On Going Research Works

Mo Mingzhen

CS&E, CUHK

Outline

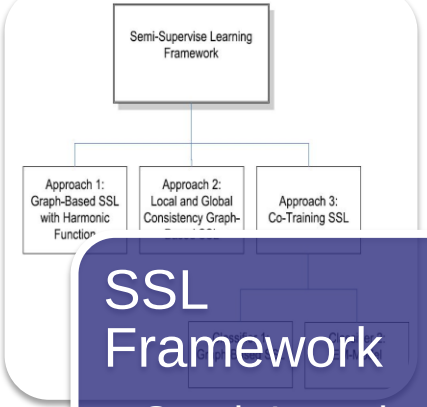
- Previous Work
- Current Work
  - Privacy Preserving
  - Multiple Graphs Combination
- Summary

PREVIOUS WORK

Previous Work

Social Network Data

- Basic Profile Info.
- Relational Info.



SSL Framework

- Graph Learning
- Co-Training Learning

Objective

- Breaking Security

Multiple Graphs Combination

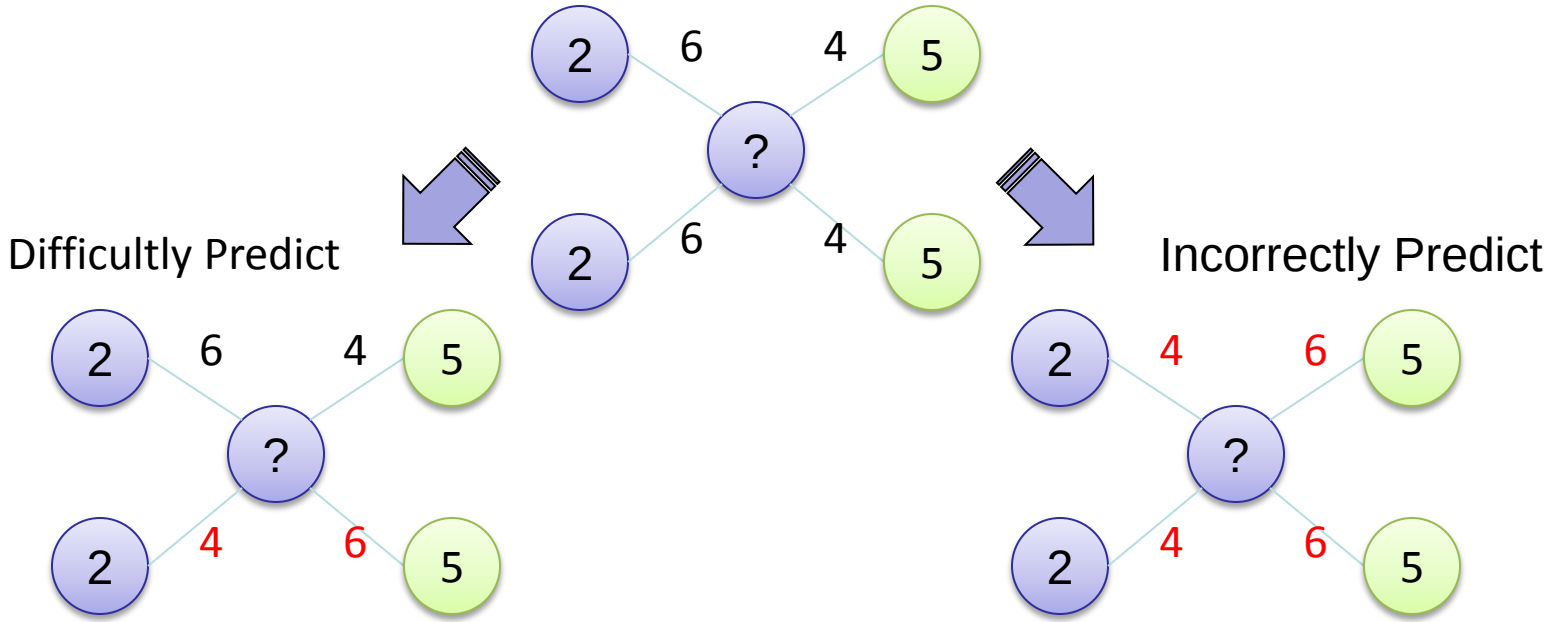
Privacy Preserving

CURRENT WORK

Privacy Preserving

Privacy Preserving

- Objective
  - Be difficult to predict
  - NOT to predict incorrectly

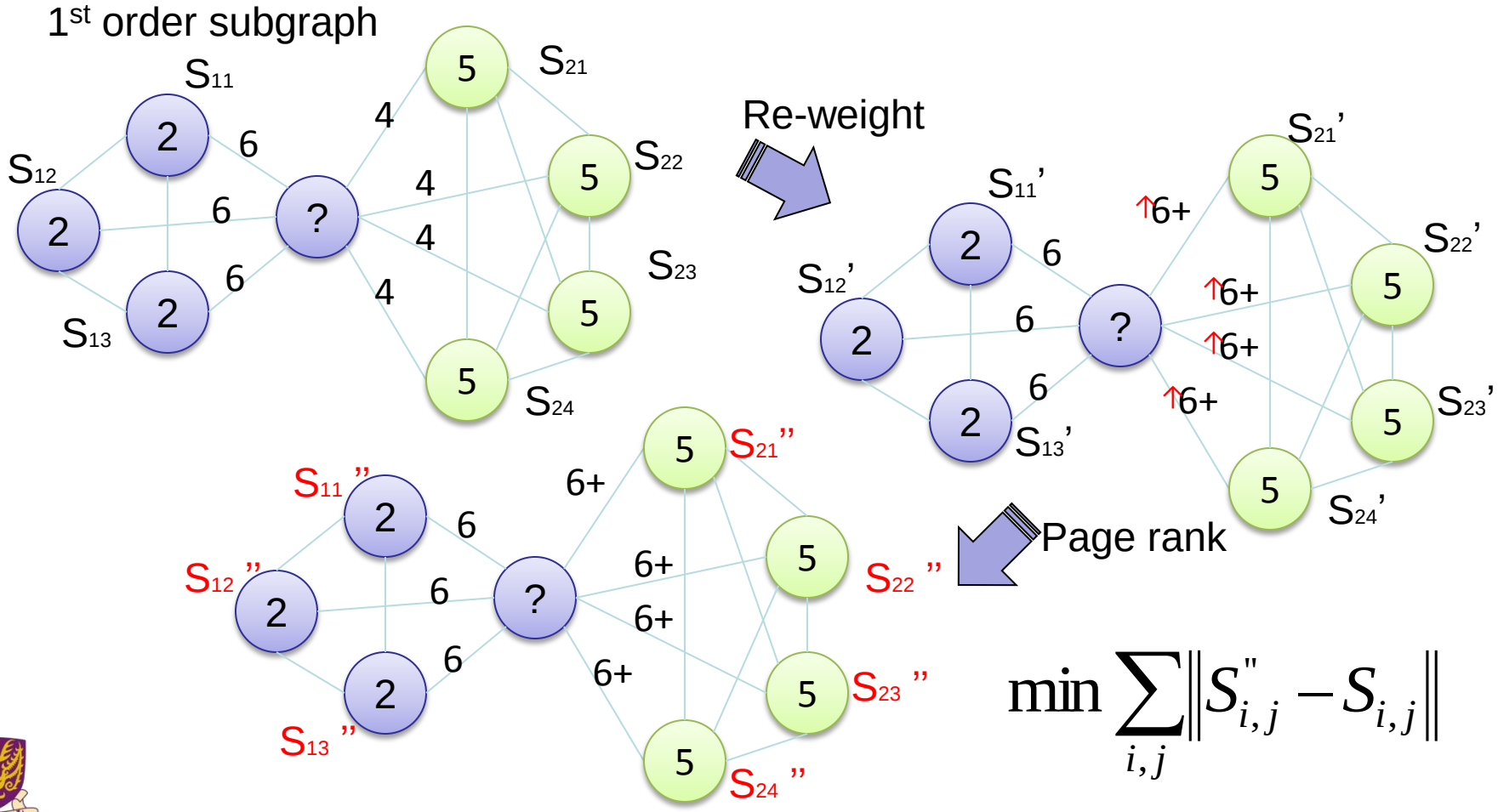


Privacy Preserving

- Methods
  - $G(E, V)$
  - Edge: Changing weights of graph
  - Node: Supernode

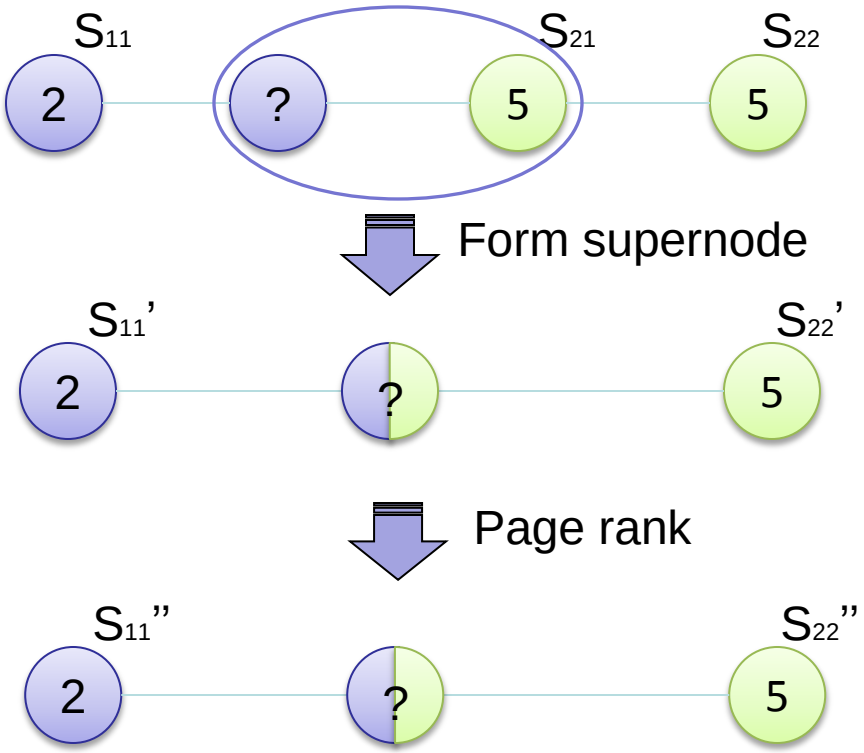
Privacy Preserving

- Changing weights of graph



Privacy Preserving

- Supernode



$$\min \sum_{i,j} \left\| S''_{i,j} - S_{i,j} \right\|$$

Privacy Preserving

