# Midterm Report

# SIGMOD programming contest 2014

## Progress Report

We have got 4 queries working in a non-optimized manner. The queries have been tested against the small dataset supplied by the organizers of the programming contest. Timing results have been collected.

#### Query Type 1 (Shortest Distance Over Frequent Communication Paths)

<short description of approach taken>

#### Query Type 2 (Interests with Large Communities)

<short description of approach taken>

#### Query Type 3 (Socialization Suggestion)

<short description of approach taken>

#### Query Type 4 (Most Central People)

<short description of approach taken>

## Refined Project Milestones

#### April 1 (Mid-term report)

Done: Got all 4 queries working in a non-optimal fashion

#### April 12

Optimize the first two queries

#### April 24

Optimize all the queries individually

#### April 29 (Project presentation)

Try to bunch optimizations such that all of them run well together.

#### May 8

Finish the project report

## Timing Results (all times in s)

#### Query 1

Load time 7354.18300009

Individual runs:

576 400 -1 time 38.6770000458

58 402 0 time 34.5739998817

266 106 -1 time 33.6689999104

313 523 -1 time 34.3120000362

858 587 1 time 54.8840000629

155 355 -1 time 34.9470000267

947 771 -1 time 28.3529999256

105 608 3 time 82.4440000057

128 751 -1 time 31.8029999733

814 641 0 time 46.3350000381

Total q1 time 420.001000166

#### Query 2

Load time 32.6449999809

Individual runs:

3 722846 time 80.8550000191

4 723249 time 61.5350000858

3 723633 time 56.0460000038

3 724039 time 60.3880000114

5 724459 time 41.0669999123

3 724792 time 34.0409998894

3 725171 time 38.2730000019

7 725546 time 23.4509999752

3 726051 time 14.3300001621

4 726492 time 8.10899996758

Total q2 time 418.09800005

#### Query 3

Load time 27.6159999371

Individual runs:

3 2 Asia time 473.381000042

4 3 Indonesia time 1.94899988174

3 2 Egypt time 0.444999933243

3 2 Italy time 0.311000108719

5 4 Chengdu time 0.137000083923

3 2 Peru time 0.157999992371

3 2 Democratic\_Republic\_of\_the\_Congo time 0.108999967575

7 6 Ankara time 0.0980000495911

3 2 Luoyang time 0.0899999141693

4 3 Taiwan time 0.131000041962

Total q3 time 476.811999798

#### Query 4

Load time 440.994999886

Individual runs:

3 Bill\_Clinton time 9849.704

4 Napoleon time 6724.16999984

As the runs of the 4th query take so long, we are unable to produce timing results for the rest. Needless to say, we need to optimize the 4th query quite a lot.