Organizational Chart Inference

Speaker: Jim-An Tsai

Advisor: Jia-ling Koh

Author: Jiawei Zhang, Philip S. Yu, Yuanhua Lv

Date: 2016/1/7

Source: KDD'15

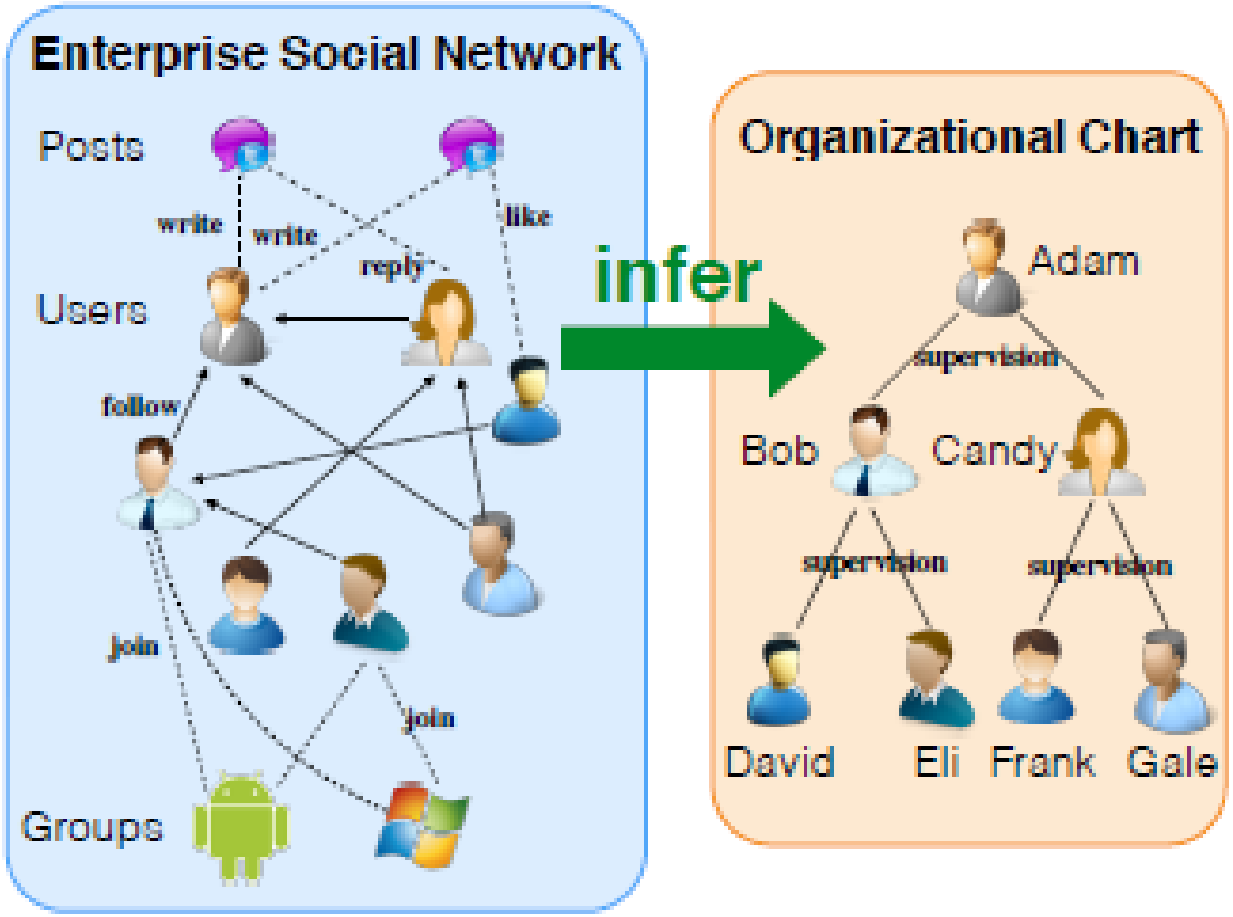
Outline

- Introduction
- Problem Proposed
- Method
- Experiment
- Conclusion

Introduction

Enterprise Social Network

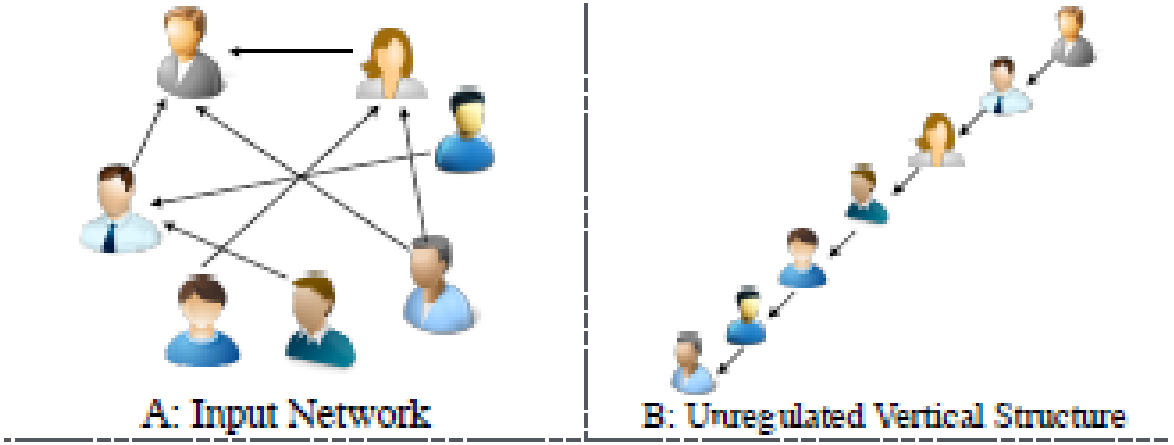
