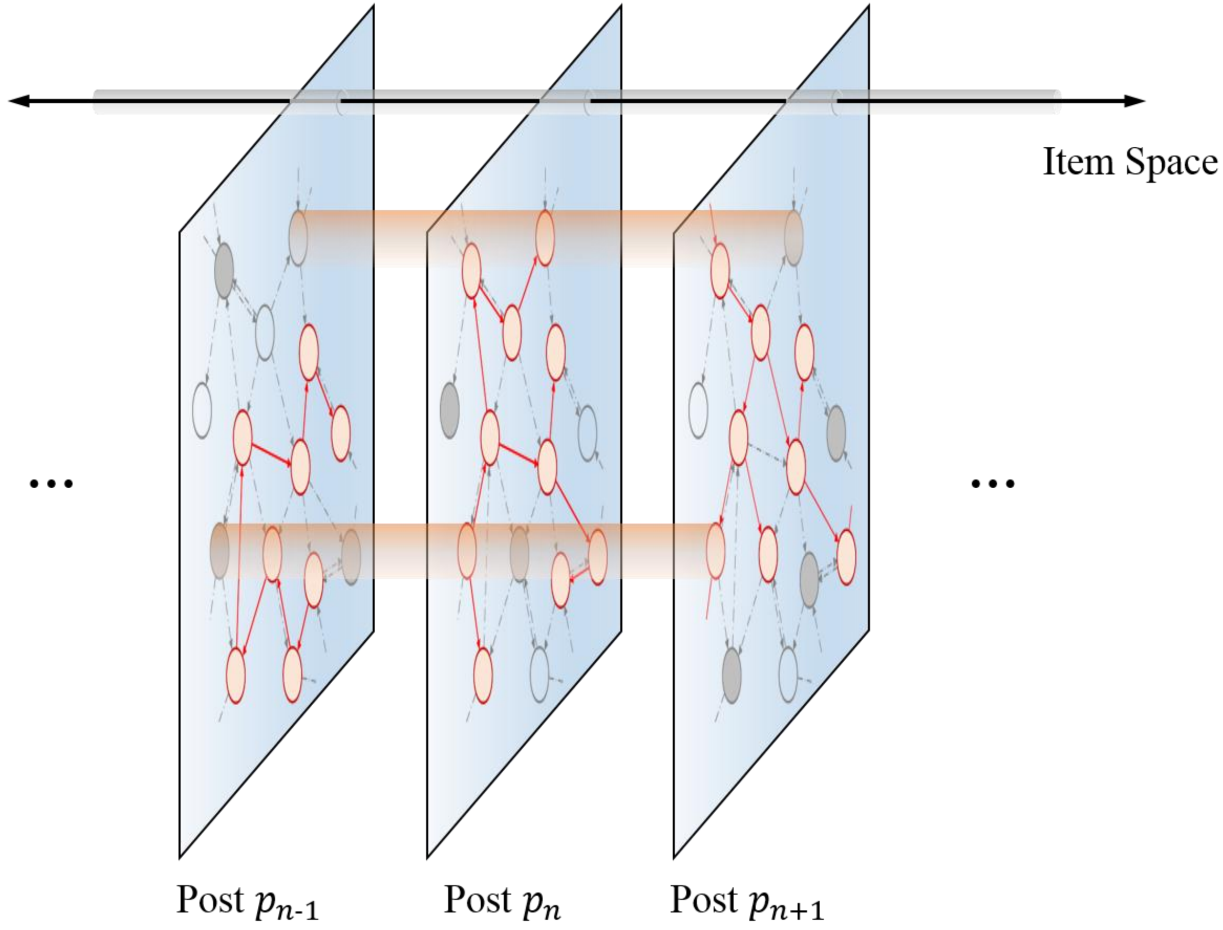
ReTrend - Training

- Dynamic inference on the most likely retweet-tree structure



ReTrend - Training

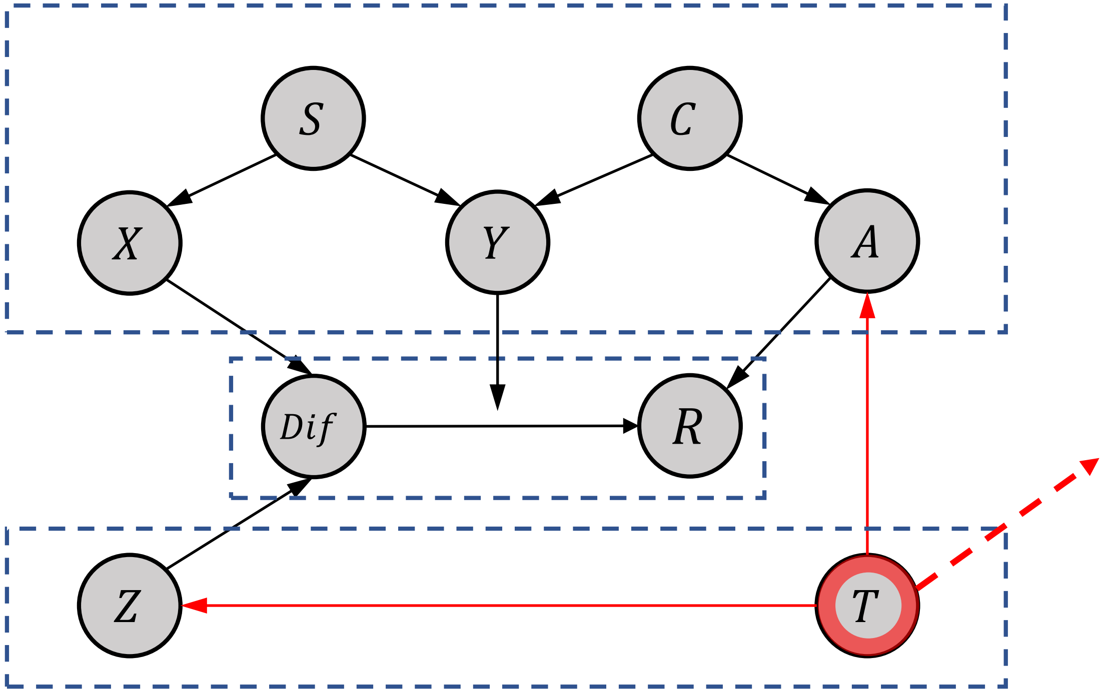
- **AND**, it is post-transcending



