



Skill Inference with Personal and Skill Connections

Zhongqing Wang¹, Shoushan Li¹, Hanxiao Shi², and Guodong Zhou¹

¹ Natural Language Processing Lab, Soochow University, Suzhou, China

² School of Computer Science and Information Engineering, Zhejiang Gongshang University, China
 {wangzq.antony, shoushan.li}@gmail.com, hxshi@mail.zjgsu.edu.cn, gdzhou@suda.edu.cn

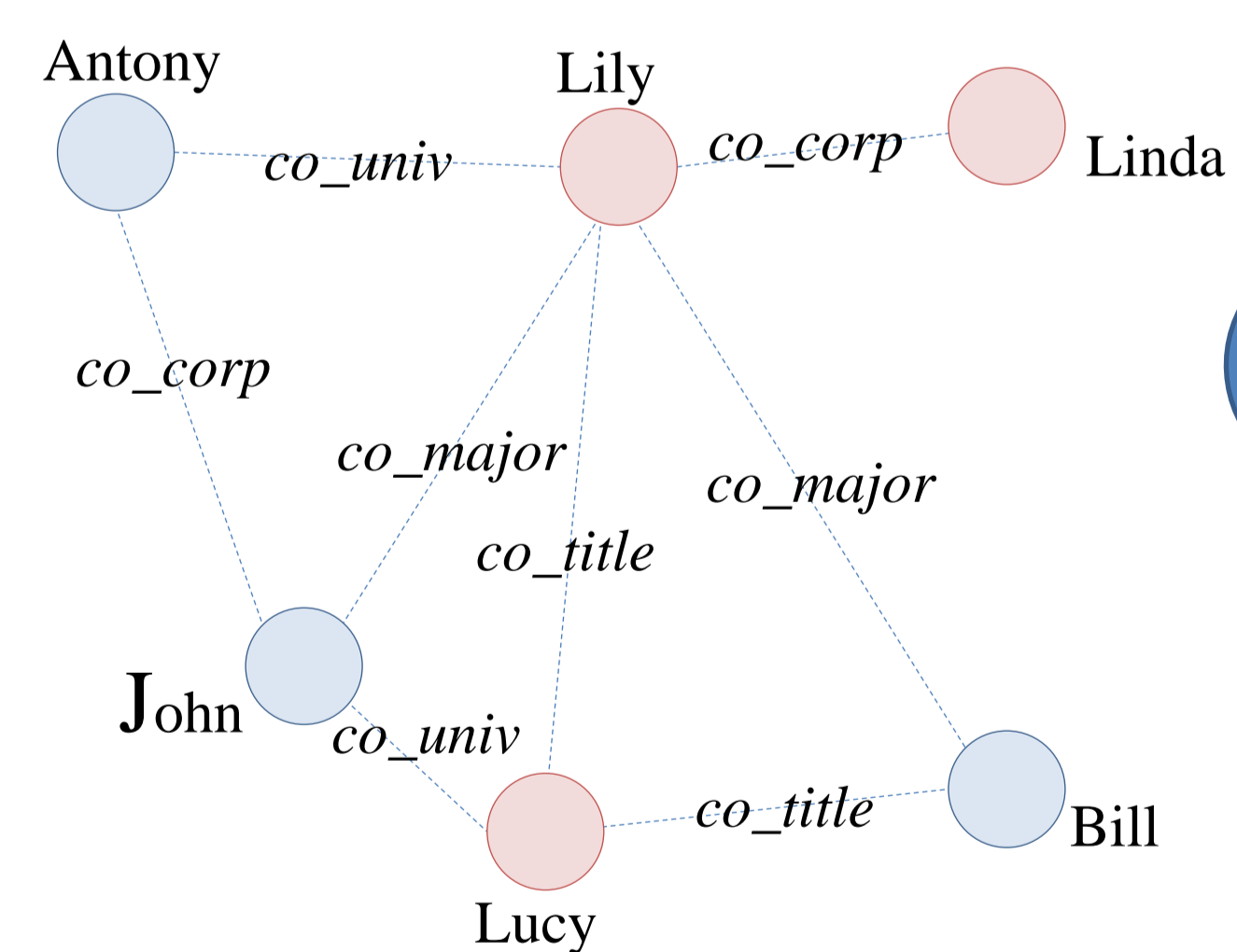
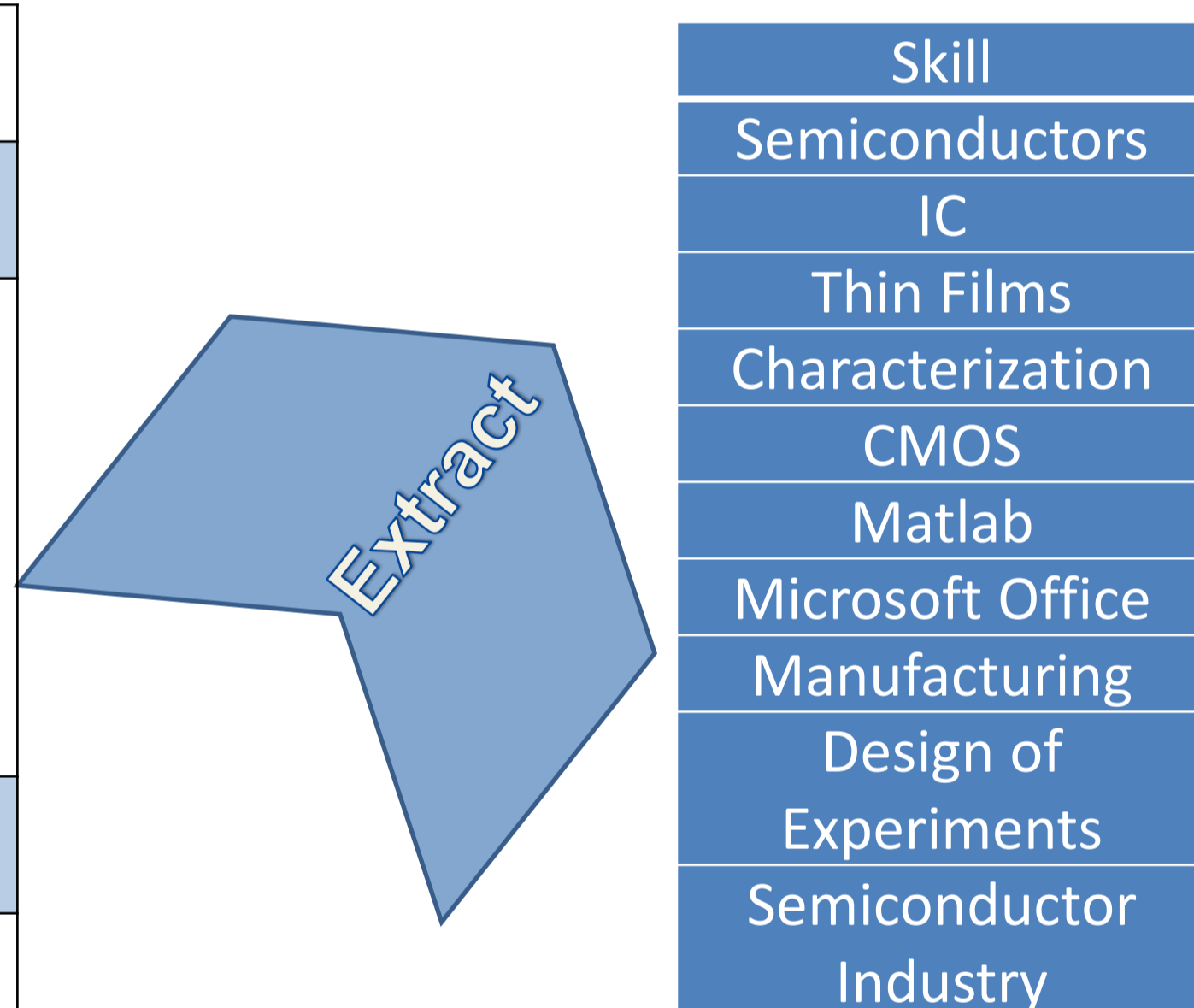