Inference of Organizational Chart



Outline

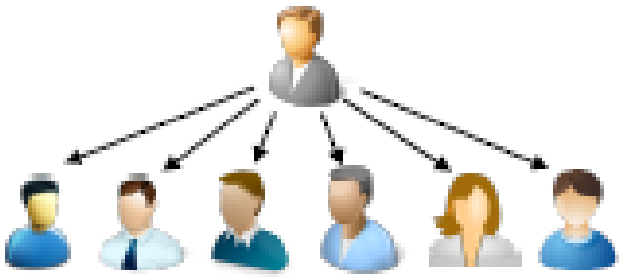
- Introduction
- Problem Proposed
- Method
- Experiment
- Conclusion

Macro-level



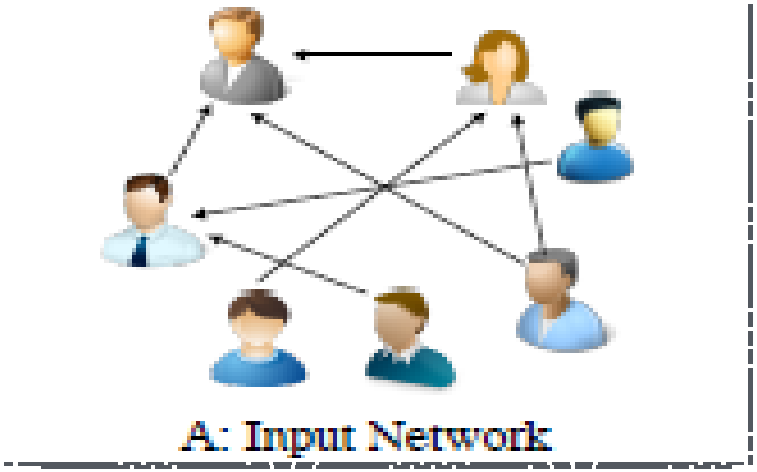
A: Input Network

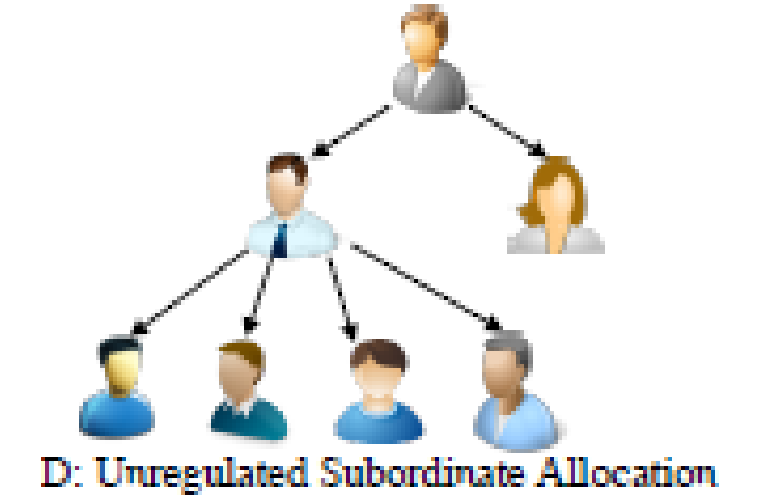
B: Unregulated Vertical Structure



C: Unregulated Horizontal Structure

Micro-level





CREATE(CHART RECOVER)