Modification

Matrix Factorization - Drawbacks

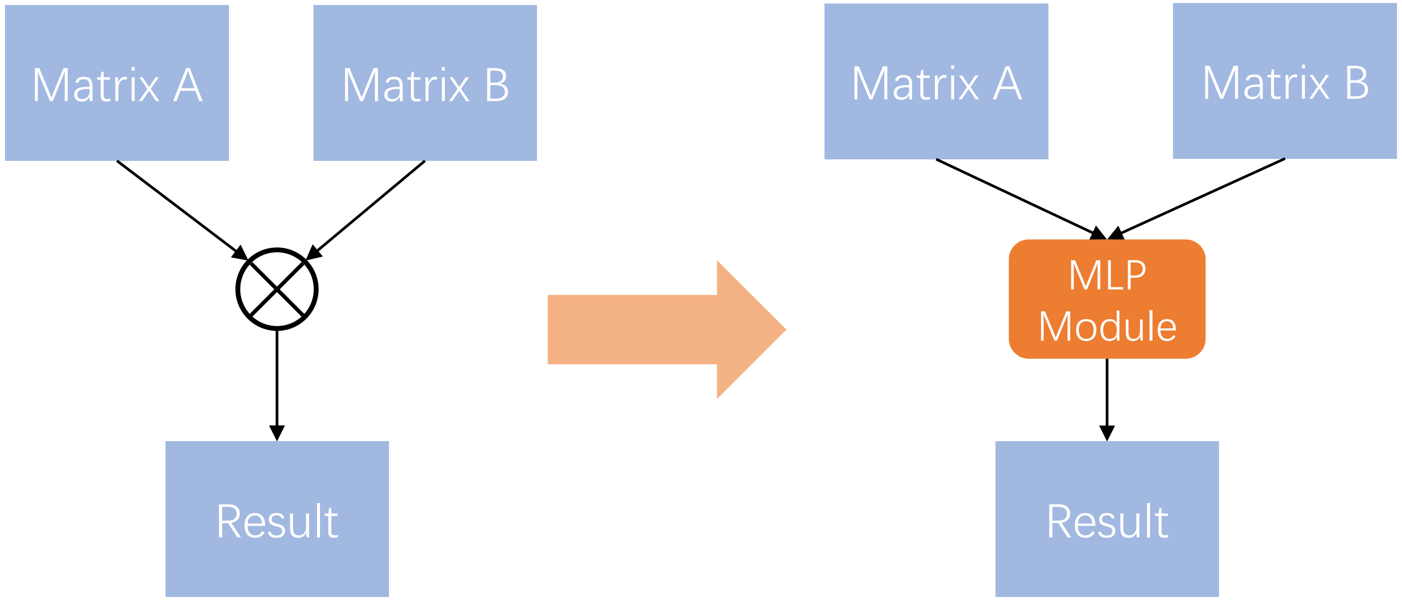
- Simple and fixed inner-product: Low Non-linearity[4]
- Complex inference in low-dimensional latent space
- Too much constraints



