# Fast Information Cascade Prediction Through Spatiotemporal Decompositions

**Huanyang Zheng** and Jie Wu
Department of Computer and Information Science
Temple University, USA

# Introduction

Online social network:

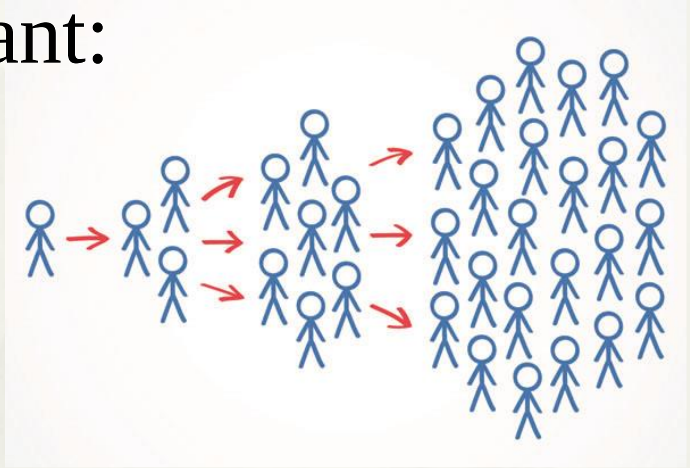
- Fundamental medium for information spreading
- Share startling news, creative ideas, and interesting stories


Information cascade:

- If Alice shares a photo, Bob may scan this photo and then further share it with his/her followers later
- Iterative information propagations

# Introduction

Cascade predictions are important:



- Control of online rumors
- Forecast of marketing strategies

Challenges:

- When will a user further propagate the information?
- How should we process the social topological and time information?

# Dataset observations

Flickr dataset:

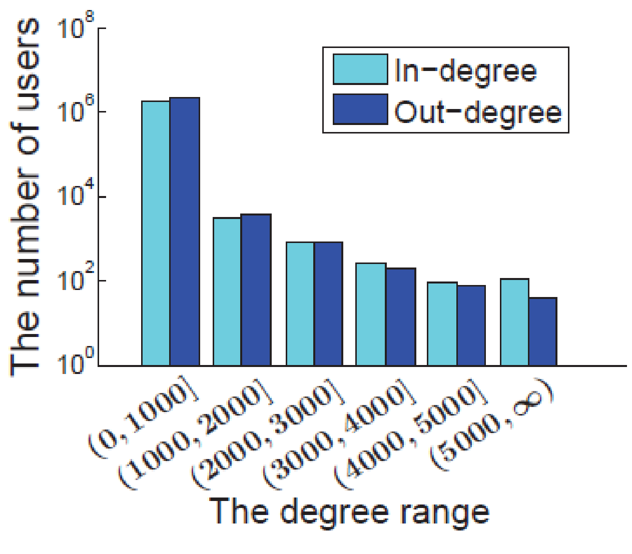
- An online social network site for sharing photos among users
- Photos can be labeled by “favorite-mark” (cascade)

| | |
|---|---|
| Time period (two periods) | 11/02/2006 to 12/03/2006<br>02/03/2006 to 05/18/2007 |
| # Links | 17,034,807 to 33,140,018 |
| # Users | 1,487,058 to 2,302,925 |
| # Photos | 11,267,320 |
| # Favorite marks | 34,734,221 |
| # Popular photos | 14,002 |
| Most popular photo | Marked by 2,998 times |
| Largest in / out-degree | 21,001 / 26,367 |

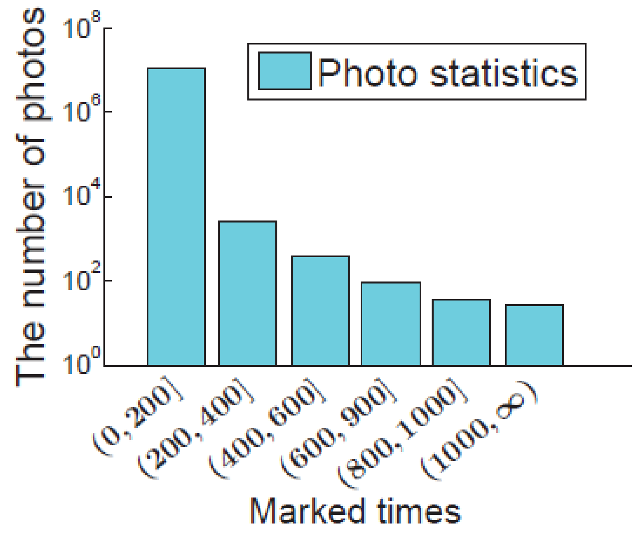
# Dataset observations

## Dataset observations:

- A large amount of data!
- Social topological information
- Time information (cascade time)




(a) User degree distribution.




