COMPSCI 2XB3:Computer Science Practice and Experience: Binding Theory to Practice
Project Proposal Template

Project Title:	*Your_best_friend.io*
Lab Section Number:	**L02**
Student Name:	**Yunbing (Peter) Weng**
Student Number:	**400158853**
Student Email:	**Wengy12@mcmaster.ca**

By virtue of submitting this document I electronically sign and date that the work being submitted is my own individual work.

Abstract

This project will be based on two graphs provided by Stanford University: Gowalla[1]. So one of them is time and location of check-in, another is friendship network.
The problem that will be solved in this project is to find out who spend most time with you, for example this week, or someone start to spend time with others but telling you that they don't have time? In order to solve this problem, I will use two dataset[2,3] at the same time, when we want to check some user, we will search for his friends, find where his friends visited last week and them sort the array of friends by the number of common check-in places.

1.Objective

Actually its useless project, don't think that many people will be interesting in it but it is very good practice in using sorting, searching and graph theory at the same time. Also the graphs are big enough to think about optimization.

2. Motivation

I remember that sometimes I'm asking myself question, why someone is not going out with me and say that he(she) has hard time in study process. Is this just polite way to say that I don't want to go out with you, or he(she) really have some problems? There is other way to use this project: analyze whole friendship network and have some statistic about how many groups of friends do people usually have? Does it relate to age or something else? May be there is some interesting phenomena. So, this can be targeted to teenagers, who whom all this is important, and also for me, just in sense of scientific interest.
I expect to have working first part (sort friends) and also hope to find something interesting about second part.

3. Prior Work

Actually I'm not sure that there are something similar. But for second part, there is one good app[4] to show the graph of friends on VK[5] (Russian social media like Facebook) and they put the edges (friends) closer if you more often massage them and it puts edges with friendship together and it seems very interesting.

4. Input/output and proposed solutions

http://snap.stanford.edu/data/loc-gowalla_edges.txt.gz[2]
this is friendship network of Gowalla network
http://snap.stanford.edu/data/loc-gowalla_totalCheckins.txt.gz[3]
Check-in time and location of all users

For first part of output it will be sorted array of friends and how many times you goes out together on the week, for second I will try to make some sort of visualization to make it understandable for me.

The first step is very easy, the way it gives us information is "id1 id2" on each line, and first id is sorted. So it is easy to find someone. And then, we got array of his friends, using the second dataset, which has
"`[user][check-in time][latitude][longitude][location id]`"
This information on each line, and its sorted by user_id and then by time.
I plan for each user, get all location within sometime, and add it into another array which will represent the location_id. That would be 2D array and contain Location, [user][time]
Then, I will sort it

5. Algorithmic challenges:
Definitely some custom made sorting system. I don't really know about runtime at this moment, but probably will try to optimize algorithm as good as I can. May de will try to use self-balanced tree somewhere to store information

6. Project plan

Week 1 – do the first step (get array of friends)
Week 2 – do second step (get 2D array with Location and [user][time])
Week 3 – sort it
Week 4 – debug
Week 5 – optimize algorithm and probably try to use more efficient datatype
Week 6 – debug again and if have time, do second part of program
Week 7 – debug and make it presenteble

References
[1] **http://snap.stanford.edu/data/loc-Gowalla.html**
[2] **http://snap.stanford.edu/data/loc-gowalla_edges.txt.gz**
[3] **http://snap.stanford.edu/data/loc-gowalla_totalCheckins.txt.gz**
[4] **https://habr.com/ru/post/243229/**
[5] **https://vk.com/**