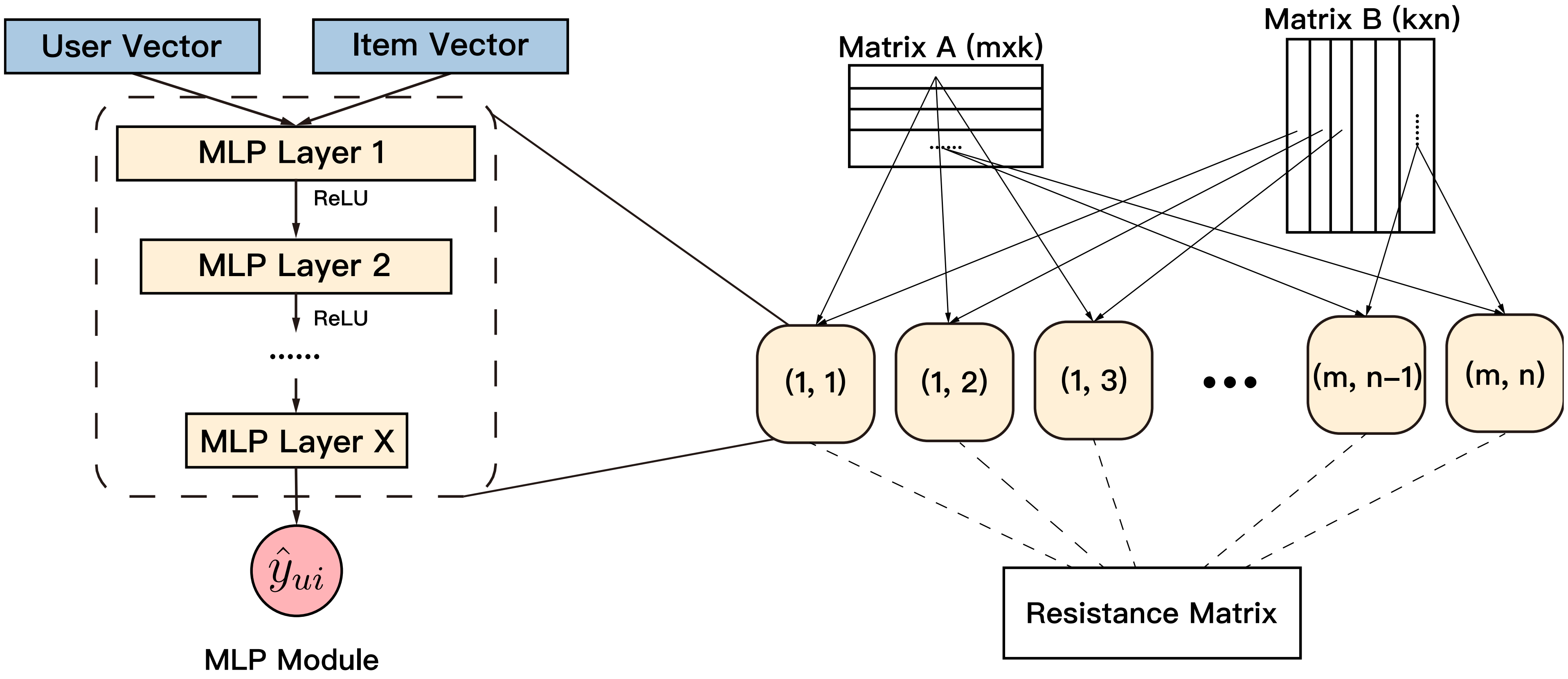
Pure linear operation:
Empirically **low performance**