(b) Favorite mark distribution.

# Ideas

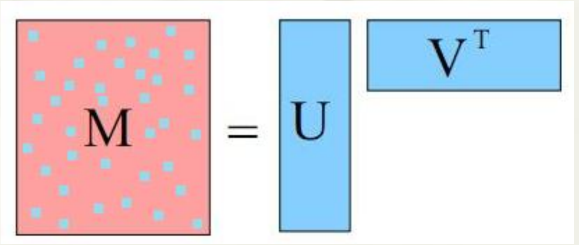
Objective – predict the number of propagated users at a future time slot

Idea – decompose the spatiotemporal cascade information to user characteristics

- Conduct predictions based on user characteristics
- Reduce the time complexity of the algorithm

Detail – convert matrix to vectors

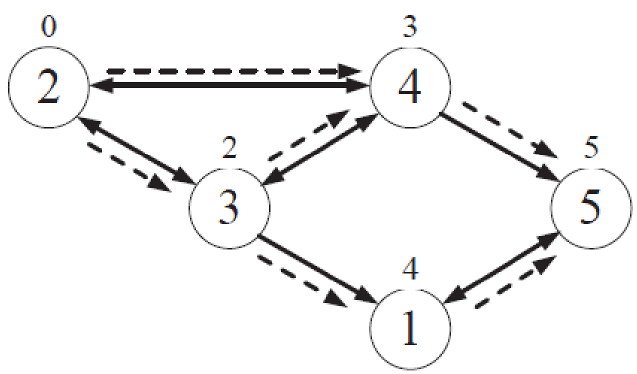
- Cascade information – a matrix
- User characteristics – two vectors



# Ideas: Spatiotemporal Information

## Spatiotemporal cascade information

- A time matrix also includes the space information:
  - ❑ Nodes that are closer within the social topology are more likely to be propagated at closer times.
- Let $t_{ij}$ be the time when user ***j*** starts to propagate information after having been influenced by user ***i***.




(a) A spatiotemporal cascade.

$$T = \begin{bmatrix} \infty & \infty & \infty & \infty & 5 \\ \infty & \infty & 2 & 3 & \infty \\ 4 & \infty & \infty & 3 & \infty \\ \infty & \infty & \infty & \infty & 5 \\ \infty & \infty & \infty & \infty & \infty \end{bmatrix}$$

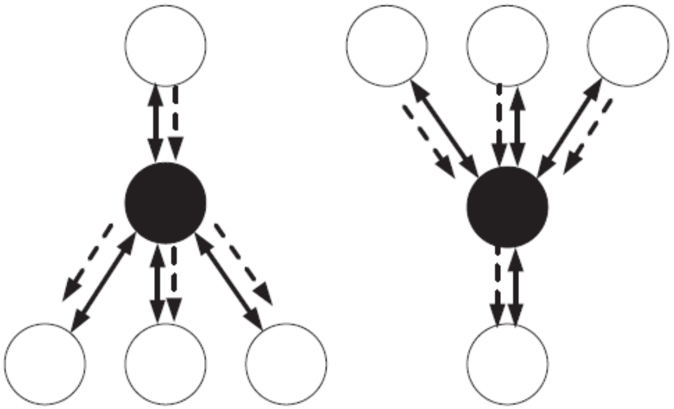
(b) The time matrix for (a).

# Ideas: User Characteristics

User characteristics (two vectors)

- Persuasiveness (information sender)
  - ❑ Followees' abilities to propagate information
- Receptiveness (information receiver)
  - ❑ Followers' willingness to accept information.

High persuasiveness and receptiveness



Low persuasiveness and receptiveness

# Decomposition

Step 1: map the time matrix to a weighted matrix

- Mapping objective
  - ❑ Tune the weights of space and time information
  - ❑ Earlier cascades are more important (larger value)
- Use exponential functions (memoryless function)

$$T(\tau_1 = 4) = \begin{bmatrix} \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & 2 & 3 & \infty \\ 4 & \infty & \infty & 3 & \infty \\ \infty & \infty & \infty & \infty & \infty \\ \infty & \infty & \infty & \infty & \infty \end{bmatrix}$$

(a) The time matrix at $\tau_1 = 4$.

$$M(\tau_1 = 4) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.67 & 0.55 & 0 \\ 0.45 & 0 & 0 & 0.55 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(b) The corresponding mapping result.

