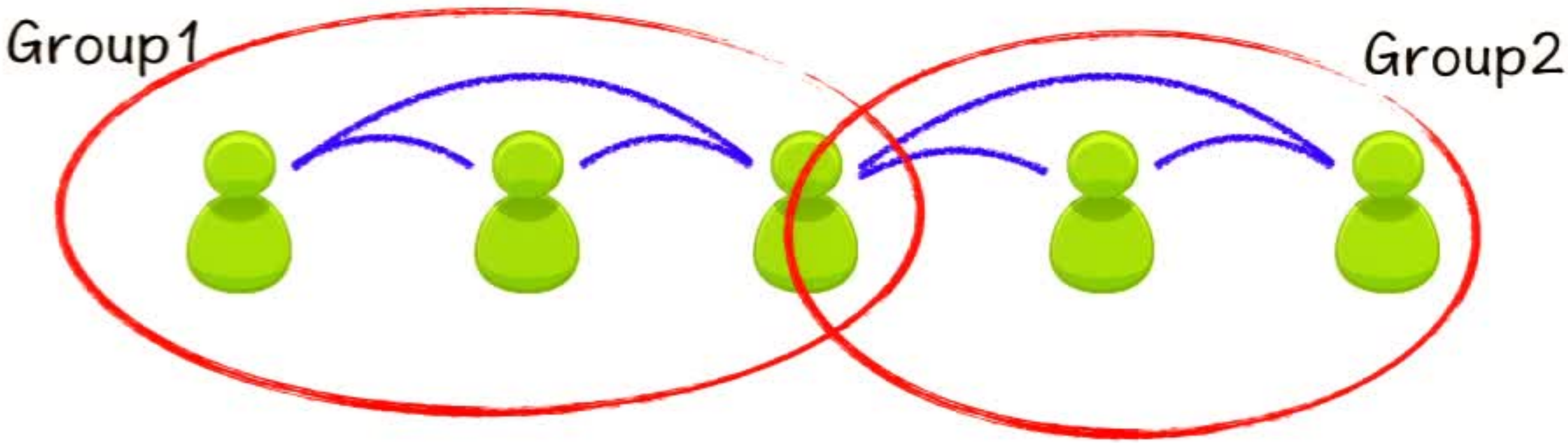
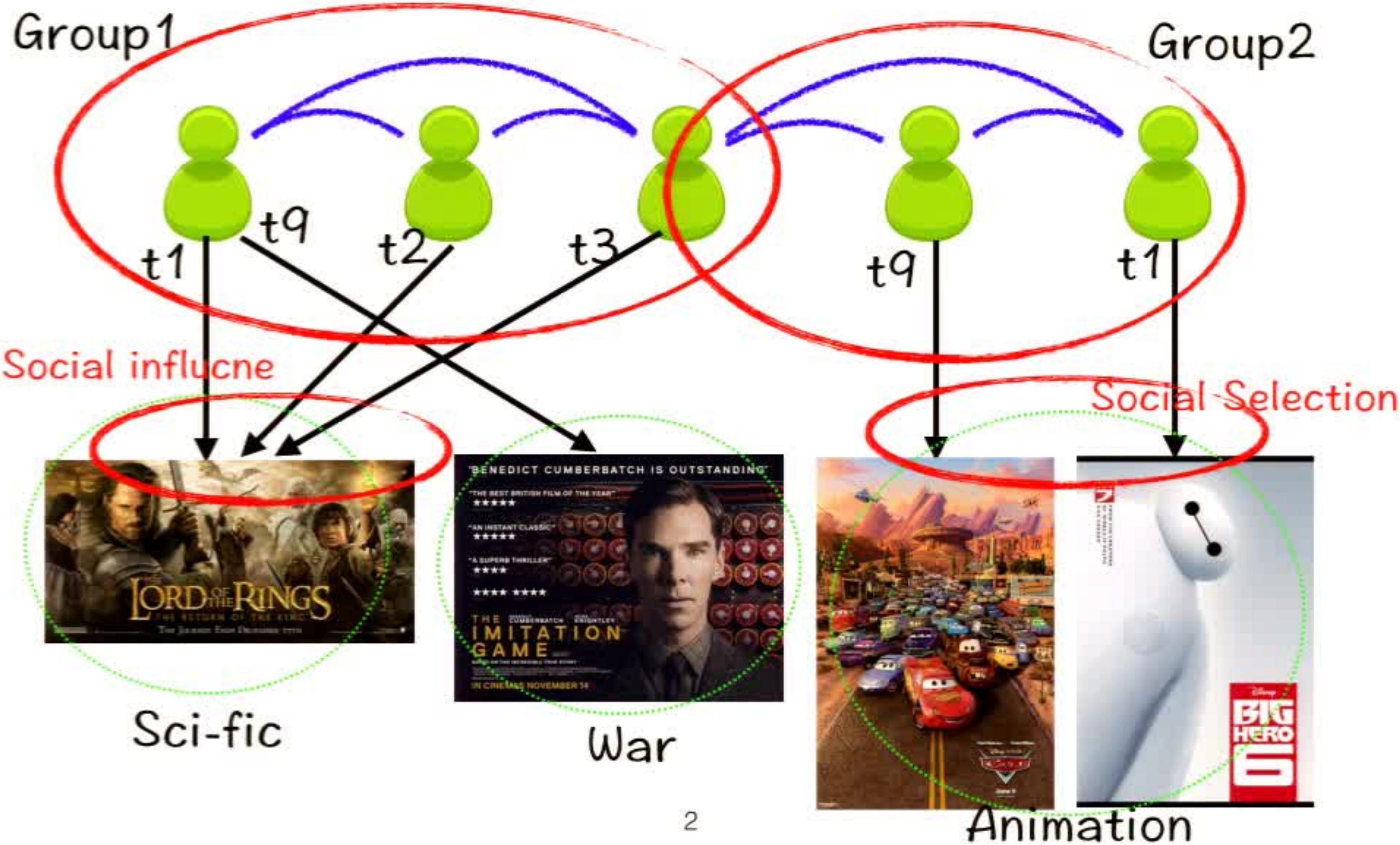
Modeling Users' Adoption Behaviors
with Social Selection and Influence

