**Detecting Tight Communities in Twitter**

**By**:

Aparna Pavithran

Chandni Shankar (cxs152830)

Mahesh

Manasa

Contents

[1 Introduction 2](#_Toc468010880)

[2 Approach 3](#_Toc468010881)

[3 Experiment and Results 3](#_Toc468010882)

[4 Future Work 4](#_Toc468010883)

[5 Conclusions 4](#_Toc468010884)

[6 References 4](#_Toc468010885)

# Introduction

Social network has increasingly become a commodity and a necessity for people around the world on one hand and on the other, catering to customer specific interests have become a business. This gives rise to a need to understand the people’s activities, their associations and their interest groups which would further help businesses like Pinterest, Twitter, Google , Facebook to customize the web pages specific to each individual’s liking. Detecting tight communities have become important to understand each person’s engagements and suggest helpful customizations that would interest the user and increase involvement in using the app.

This project aims at detecting tight communities in Twitter social network. The Twitter data contains 1000 ego-networks consisting of 4869 circles and 81362 users. The people in the Twitter social network are represented as nodes and the connection between two people are represented as edges in the graph that represents the entire Twitter data.

# Dataset Description

Twitter data was crawled from public sources. The dataset includes node features (profiles), circles, and ego networks. Dataset contains .txt file with node id and follower id of twitter users.

Twitter data is obtained from 1,000 ego-networks, consisting of 4,869 circles and 81,362 users. The ego-networks has 81306 nodes and the directed edges indicate person a follows person b. The set of circles for each node is the all the groups the user is a part of and each circle is a set of nodes indicating all the users of that group and their connection. These circles were obtained by using the features mentioned in the feature name file.

We used the file that contained all the edges i.e. all the connections between every user in the Twitter data to construct a graph of the social network with each user as a node and the connection of that user with other users as an edge in the graph.

# Related Work

# Proposed Solution

# Experiment and Results

# Future Work

# Conclusions

# References