# Decomposition

Step 2: singular value decomposition (SVD)

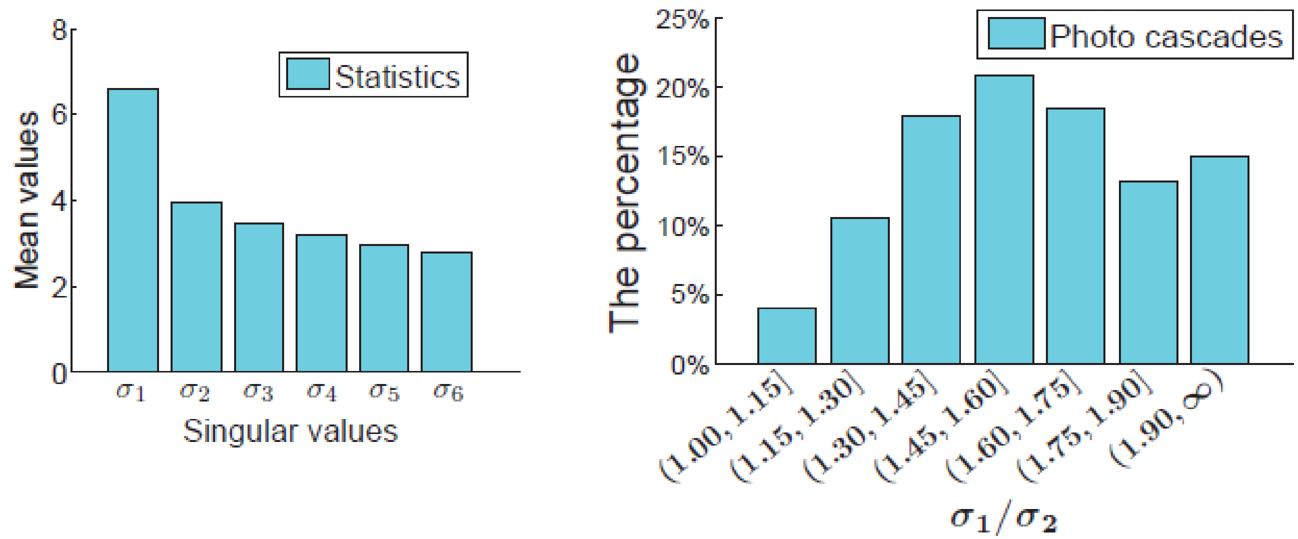
- Approximately reconstruct the weighted matrix (the tuned time matrix) by two vectors
- Two vectors represent persuasiveness and receptiveness, respectively
- Larger value in the matrix (earlier cascades)
  - ❑ Result in larger persuasiveness and receptiveness

$$M_1 = \sigma_1 u_1 v_1^* = 0.98 \cdot \begin{bmatrix} 0.00 \\ 0.83 \\ 0.56 \\ 0.00 \\ 0.00 \end{bmatrix} \cdot \begin{bmatrix} 0.26 \\ 0.00 \\ 0.57 \\ 0.78 \\ 0.00 \end{bmatrix}^* = \begin{bmatrix} 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.21 & 0.00 & 0.46 & 0.63 & 0.00 \\ 0.14 & 0.00 & 0.31 & 0.42 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \end{bmatrix}$$

# Decomposition

Information loss in the decomposition

- Can be revealed by the largest singular values




(a) Statistics on $\sigma_i$.  (b) $\sigma_1/\sigma_2$ of photo cascades.

- Information loss is limited!