MLP Module - Optimization for MF

- Replace multiplication with a simple MLP module.
- Level up non-linearity



MLP Module - Detail



Experiments - Dataset

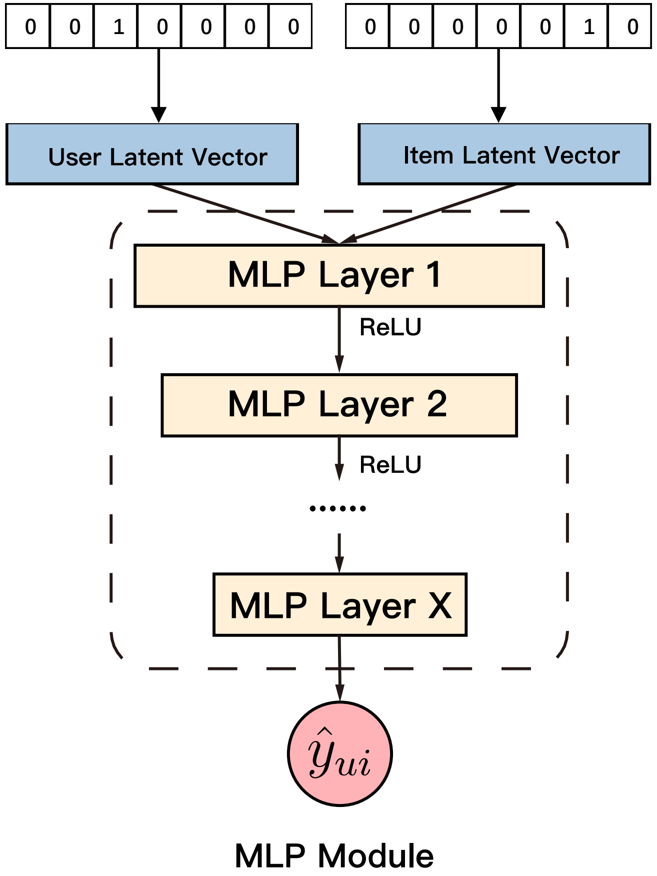
- **Real-world dataset from Twitter**
- More than 90,000 users and 99,696,204 tweets related[1][2].
- 440,000+ subscribes.
- 2,370,000+ retweet behaviors.
- 18,210,000+ un-retweet behaviors.
- 18,210,000+ resistance tuples.
- 2,170,000+ contagion tuples.

[1] https://www.aminer.cn/data-sna#Twitter-Dynamic-Net
[2] https://www.aminer.cn/data-sna#Twitter-Dynamic-Action

Experiments - Implementation detail

- For ReTrend:
  - Indicator matrices
  - Normalization for *R, S, C, T*
  - Latent Feature: 30
  - SGD:
    - Batch size: 1000
    - Training epoch: 100
    - Learning rate: 0.03; 0.03*value(loss function)/500 when loss function is below 500
- For MLP Module:
  - Trained for 20 epochs implemented by Keras

Experiments - Performance



- Plausibility Validation for MLP:
  - Plain features: only the identity of user and item
  - Embedding: latent vector for user and item

Layer Settings	Accuracy-3	Accuracy-2
{128,64,32,16,8}	0.882	0.997
{64,32,16,8}	0.783	0.996

