

HACKATHON PROBLEM STATEMENTS

Marks: 50M

Deadline: 3rd May

Instructions

1. Preferably use **Python3** for this Hack.
2. The final submission must consist of all your source code, data collected and **One PDF report** containing all your findings.
3. In case of any doubts/issues, please post on the Moodle forum.
4. Do not copy. This could result in an '**F**' grade in the course.

Bitcoin Alpha trust weighted signed network [30 Marks (5M * 6)]

This is who-trusts-whom network of people who trade using Bitcoin on a platform called [Bitcoin Alpha](#). Since Bitcoin users are anonymous, Pennsylvania Police need to maintain a record of users' reputation to prevent transactions with fraudulent and risky users. Members of Bitcoin Alpha rate other members in a scale of -10 (total distrust) to +10 (total trust) in steps of 1. For further insights and dataset refer to [link](#). For any definitions and formulas refer to [paper](#). Answer the following questions:

- 1) Calculate and plot user's fairness, goodness and reliability score.
- 2) How will you interpolate fairness, goodness and reliability score for each user to come up with a single value.
- 3) Find the set of 10 users who are highly trustable and non-trustable using the cumulative value obtained in above part.
- 4) Find the strongly connected component from the graph and state your observations on that.
- 5) Given a situation: Person A and Person B both has same negative score. Person A scored low negative score but from many users unlike Person B who has less degree compared to A but high negative score. Which person according to you can't be trusted.
- 6) Try to draw other similar inferences using plots.

Gowalla [20 Marks (5M * 4)]

Gowalla is a location-based social networking website where users share their locations by checking-in. The friendship network is undirected and was collected using their public API, and consists of 196,591 nodes and 950,327 edges. New York Cyber Police have collected a total of 6,442,890 check-ins of users over the period of Feb. 2009 - Oct. 2010. For further information and dataset refer to [link](#). Answer the following questions to help the police in predicting mobility of user which could help them in future crisis.

- 1) Plot the distribution of distance from home which user tends to travel.
Home for a user is the average of location where a user check-ins.
- 2) Plot the probability that a user will travel to friend's home vs the distance of check-in points of the user. Also explain the way you are calculating the probability.
- 3) Consider a user A and see whether user A's motion of trajectory at time t is affected by his/her friends check-in locations before the time t .
- 4) State the limits of using friendship data in predicting mobility of user.