# Cascade prediction

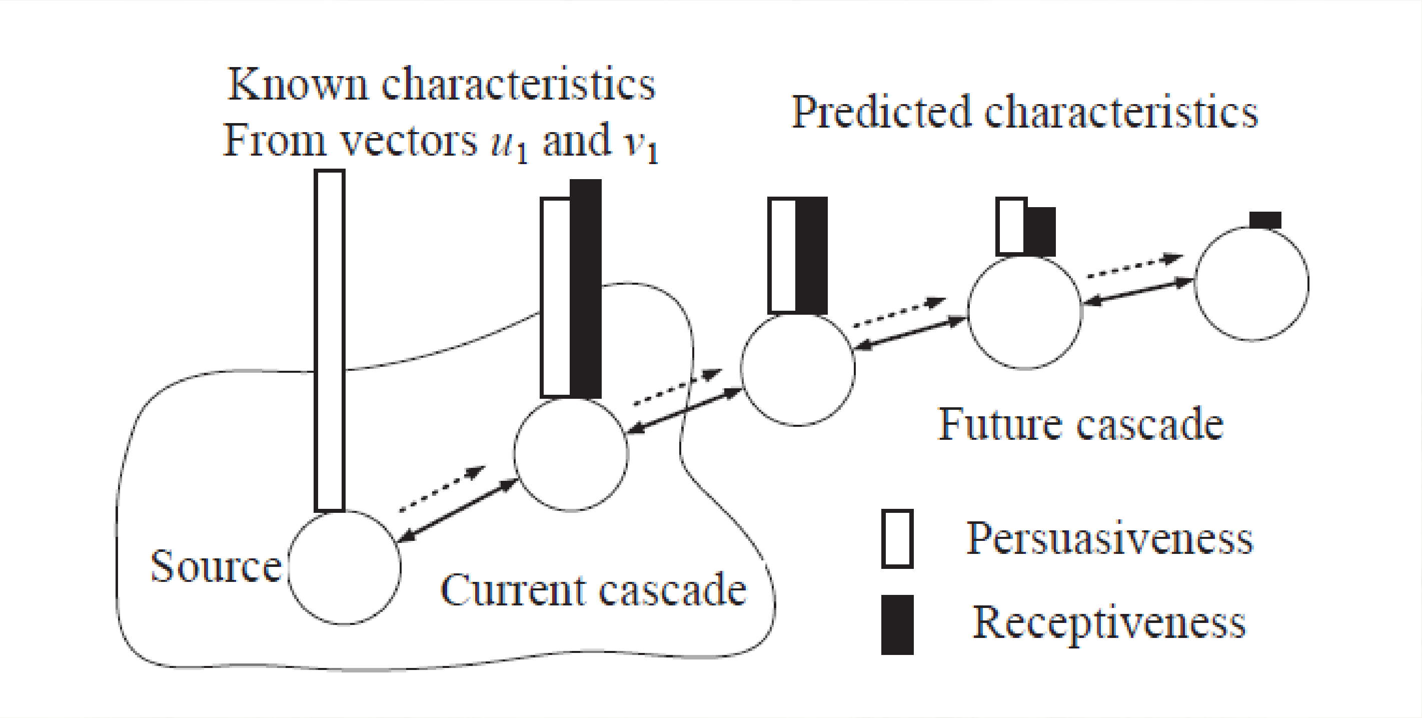
The pattern of persuasiveness

- If a node with a high out-degree is spatially far away from the information source, it may not be propagated, and thus it cannot positively propagate the information further (i.e., low persuasiveness).
- In the case of a temporal remote node, it also has low persuasiveness, since its followers may have been propagated by other nodes.

A similar rule works for the receptiveness.

# Cascade prediction

The pattern of the cascade




Persuasiveness and receptiveness should decay with respect to their spatiotemporal distances to the source

# Cascade prediction

## Non-historical predictions

- Predict persuasiveness and receptiveness hop by hop
- Along the shortest path tree from the source to the other nodes

## Historical predictions

- Use historical data as predictions

## Assemble predicted persuasiveness/receptiveness

- Recover the time matrix as the final prediction

# Evaluations

We focus on cascades of popular photos that are marked "favorite" more than 100 times

- Photos of different levels of popularity stand for cascades of different types

Each photo may be involved in multiple cascades that are independent of each other

- Only the largest cascade is selected

Define $\tau_1$ as the current time, and $\tau_2$ as the future time for the cascade prediction