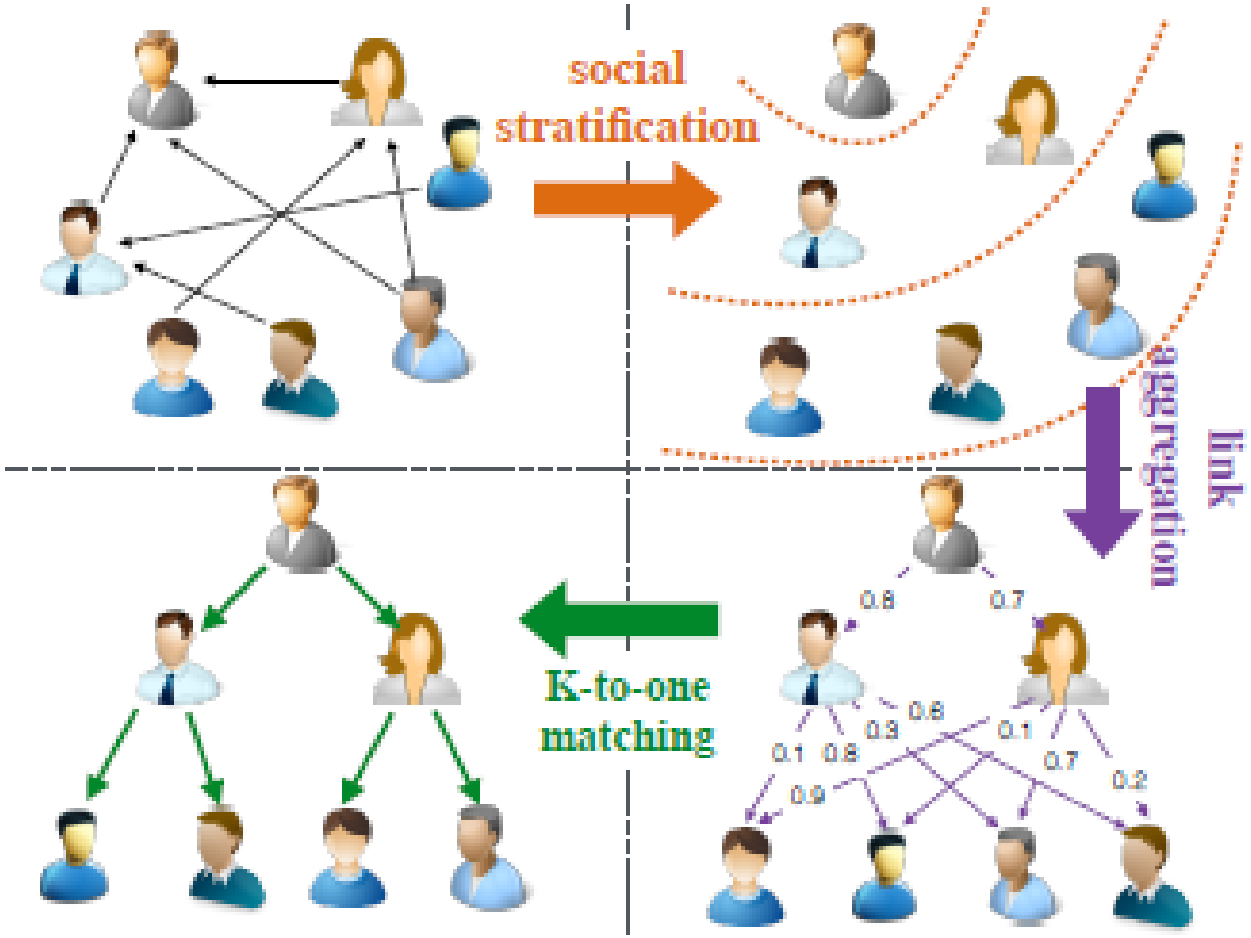


Figure 3: The framework of Create.

Outline

- Introduction
- Problem Proposed
- Method
- Experiment
- Conclusion

Regulated Social Stratification

- Social class

$$c(u) = \begin{cases} 1, & \text{if } u \text{ is the CEO;} \\ c\left(m(u)\right) + 1, & \text{otherwise.} \end{cases}$$

where $m(u)$ represents the direct manager of u.