- Current Result
  - StudiVZ data
    - Size: 1400
    - 6 Classes
    - 80% - 90% labeled
  - Learning Methods
    - Supuevised Learning (SL)
    - SSL
      - Basci Graph
      - Local & Global Consistency (LGC) Graph
  - Average of accuracy

Accuracy	SL	Basic Graph	LGC Graph
Before	75%	95%	95%
After	0%	5%	20%

Privacy Preserving

- On-going Part
  - Current methods are local strategies (1st order subgraph)

    and that's the angle of view of network security
  - Think of global strategy, considering Laplacian Smoothing (Machine Learning)

Accuracy	SL	Basic Graph	LGC Graph
Before	75%	95%	95%
After	0%	5%	20%

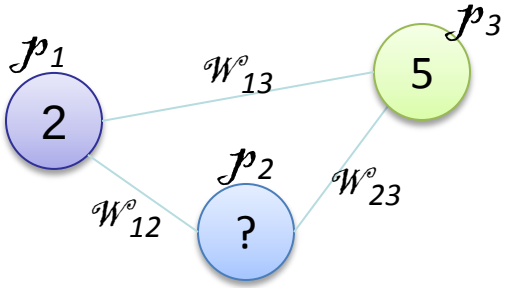
Because of global consistency of learning

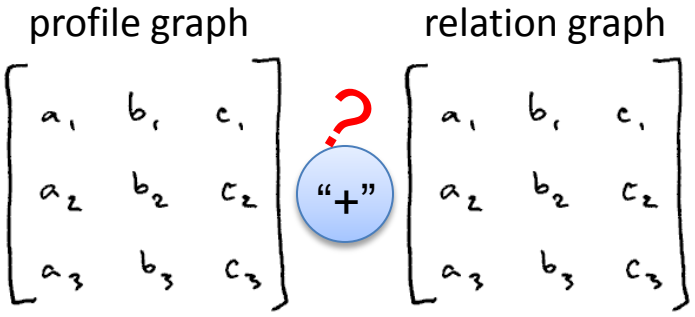
CURRENT WORK

Multiple Graphs Combination

Multiple Graphs Combination

- Objective
  - Better combine profile info. graph and relational info. graph.





  - Looking for a more reasonable and effective method, besides linear combination and Co-Training framework.

Multiple Graphs Combination

- Method
  - Similar Matrix Factorization

$$\min_{L,F} \frac{1}{2}\left\|X - LF^T\right\|^2 + \frac{\alpha}{2}\left(\|L\|^2 + \|F\|^2\right) + \frac{\beta}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\left\|L_{i*} - L_{j*}\right\|^2$$

Cost function: profile info.

Regularization: relation info.

L (nxc) is the label matrix, F (fxc) is a feature matrix,
X (nxf) is the profile info. matrix
A (nxn) is the adjacency matrix for relational info.

n : # of data, c : # of class, f : # of feature

Multiple Graphs Combination

- On-going Part
  - Writing codes to solve this optimization problem
  - Apply on the real data sets

SUMMARY

Summary

- Two on-going works, based on a previous one
- Network, Privacy Preserving
  - Finding a global strategy
- Learning, Graphs Combination
  - Writing codes and doing experiments

Thank You