WhatsHere
Geoffrey Hayes
Faculty Advisor: Jianbo Shi

Abstract:
- WhatsHere is a web application allowing users to find events in their community. The system uses a crawler to mine data from newspapers and websites across the Internet.

WhatsHere:
- We have designed a web application which allows users to view events within his local area. The user may view details and rate events as the system learns his preferences.

What's Here Application System:
- The application utilizes components from a diverse array of servers. deCarta provides professional web services for map and satellite imagery. We use Apache Tomcat and mySQL for application and database servers, running on an Amazon EC2 server.

Abstract:
- WhatsHere is a web application allowing users to find events in their community. The system uses a crawler to mine data from newspapers and websites across the Internet.

WhatsHere:
- We have designed a web application which allows users to view events within his local area. The user may view details and rate events as the system learns his preferences.

What's Here Application System:
- The application utilizes components from a diverse array of servers. deCarta provides professional web services for map and satellite imagery. We use Apache Tomcat and mySQL for application and database servers, running on an Amazon EC2 server.

Recommender System:
- Generating quality recommendations helps users discover and explore novel events, a keystone of our system. WhatsHere utilizes the user's past preference history along with aggregate analysis of a plethora of other users who have rated alternate events.

- We dynamically generate a Markov network (undirected Bayesian network) to determine the likelihood of a user enjoying a given event based on the past history of all users.

- We explore serendipity, choosing events using a biased random

Adaptive Context Curves:
- Collaborative recommending provides a framework but falls short when a source has limited rating information. We utilize other context information to provide crisper results.
- Proximity: is the event close enough to the user?
- Punctuality: is the event soon enough for the user? Is the event too long away for the user?
- Time of Day: does the user enjoy events at this time of day?

Experimental Setup:
- We synthesized a data set composed of 500 test users and 50 event sources. Each source was grouped into one of five categories and each user was given some preference toward each source.
- A Receiver Operating Curve (ROC) is used in signal detection theory as a measure of ability to detect signal from noise.
- A random detector will generate a ROC curve diagonal from origin to top-right while better detection bows the curve towards the top-left.

Experimental Results:
- ROC curves are shown for a single user, when each user in the system had rated 5, 10, 15 and 30 sources.
- In the 5-rates curve, the recommender system performed worse than random.
- By 30-rates, the system had very high accuracy.
- The fact that the system performed worse at 15-rates than 10-rates shows early instability.

Novel User
- A novel user was added to the 30-rates test and rated 5 sources (one from each category)
- The system had a 97% accuracy rating for that user

Conclusions:
- WhatsHere delivers an efficient web application for users looking for events in their communities.
- Our recommender system utilizes broad trends and local context to choose optimal events from a large search space.

Senior Project Poster Day 2008, CIS Dept. University of Pennsylvania