Regulated Social Stratification

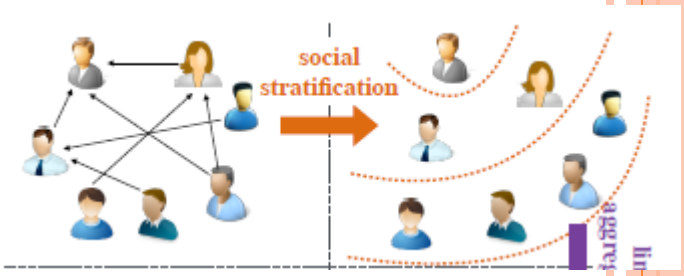
- Class Transcendence Social Link

$$p\left(c(u), c(v)\right) = \begin{cases} 0, & \text{if } c(u) > c(v) \\ c(v) - c(u) + 1, & \text{otherwise.} \end{cases}$$

The *class transcendence penalty* introduced by all social links (i.e., $\mathcal{S}$) in the ESN can be represented as

$$\begin{aligned} p\left(c(\mathcal{U})\right) &= \sum_{(u,v)\in\mathcal{S}} p\left(c(u), c(v)\right) \\ &= \sum_{(u,v)\in\mathcal{S}} \max\{c(v) - c(u) + 1, 0\}. \end{aligned}$$

Regulated Social Stratification



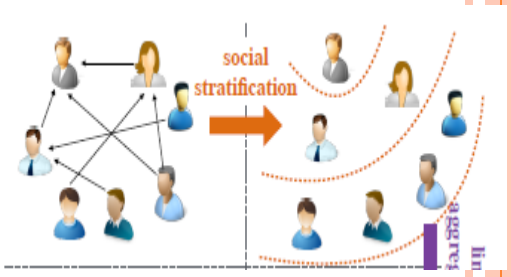
- Matthew Effect based Constraint

The *matthew effect based constraint* on users u and v can be represented as $c(u) \leq c(v)$ if $|\Gamma(u)| \geq |\Gamma(v)|$.

- Chart Depth Regulation Constraint

$$\sum_{u \in \mathcal{U}} c(u) \geq \alpha \cdot |\mathcal{U}|,$$

Regulated Social Stratification

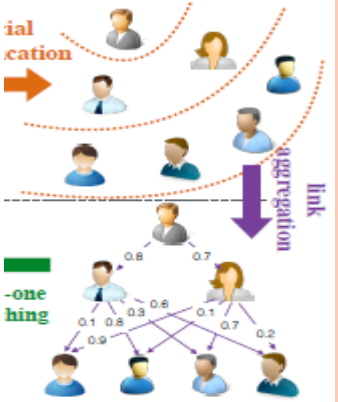


- Optimal regulated social stratification

$$c^*(\mathcal{U}) = \arg \min_{\{c(u_1), c(u_2), \cdots, c(u_{|\mathcal{U}|})\}} \sum_{(u,v)\in\mathcal{S}} p\left(c(u), c(v)\right) + \sum_{u\in\mathcal{U}} c(u),$$

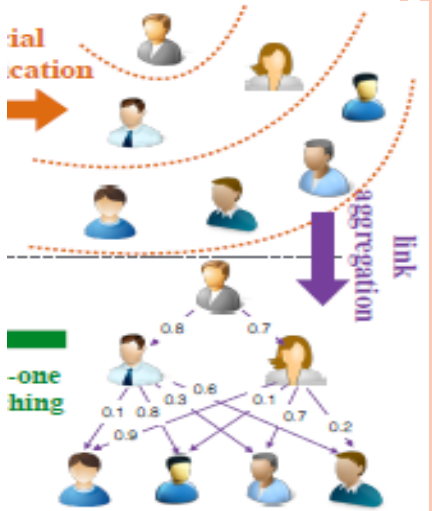
$$\begin{aligned}
s.t.,\ & p(c(u), c(v)) \geq c(v) - c(u) + 1, \forall (u,v) \in \mathcal{S}, \\
& p(c(u), c(v)) \geq 0, \forall (u,v) \in \mathcal{S}, \\
& c(u) \leq c(v), \forall u, v \in \mathcal{U}, \text{ if } |\Gamma(u)| \geq |\Gamma(v)|, \\
& \sum_{u\in\mathcal{U}} c(u) \geq \alpha \cdot |\mathcal{U}|, \\
& c(u) = 1, \text{ if } u \text{ is the CEO}, \\
& c(u) > 1, c(u) \in \mathbb{Z}^+, \forall u \in \mathcal{U} \setminus \{\text{CEO}\}, \\
& p\left(c(u), c(v)\right) \in \mathbb{Z}, \forall (u,v) \in \mathcal{S}.
\end{aligned}$$

