

# Project Proposal

*Project name*

***Ting Zhang and Wen Yi***

School of Industrial Engineering, School of Electrical and Computer Engineering  
zhan1013@purdue.edu, yi35@purdue.edu

September 29, 2015

## 1 Introduction

*What is the project*

*Motivation*

## 2 Related work

### 2.1 Relationship Inference

Relationship inference in social networks has been studied in various fields and domains. In this context, we refer to friendship inference between pairs of individuals. Representing social networks with topology structures provides insights to predict relationships between individuals based on topology and probability distribution of the links in the topology. [Liben-Nowell and Kleinberg](#)

### 2.2 Community Detection

In previous research such as the paper by [Xie and Szymanski](#), from the analysis of individual data collected, community detection in social networks focus on grouping the people into overlap or distinct communities while the people from same community have stronger general relationships or share the similar characteristics. Further more, in the paper by [Mislove et al.](#), community detection were used to inferring the profile information of people based on the profile information from other members in the same community, while some of the members in community have profile information vacant.

## 3 Problem formulation

*Describe your project as a machine learning problem, identify inputs objects, labels, possible features*

## 4 Data and Evaluation plan

*Describe the data you intend to use. Mention if there is an existing data source you intend to use, or if annotation is required*

*How will you evaluate your algorithm? What is a reasonable baseline?*

### Submission Instructions:

*delete this section when submitting*

You are required to use L<sup>A</sup>T<sub>E</sub>X to type your solutions to questions, and report of your programming as well. Other formats of submission will **not** be accepted. A template named “homework.tex” is also provided for your convenience.

After logging into data.cs.purdue.edu (physically go to the lab or use ssh remotely, you are all granted the accounts to CS data machines during this class), please follow these steps to submit your assignment:

1. Make a directory named ‘*your Name\_your Surname*’ and copy all of your files there.
2. While in the upper level directory (if the files are in /homes/dan/dan\_goldwasser, go to/home-s/dan), execute the following command:

```
turnin -c cs578 -p PROPOSAL *your_folder_name*
```

(e.g. your prof would use: `turnin -c cs578 -p PROPOSAL dan_goldwasser` to submit his work)

Keep in mind that old submissions are overwritten with new ones whenever you execute this command.

3. You can verify the contents of your submission by executing the following command:

```
turnin -v -c cs578 -p PROPOSAL
```

Do **not** forget the -v flag here, as otherwise your submission would be replaced with an empty one.

## References

David Liben-Nowell and Jon Kleinberg. The link-prediction problem for social networks. 58(7): 1019–1031. ISSN 1532-2882. doi: 10.1002/asi.v58:7. URL <http://dx.doi.org/10.1002/asi.v58:7>.

Alan Mislove, Bimal Viswanath, Krishna P. Gummadi, and Peter Druschel. You are who you know: Inferring user profiles in online social networks. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, WSDM '10, pages 251–260. ACM. ISBN 978-1-60558-889-6. doi: 10.1145/1718487.1718519. URL <http://doi.acm.org/10.1145/1718487.1718519>.

Jierui Xie and B.K. Szymanski. Community detection using a neighborhood strength driven label propagation algorithm. In *2011 IEEE Network Science Workshop (NSW)*, pages 188–195. doi: 10.1109/NSW.2011.6004645.