Contact Tracing

Problem Description

Under the Covid-19 situation, we could collect contact traces from mobile apps via Bluetooth beacons. Suppose we have such contact traces as below:

PID1	PID2	Time
11	10	32
21	11	42
12	13	55
13	21	66

In the traces, person with ID 11 contacted another person with ID 10 at time 32.

Now you are required to handle such traces which may have up to billions of records.

Task

Write a program to achieve the following functions:

1. For a particular person A, find all other persons who contacted A after a particular time T.
2. Parallelise the task with multithreading and show the performance improvement related to the number of threads used.
3. Suppose you have a group of affected persons with their affected times, find out other persons who have risks of affection. Discuss the possible approaches to lable the risks and implement one approach in your program.
4. Write a report (less than one page) to show the performance of the parallel implementation, and explain your risk-labeling approach and its implementation.

Relates to Objectives

1.1, 1.2, 2.1, 2.2, 2.6, 3.1, 3.4, 3.5, 3.6, 4.1, 4.2, 4.3, 4.7, 4.8

(Pair)