Ziqi Liu
Department of Computer Science
Xi'an Jiaotong University

Movie Adoption



Movie Adoption

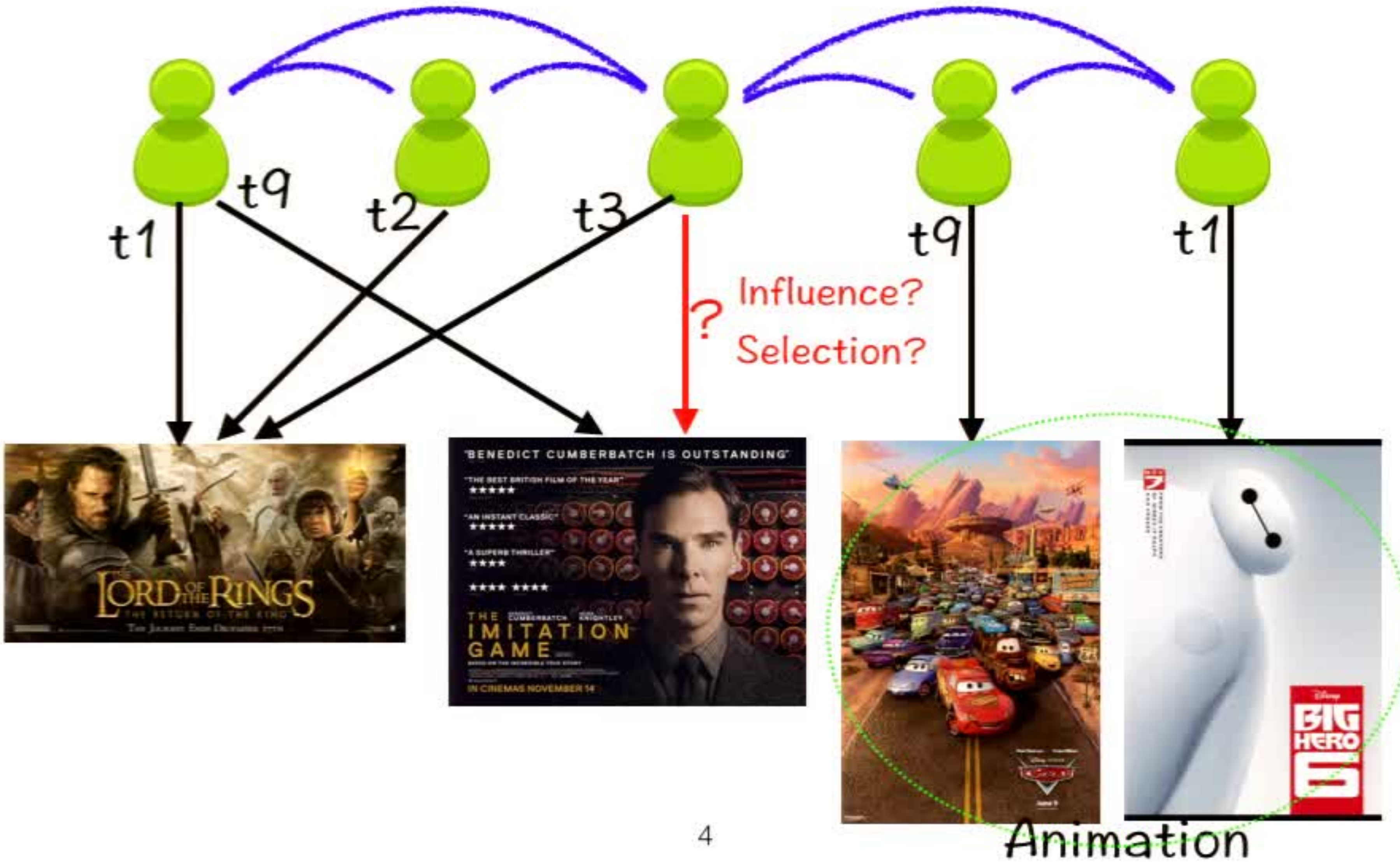


Two Social Factors

- *Social Selection*, which refers to the phenomenon that users tend to form new links to others who are similar to them; Basis of CF

- *Social Influence*, which refers to the phenomenon that users' decisions are usually affected by their friends or idols. Basis of Diffusion

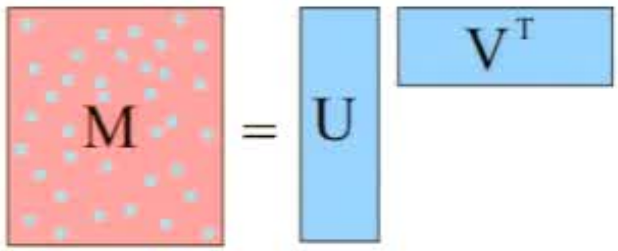
Movie Adoption



Selection and CF

$$a_{i,j} \sim P(a_{i,j}|\boldsymbol{\theta_i^T} \cdot \boldsymbol{\phi_j})$$

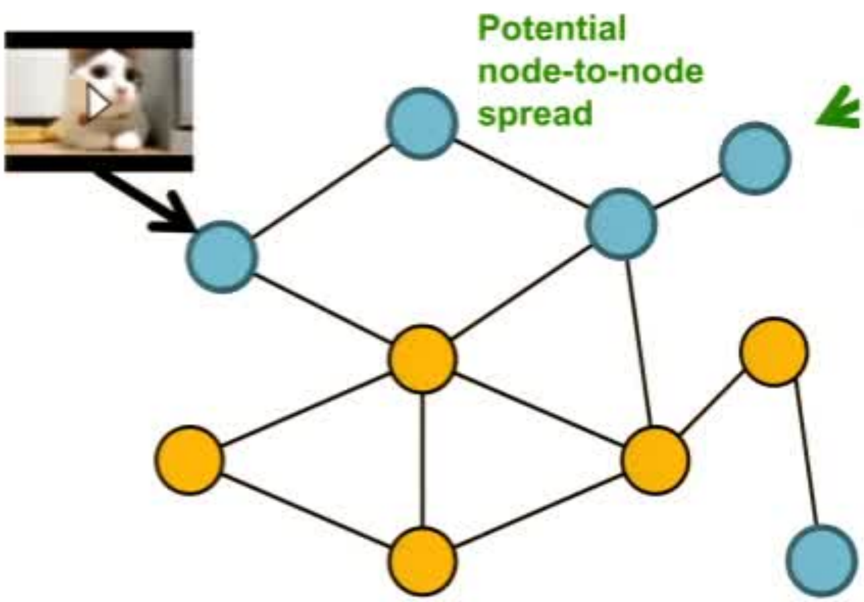
1. Do regression over the latent factors



Matrix Factorization

2. A generative perspective Latent variable model
   - dyadic responses with only positive instances can be naturally modeled by a generative process
   - stress the interpretability of the adoption process
   - intend to uncover the latent structures underlying the process

Influence and Diffusion



$$a_{i,j}^t \sim P\Big(a_{i,j}^t \mid \underbrace{\sum_{f \in i' friends} \sum_{\delta=1}^{\Delta} exp^{-\delta/\nu} \cdot \psi_{fi} \cdot I_{t-\delta}(f,j)}_{\text{Intensity}}\Big)$$

- For discrete time series, usually a logistic regression function is learned
- For continuous time series, a Poisson process is learned