Supervision Link Inference with Social Meta Paths Aggregation



- Social Meta Paths in Enterprise Social Networks
  - *Follow*: User $\xrightarrow{follow}$ User, whose notation is "$U \rightarrow U$" or$\Phi_1(U,U)$.
  - *Follower of Follower*: User $\xrightarrow{follow}$ User $\xrightarrow{follow}$ User, whose notation is "$U \rightarrow U \rightarrow U$" or $\Phi_2(U,U)$.
  - *Common Followee*: User $\xrightarrow{follow}$ User $\xrightarrow{follow^{-1}}$ User, whose notation is "$U \rightarrow U \leftarrow U$" or $\Phi_3(U,U)$.
  - *Common Follower*: User $\xrightarrow{follow^{-1}}$ User $\xrightarrow{follow}$ User, whose notation is "$U \leftarrow U \rightarrow U$" or $\Phi_4(U,U)$.
  - *Common Group Membership*: User $\xrightarrow{join}$ Group $\xrightarrow{join^{-1}}$ User, whose notation is "$U \rightarrow G \leftarrow U$" or $\Phi_5(U,U)$.
  - *Reply Post*: User $\xrightarrow{write}$ Post $\xrightarrow{reply}$ Post $\xrightarrow{write^{-1}}$ User, whose notation is "$U \rightarrow P \rightarrow P \leftarrow U$" or $\Phi_6(U,U)$.
  - *Like Post*: User $\xrightarrow{write}$ Post $\xrightarrow{like^{-1}}$ User, whose notation is "$U \rightarrow P \rightarrow P \leftarrow U$" or $\Phi_7(U,U)$.

Supervision Link Inference



- DP-intimacy(Directed Path-Intimacy)
- i={1,2,3,4,5,6,7}

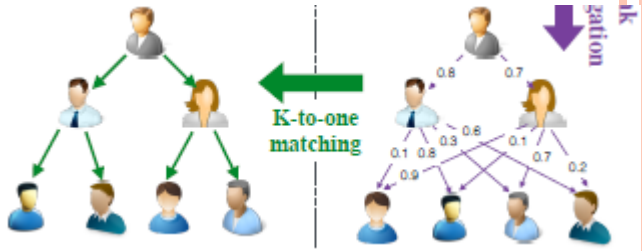
$$\text{DP-intimacy}_i(u,v) = \frac{|\mathcal{PATH}_i(u \rightsquigarrow v)| + |\mathcal{PATH}_i(v \rightsquigarrow u)|}{|\mathcal{PATH}_i(u \rightsquigarrow \cdot)| + |\mathcal{PATH}_i(v \rightsquigarrow \cdot)|},$$

where $\mathcal{PATH}_i(u \rightsquigarrow v)$ denotes the instance set of meta path $\Phi_i(U,U)$ going from u to v in the ESN.

$$intimacy(u,v) = \frac{e^{\sum_{(i)} \omega_i \text{DP-intimacy}_i(u,v)}}{1 + e^{\sum_{(i)} \omega_i \text{DP-intimacy}_i(u,v)}} \in [0,1],$$

where the value of ω_i denotes the weight of social meta path Φ_i and $\sum_i \omega_i = 1$.

