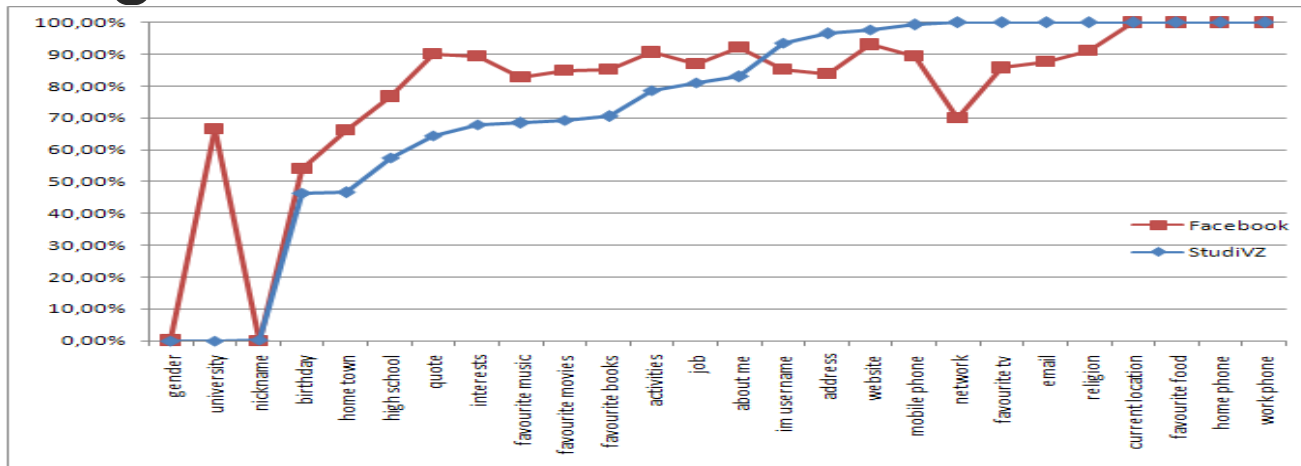


# Easy Exposure of Privacy in Social Network through Semi-Supervised Learning

Michelle, Mingzhen, Dingyan, Baichuan

# Motivation

- Privacy Preserving: Three Hundred Million Active Users in Facebook
- Semi-Supervised Learning: few users have completed data (labeled) and a large number of users do not.



# [ Characteristic of Data ]

- Two independent Views: Personal and Relationship Information
- Relationship Information: Graph Structure
- Personal Information: Statistic Models

# [ Methods ]

- Co-Training: combine confident predicting results from two views
- Graph-Based SSL: using harmonic restricting function
- Advanced Graph-Based SSL: local and global consistency

# Information Integration

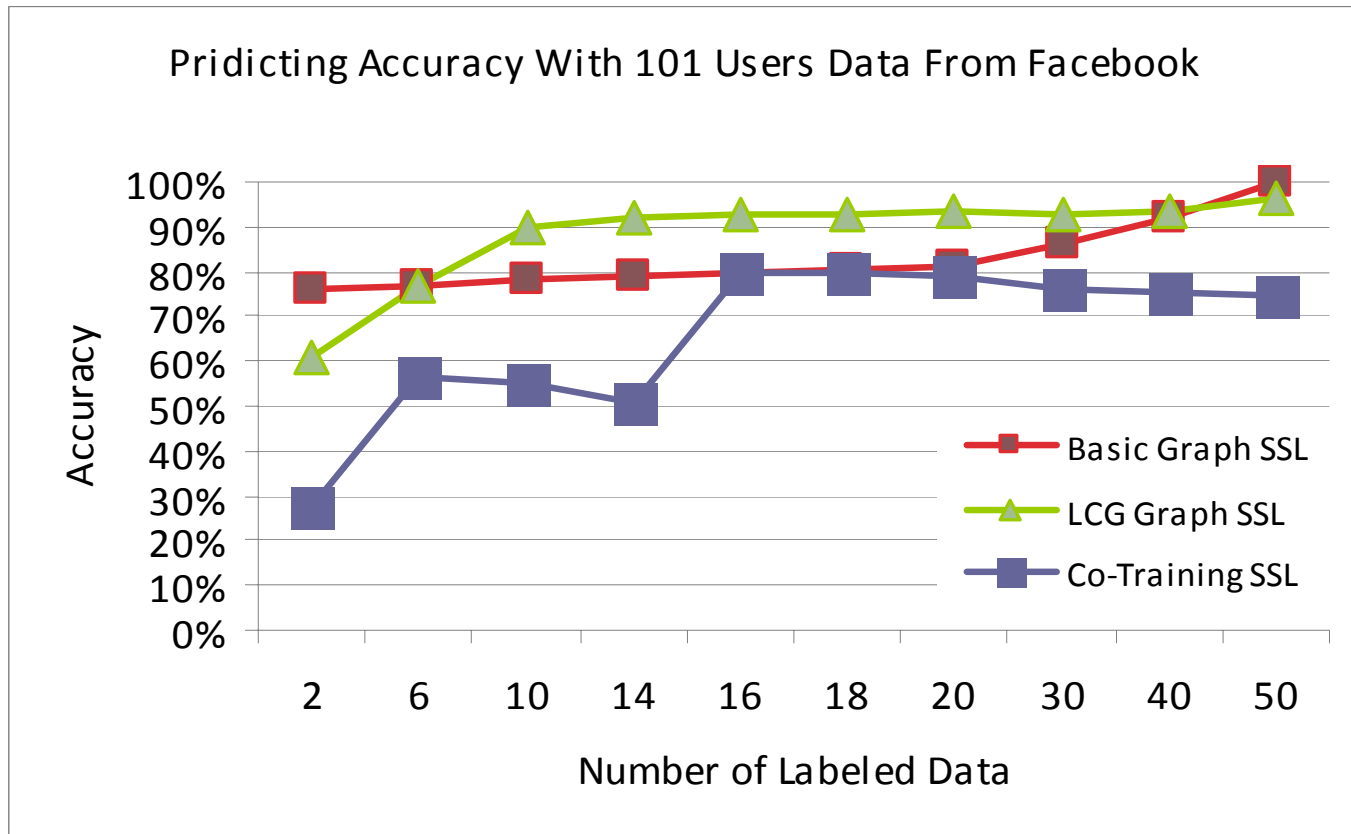
- View of Personal Information

$$P_i = f(\text{age}) \times (\lambda W_{\text{telephone}} + (1 - \lambda) W_{\text{location}})$$

- View of Relationship Information

$$W_{i,j} = \lambda_1 W_{\text{friendship}} + \lambda_2 W_{\text{group}} + \lambda_3 W_{\text{network}}$$

# [ Result ]



# [ More Experiments ]

- A Germany Social Network: data size is about 1,000
- Another Facebook Experiment: data size is about 200,000
- [Option] Self-building data: data size is about 10,000



Thank You