Introduction

- In this paper, we propose a factor graph based approach to automatically **infer skills from personal profile** incorporated with **both personal and skill connections**.
- We propose a **joint prediction factor graph (JPGF) model** to collectively infer personal skills with help of **personal connection factor, skill connection factor**, besides the normal textual attributes.

Data Collection and Definition

John Smith
Summary
Machine learning researcher and engineer on many fields: Query understanding...
Experience
Applied Researcher Apple Inc., Query recognition and relevance ...
Education
MIT Ph.D., Electrical Engineering, 2002 – 2008 ...



We collect our data set from **LinkedIn.com**

- Textual information.** We use texts of *Summary* and *Experience* as the textual information.
- Personal connections.** We extract four kinds of personal relationships between two persons (*co_univ*, *co_maj*, *co_corp*, *co_title*).
- Skill connections.** We extract skill connections from same person

Joint Prediction Factor Graph Model

Textual attribute function

$$\frac{1}{Z_1} \exp \left(\sum_i \sum_k \alpha_k f_k(x_{ik}, y_i) \right)$$

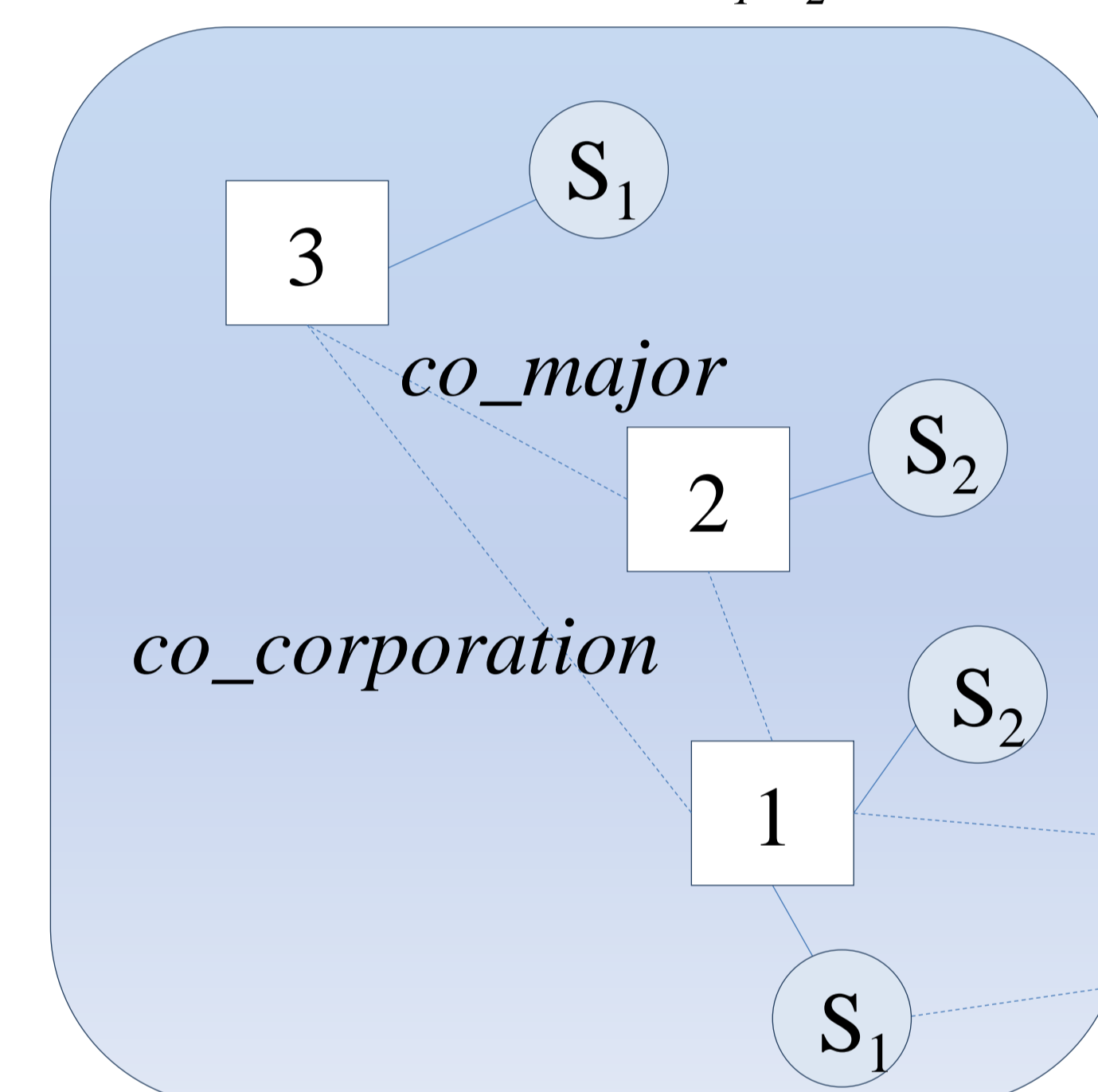
Personal connection factor function

$$g(y_i, y_j) = \exp \left\{ \beta_{ij} (y_i - y_j)^2 \right\}$$

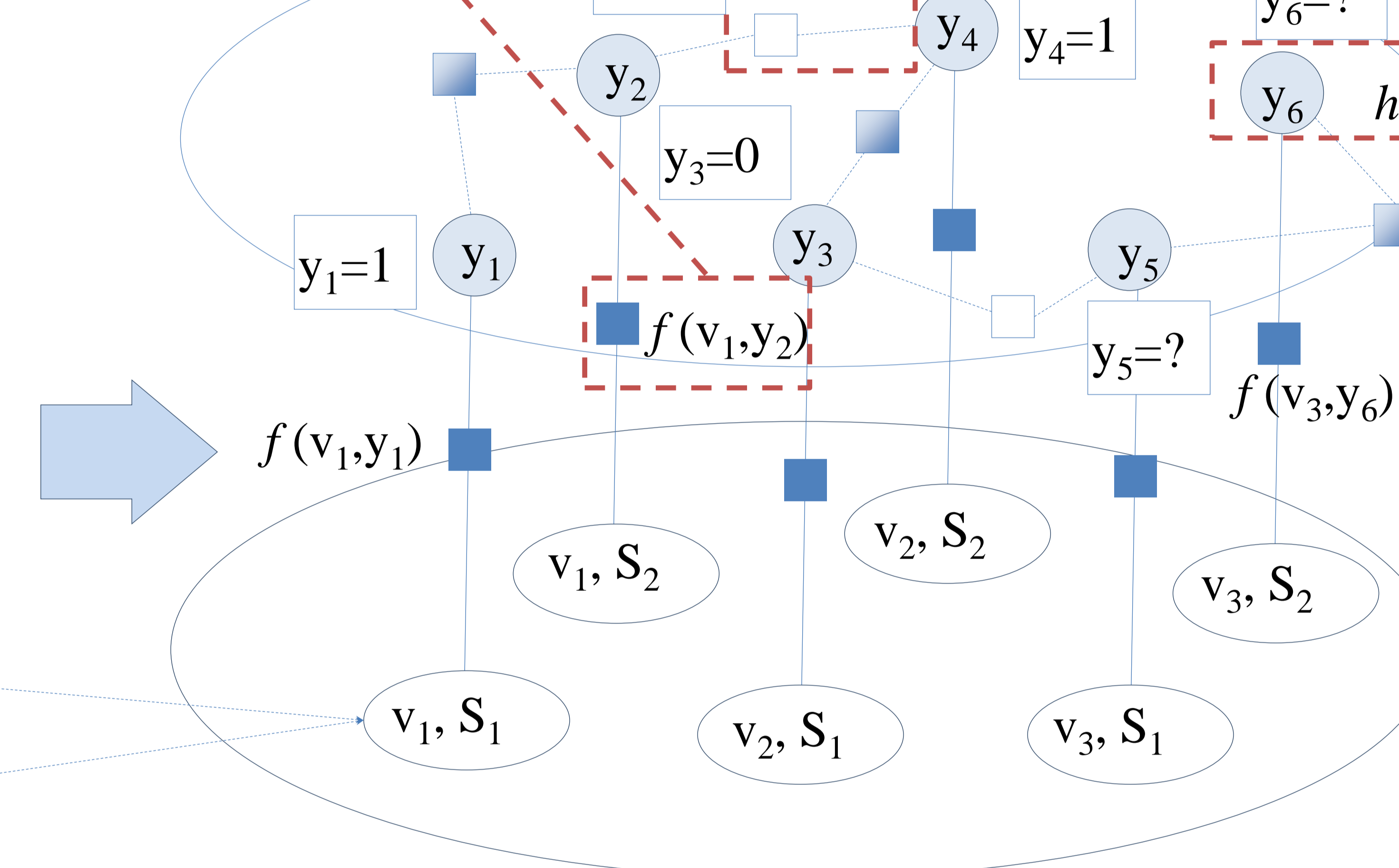
Skill connection factor function

$$h(y_i, y_j) = \exp \left\{ \gamma_{ij} (y_i - y_j)^2 \right\}$$

Nodes of people with different kinds of skills (S_1, S_2)



JPGF model



Learning objective function

$$\begin{aligned} L(\theta) &= \log_{\theta} P(Y | X, G) \\ &= \frac{1}{Z_1} \sum_{i=1}^{|E|} \sum_{j=1}^d \alpha_j f(x_{ij}, y_i) \\ &\quad + \frac{1}{Z_2} \sum_i \sum_{j \in NB(i)} \exp \left\{ \beta_{ij} (y_i - y_j)^2 \right\} \\ &\quad + \frac{1}{Z_3} \sum_i \sum_{j \in SAME(i)} \exp \left\{ \gamma_{ij} (y_i - y_j)^2 \right\} - \log Z \end{aligned}$$

Experimentation

- we randomly select 2,000 profiles as training data and 1,000 profiles as testing data.
- we select balanced testing and training samples for each skill.
- For performance evaluation, we adopt Precision (P.), Recall (R.) and F1-Measure (F1.).