Regulated Social Class Matching



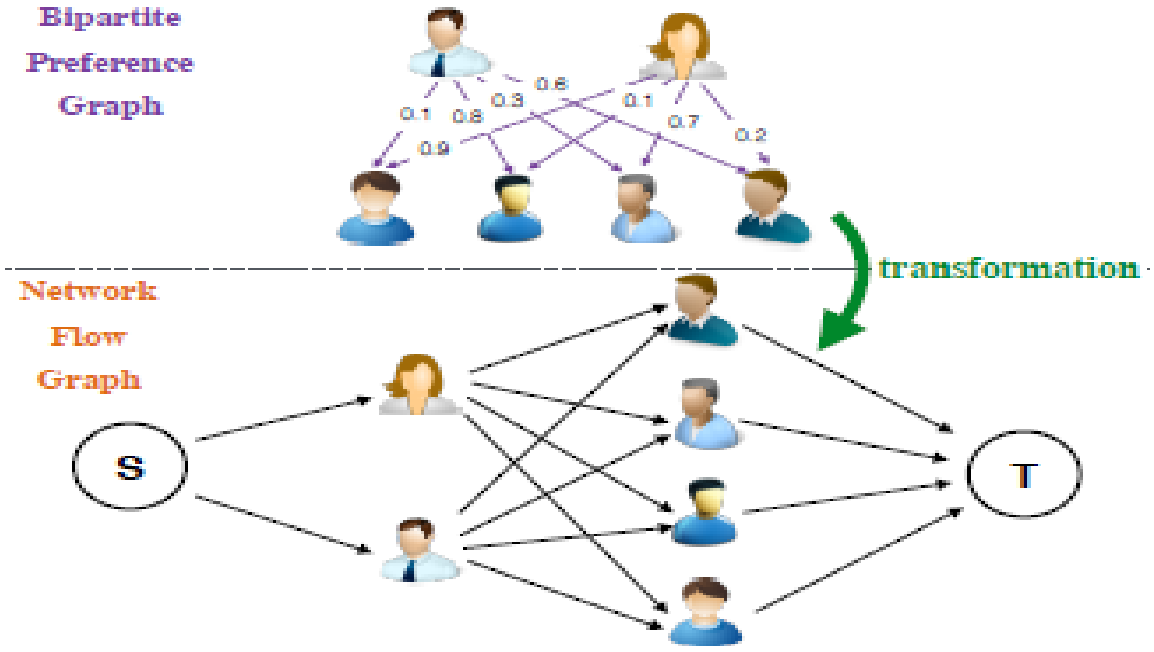


Figure 4: An example of K-to-one matching.

Outline

- Introduction
- Problem Proposed
- Method
- Experiment
- Conclusion

Experiment Dataset

100k employees

Social Stratification Results

- 先來實驗出哪個 α 最佳

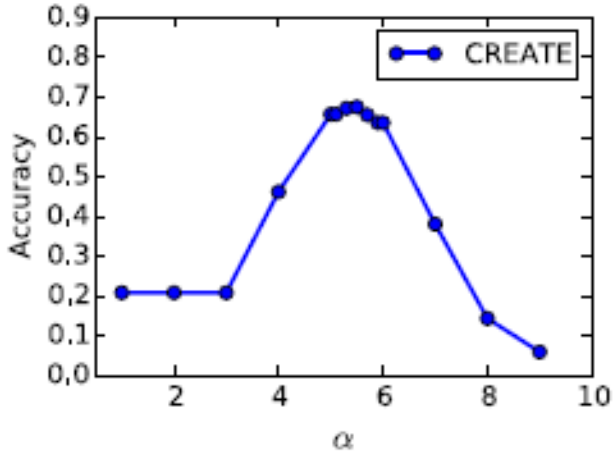
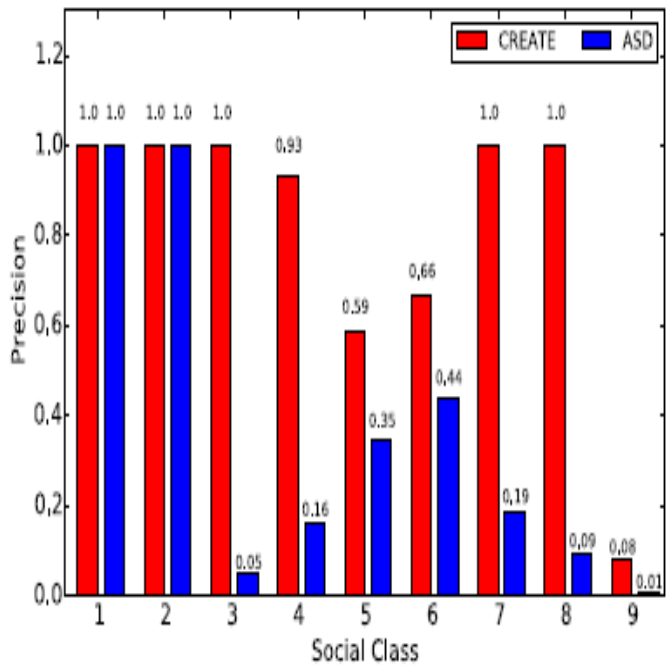


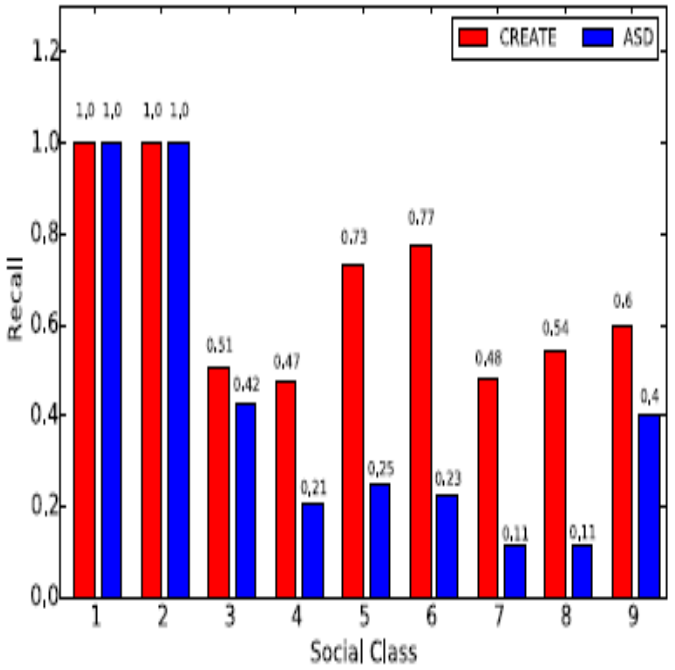
Figure 5: Sensitivity analysis of parameter α.

