**COS 424 Project Proposal** 4/6/2012

*Members*: Stephen Beard, Joey Dodds, Soumyadeep Ghosh, Matt Zoufaly

*Project Name*: *Nightout*

**Problem**

Exploring where you are isn’t as easy as it sounds. Word of mouth combined with reviews from places like Google and Yelp help to find good places to eat and fun things to do. However, it would be nice to have a simpler approach: create an application that finds things for you to do based on your interests and tastes. What we’re going to create is an application that will have two features:

1. A rating system for activities in your area. Essentially, an activity can be anything ranging from a place to grab food, a place to go shopping, or a a parks. To put it simply, we’re going to combine a rating system from sites like Netflix with rating applications like Yelp.
2. A feature that allows multiple users to find strings of activities, food places, and businesses called ‘Outings’.

Fundamentally, people enjoy different things. So recommending activities based on what “most people” like (such as Yelp) does not necessarily lead searching users to activities they’ll enjoy. Specific interests / parameters can drastically change the recommendation. For this reason, we feel feature #1 is backbone of our application. The key motivation behind feature #2 is that activities are frequently done in sequence, and done with friends. Thus, we are aiming at creating strings of activities called “Outings”. However, in order to be effective, Outings cannot simply chain together businesses (e.g. businesses tens of miles apart cannot be in one Outing), and must take the location / method of transportation of the users into account. Finally, if the application’s feature is to be social, it has to take into account the ratings of several different users to find the best possible Outing for all users to communally take part in.

**Data**

Ideally, we would be able to populate our databases with real user data in real life. To this end, we’re creating a website that is slowly being populated. In order to evaluate our model, we can leverage existing ratings and information in internet databases (including Yelp, Google and FourSquare) to “seed” our databases with ratings. In order to do this, we’re going to create a web-crawler that will scrape businesses and user ratings from existing internet database. For each internet database, we will translate their rating system to ours (we intend to use a like / dislike approach) and insert the data as if we had collected it ourselves.[[1]](#footnote-1) This will be customized for each database, but since there are only a couple of possible rating systems (and most use a 5-star approach), this is fairly simple. We can evaluate on, essentially, an arbitrary number of users / businesses. To start, we’ll evaluate on several dozen users / businesses, refine the model, and then continually expand until we feel the results are indicative of a realistic workload (we estimate this will be around several thousand users and businesses).

**Methods**

**Evaluation**

1. Note, this is *only* for testing the validity of the model. This would not be legal in a live web application [↑](#footnote-ref-1)