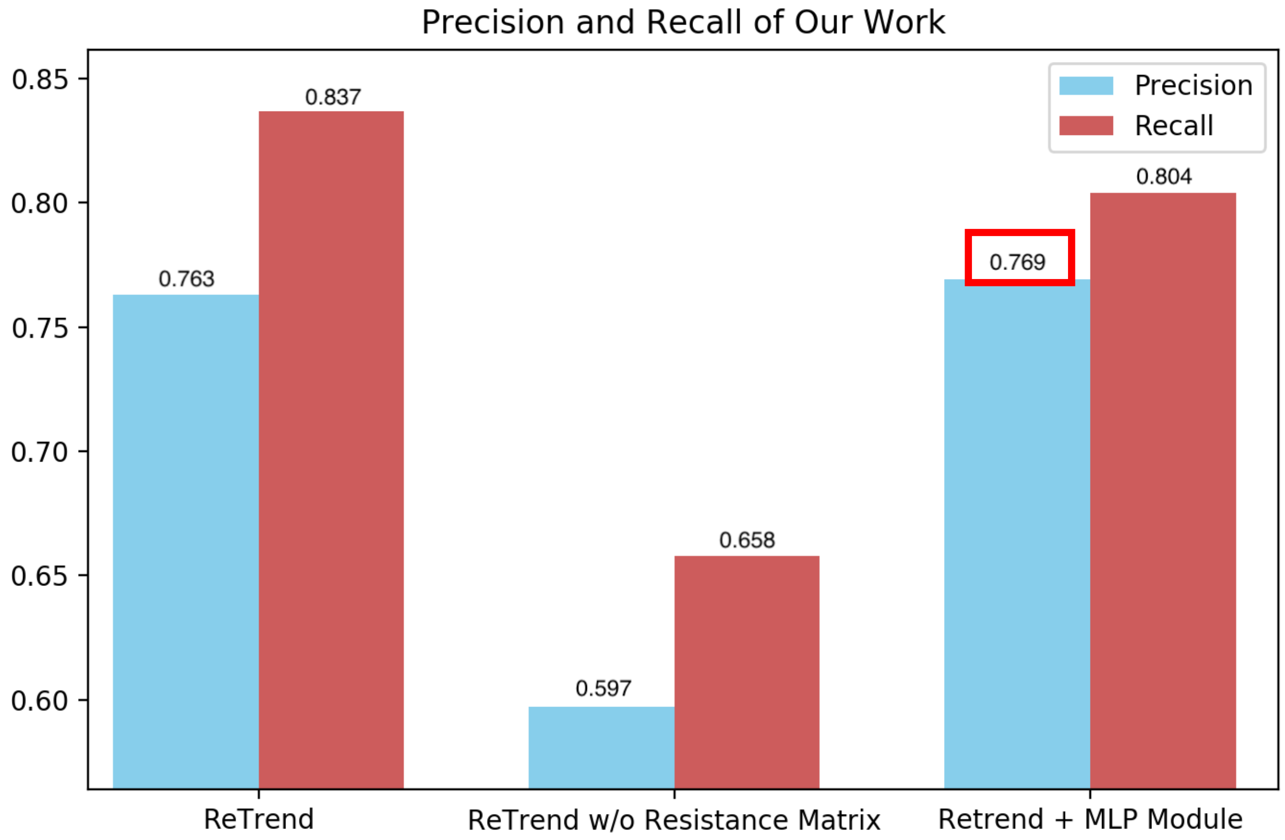
Experiments - Performance

- Baselines:
  - Random
  - Word Vector Based SVM[5]
  - Neural Collaborative Filter[4]

Method	Precision	Recall
Random	0.264	0.485
SVM	0.382	0.649
NCF	$\boxed{0.782}$	0.821
ReTrend	0.763	$\boxed{0.837}$

Experiments - Performance

- ReTrend+MLP:



Reference

- [1] Jiang, Bo, Jiguang Liang, Ying Sha, and Lihong Wang. "Message clustering based matrix factorization model for retweeting behavior prediction." In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM), pp. 1843-1846. ACM, 2015.
- [2] Cui, Peng, Fei Wang, Shaowei Liu, Mingdong Ou, Shiqiang Yang, and Lifeng Sun. "Who should share what?: item-level social influence prediction for users and posts ranking." In Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, pp. 185-194. ACM, 2011.
- [3] Li, Jun, Jiamin Qin, Tao Wang, Yi Cai, and Huaqing Min. "A Collaborative Filtering Model for Personalized Retweeting Prediction." In International Conference on Database Systems for Advanced Applications, pp. 122-134. Springer, Cham, 2015.
- [4] Xiangnan He, Lizi Liao, Hanwang Zhang. "Neural Collaborative Filtering". arXiv preprint arXiv: 1708.05031, 2017
- [5] Zhang, Q., Gong, Y., Wu, J., Huang, H., & Huang, X. (2016). Retweet prediction with attention-based deep neural network. 75-84.

Thanks