$\alpha = 5.5$

Social Stratification Results



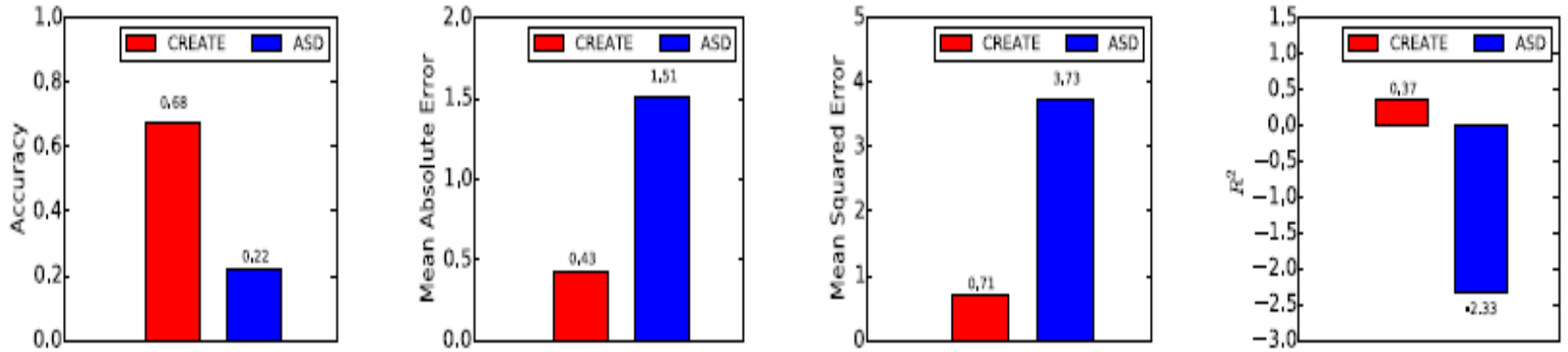
(a) Precision



(b) Recall

Figure 6: Precision and Recall achieved by Create and ASD at each social class of the organizational chart.

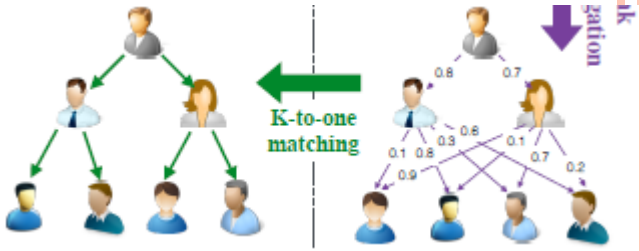
Social Stratification Results



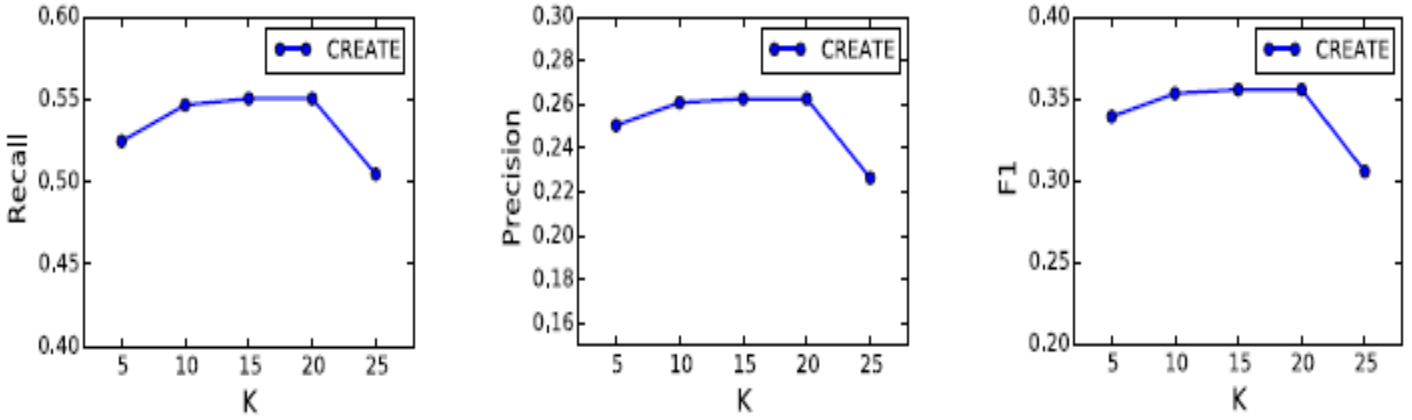
(a) Accuracy (b) MAE (c) MSE (d) R^2

Figure 7: Performance comparison of Create and ASD evaluated by different metrics.