Models

A simple framework: $a_{i,j} \sim P(a_{i,j}|\psi_i^T \cdot \Theta \cdot \phi_j)$

A generative process for generating $a_{ij:}$

1. first pick a friend f of user i with probability ψ_{if}:

   (1) if $f \neq i$: (social influcne):

   - generate a_{ij} by $<\theta_f,\phi_j^{f,t}>$

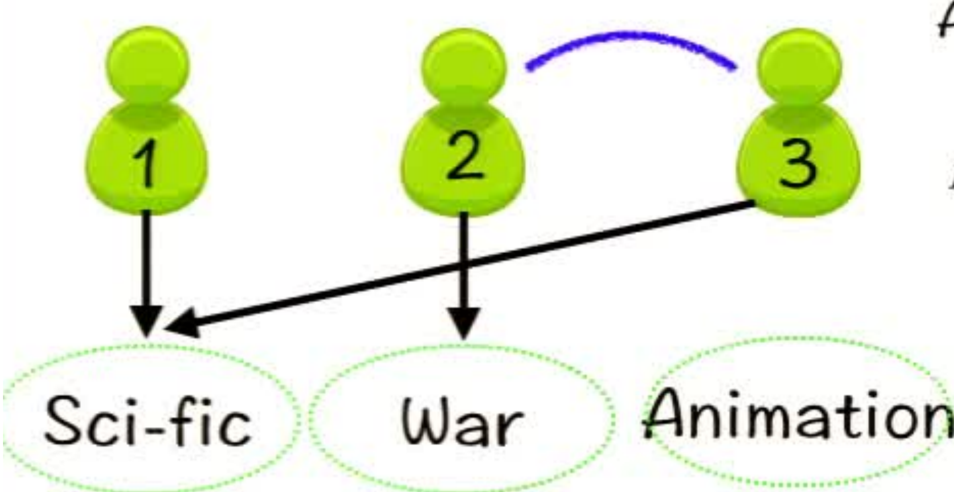
   (2) if $f = i$: (social selection)

   - generate a_{ij} by $<\theta_i,\phi_j>$

Models

A simple framework: $a_{i,j} \sim P(a_{i,j} | \psi_i^T \boxed{\cdot \Theta \cdot} \phi_j)$

1. Don't smooth user's interest by neighbors' interests
2. Smooth by both users with similar interests and neighbors



A hierarchical Polya-Urn representation

$$P(z_{i,j} = k | \boldsymbol{z}^{-(i,j)}, \eta, \alpha, \boldsymbol{r}) \propto n_{ik}^{-(i,j)} + \eta_1 \frac{m_k^{-(i,j)} + \alpha/K}{\sum_{k'} m_{k'}^{-(i,j)} + \alpha} + \eta_2 \frac{m_k^{r_i}}{\sum_{k'} m_{k'}^{r_i}} \quad ($$

Models

A simple framework: $a_{i,j} \sim P(a_{i,j} | \psi_i^T \cdot \Theta \cdot \phi_j)$

- if one adoption is caused by social selection ϕ_k
- if caused by social influence $\phi_{kj}^{f,t} | \Phi \propto \underbrace{\phi_{kj}}_{\text{Projection}} \cdot \underbrace{\sum_{\delta=1}^{\Delta} \exp^{-\delta/\tau} \cdot I_{t-\delta}(f,j)}_{\text{Intensity}} = \phi_{kj} \cdot \lambda(f,j,t)$

- project the intensity upon the latent interest space

Estimation

- Goal $\max P(\boldsymbol{a}|\mathcal{H}) = \int_\psi \int_\theta \int_\phi \sum_f \sum_z P(\boldsymbol{a}, \Psi, \Theta, \Phi, \boldsymbol{f}, \boldsymbol{z}|\mathcal{H}) d\phi\, d\theta\, d\psi$
  - Iteratively do inference of f and z, and optimize the parameters ψ and ϕ
  - Gibbs Expectation Maximization

- e-step $O(|\boldsymbol{a}| \cdot F \cdot K)$
- m-step $O(|U| \cdot F + K \cdot |V| + |\boldsymbol{a}| \cdot |V|)$

- F: average number of friends/indegree
- K: # of factors
- |U|: # of users
- |V|: # of items
- |a|: # of adoptions

Experiments

Data	# Users	# Items	# Adoptions	# Friends	# Epochs
Epinions	14,461	14,030	304,140	20	4,307
Brightkite	3,567	10,136	888,833	11	360
Flixster	19,041	8,251	1,393,906	11	360