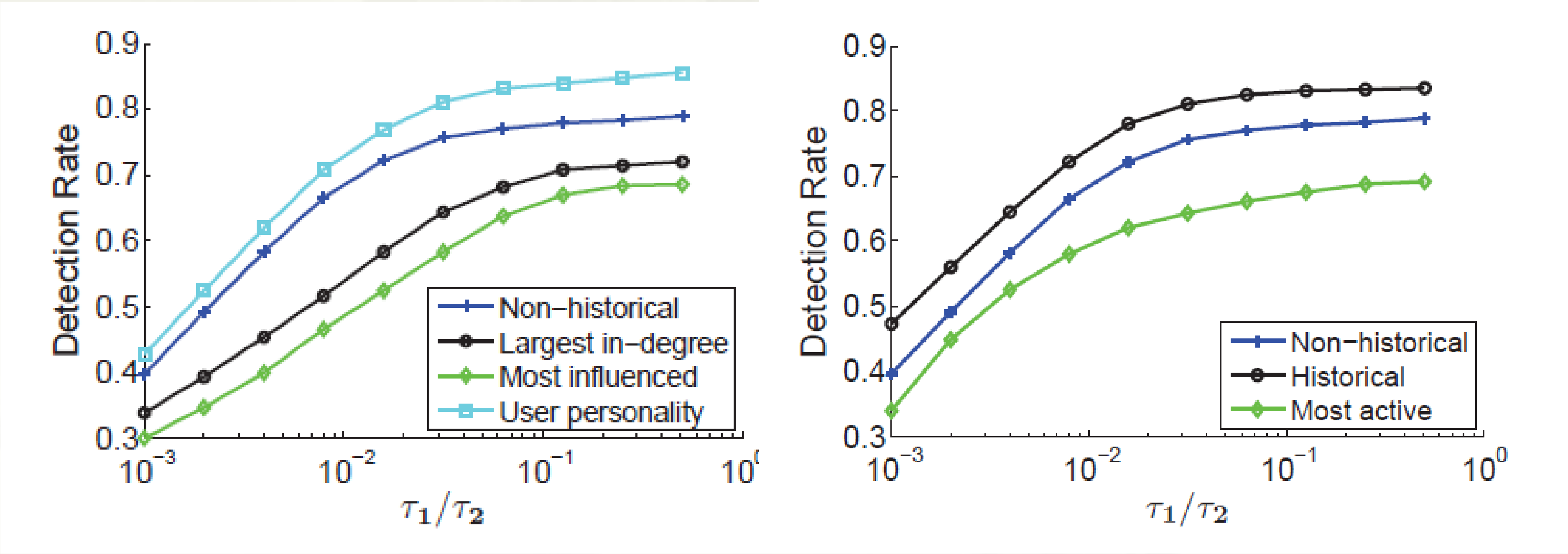
# Evaluations

Baseline algorithms:

- **Largest in-degree**: the largest in-degree node (in social topology) would be the next propagated node
- **Most influenced**: the node that has the largest number of incoming propagated neighbors would be the next propagated node
- **Most active**: the node that is the most active (propagated by former cascades for the most times), would be the next propagated node
- **User personality**: incorporate extra user personality