Management Threshold Sensitivity Analysis



- 决定K的值



(a) Recall (b) Precision (c) F1

Figure 9: Sensitivity analysis of parameter K.

Organizational Chart Inference Results

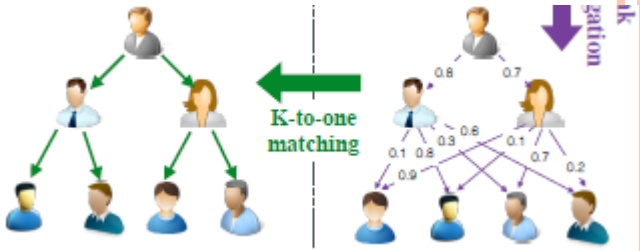
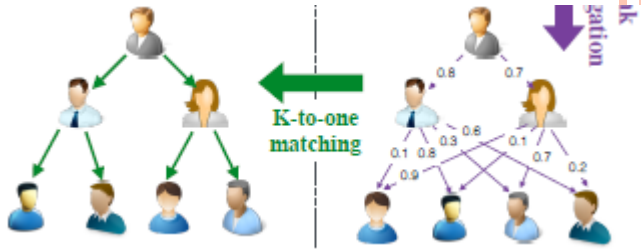
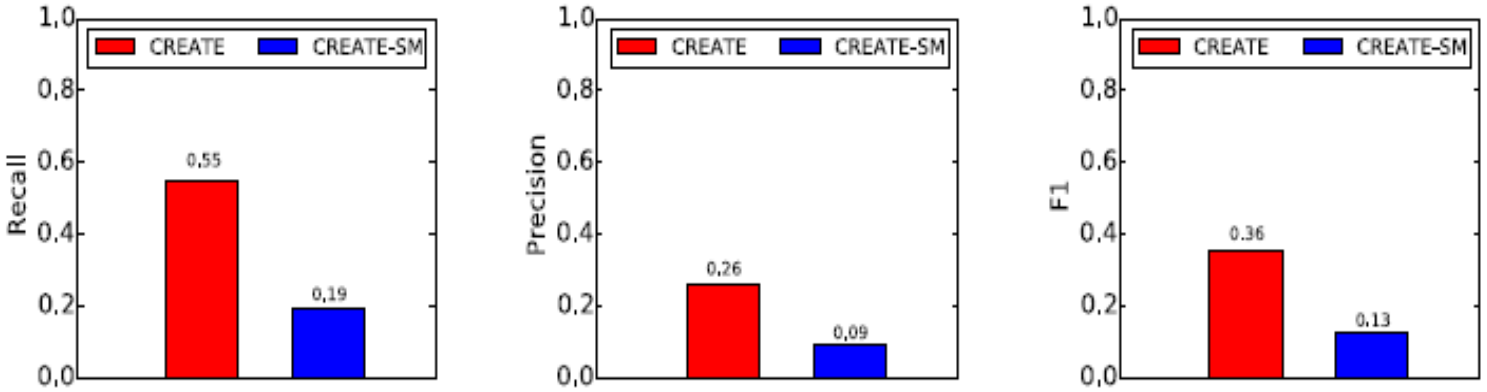


Table 1: Performance comparison of different organizational chart inference methods.

Method	Metrics	
	AUC	Precision@100
CREATE(K = 10)	0.856	0.830
CREATE(K = 15)	**0.869**	**0.870**
CREATE(K = 20)	**0.869**	**0.870**
CREATE-SL	0.719	0.820
CREATE-SM(K = 10)	0.610	0.720
CREATE-SM(K = 15)	0.630	0.790
CREATE-SM(K = 20)	0.630	0.790
CREATE-S	0.627	0.740
S-CN	0.636	0.440
S-JC	0.636	0.260
S-AA	0.528	0.070

Organizational Chart Inference Results





(a) Recall (b) Precision (c) F1

Figure 8: Performance comparison of Create and Create-sm evaluated by different metrics (K = 15).

Outline

- Introduction
- Problem Proposed
- Method
- Experiment
- Conclusion

Conclusion

- 1. We have studied the organizational chart inference (IOC) problem based on the heterogeneous online ESNs.

- 2. A new chart inference framework Create has been proposed.

Thanks for Listening