Project Proposal: Improving a System

System: https://arxiv.org/pdf/1711.11458.pdf

1. What are the names and NetIDs of all your team members? Who is the captain? The captain will have more administrative duties than team members.

Captain: Philip Cori, pcori2
Team Member 1: Henry Moss, htmoss2
Team Member 2: Kyle Maxwell, kylem6

2. What system have you chosen? Are you adding a function or improving a function? What function?

We will improve on the recommender system proposed in *Collaborative Filtering with Social Exposure: A Modular Approach to Social Recommendation* by Menghan Wang, Xialin Zheng, and Yang Yang.

We will attempt to improve the system in 3 ways:

- Write a function to measure closeness between friends, as suggested by the paper.
- Write a function that performs matrix factorization for the the Rating Component
- Improve model based on recent social network analysis techniques (examples listed are social contagion and/or social structural influence). One possibility is to weigh items by an IDF term based on the items' number of occurrences in exposures between friends.

3. If you are adding a function, why is the new function important or interesting? How will it benefit the users? If you are improving a function, what are the main limitations of the current function? How are you going to improve it? How will your improvements benefit the users?

Implementing these new functions could improve the MAP, recall, and NCDG metrics. Firstly, a limitation of the current algorithm is that it assumes all friends are equally close. Therefore, if we can successfully measure "closeness" between friends, this can allow us to give more accurate recommendations. Secondly, the current algorithm does not use matrix factorization for the rating component. Implementing this can improve our metrics. Lastly, if we are able to use IDF weighting to filter out "main-stream" recommendations, our users will get more unique recommendations. This will make the recommendations that they receive from similar users more genuine and useful.

4. If you are adding a function, how will you demonstrate that it works as expected? If you are improving a function, how will you show your implementation actually works better?

The MAP, recall, and NDCG metrics will be used to determine if our experiments improved the recommender system.

5. How will your code communicate with or utilize the system?

We will directly extend the libraries of the framework and recommender system.

6. Which programming language do you plan to use?

We will use Python to implement our extensions.

7. Please justify that the workload of your topic is at least 20*N hours, N being the total number of students in your team. You may list the main tasks to be completed, and the estimated time cost for each task.

Implementing these functions will take considerable time for three reasons. Firstly, it will take time to fully understand the approach presented in the paper. Secondly, it will take time to understand the open source framework that implements the approach discussed. Lastly, it will take time to implement and experiment with the new functions.

Each person will do:

- 2h Team Meetings
- 1h Read Paper
- 1h Getting started with framework
- 2h Research into possible feature improvements
- 10-20h Implementation of feature improvement
- 2h Evaluation, Reporting