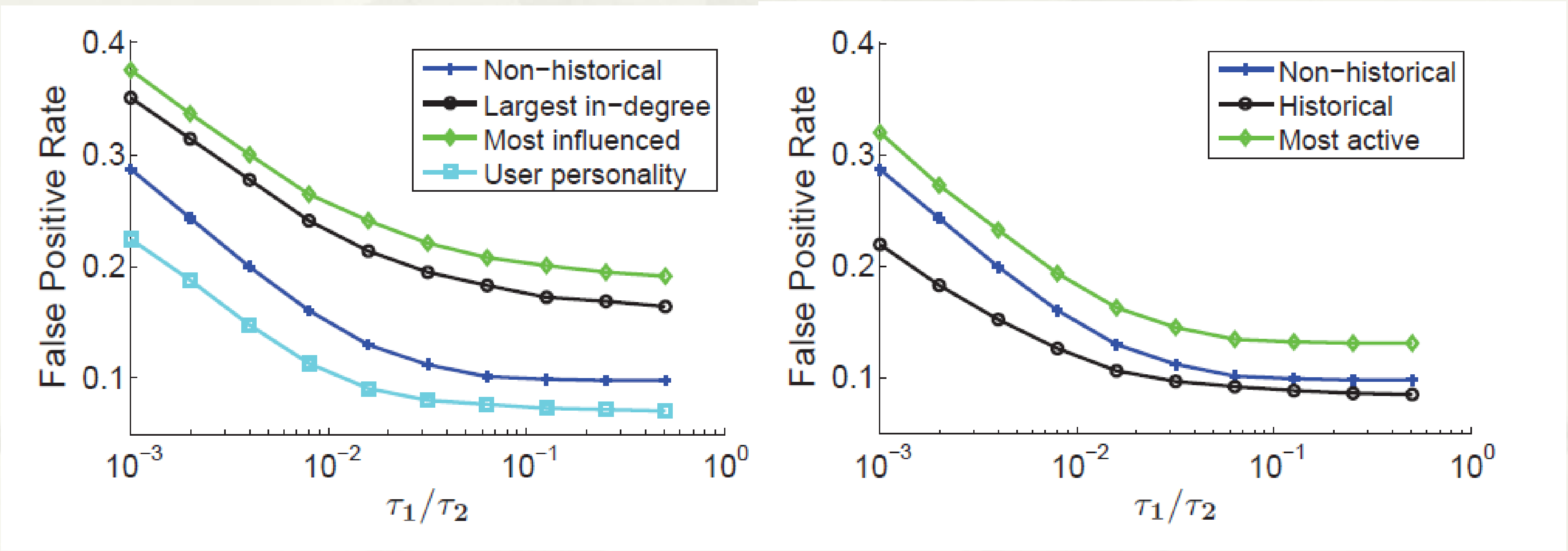
# Evaluations

Non-historical v.s. historical (detection rate):



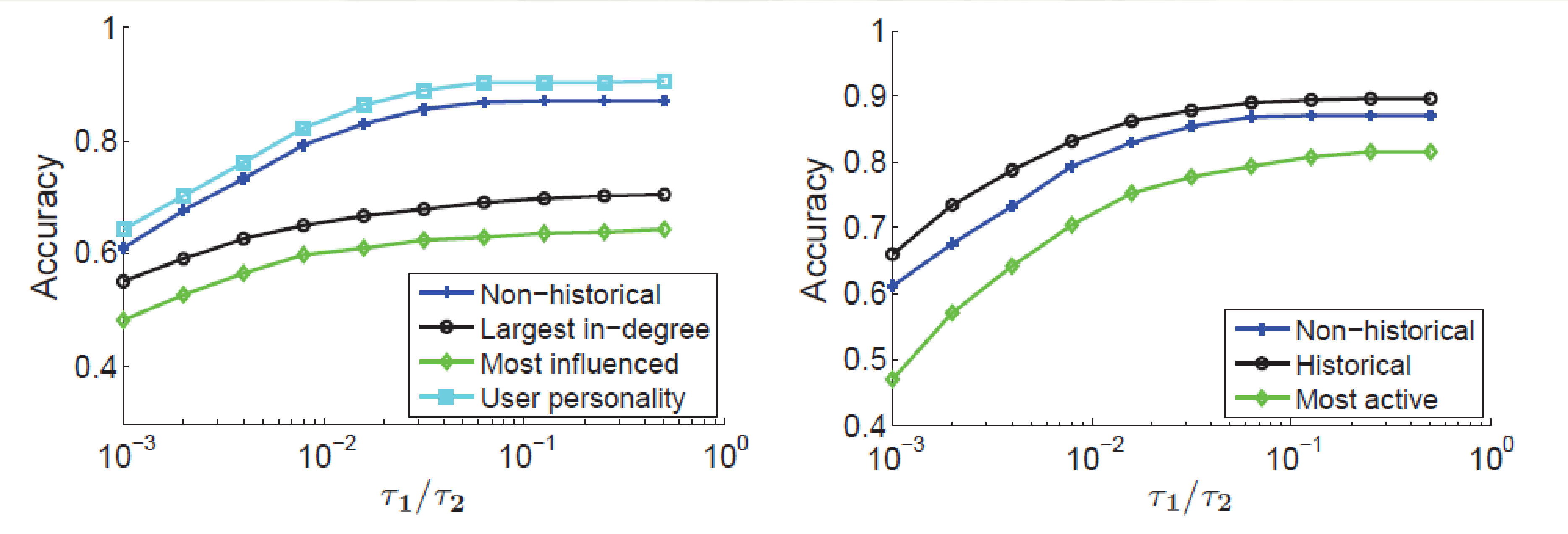
# Evaluations

Non-historical v.s. historical (false positive rate):



# Evaluations

Non-historical v.s. historical (accuracy):



# Evaluations

Evaluation summary:

- For non-historical predictions, our algorithm gets about 20% higher accuracy than the two baselines (for $\tau_1/\tau_2 \geq 0.1$ )
- For historical predictions, our algorithm gets about 15% higher accuracy than the baseline, and 10% higher accuracy than the non-historical algorithm
- The future of the cascade is very “predictable”. A small amount of existing information can provide very accurate future predictions

# Conclusions

Conclusions:

- Decompose the space and time cascade information into user characteristics
- The information loss in the decomposition is limited
- Use the shortest path tree to infer the trace of the information propagation

Future work

- Parallel and distributed computing

# End

Thank you!

Questions?