**Thesis Proposal for CSIS by Michael Mark**

**Title**

Vecsim optimization and Pathsim proposal

**Introduction**

To further optimize the Vector Similarity method proposed by A. Mukherjee, there are a variety of methods that we could use to do this.

1. Refine the random walks to become more realistic walking.
2. If we know the routes and directions of two different entities, we can use *Vecsim* to find their locations based on their speed and trajectory.
3. So, to do this, I propose that we create a more realistic method named *Pathsim*, based on pathing similarity between two different entities.

**Background**

Vecsim is a software which can be installed at the carrier. However, there are a few downsides to this method, including the fact that there may be DPD (discontininuity pair database) in the data logs received from the cell phone towers (or MRO records). This may cause Vecsim to require extensive data to be sent to it and extra time for the software to gather big data to figure out these DPD patterns.

Vecsim also takes into account of cell phone log data and normalizes everything. However, it does not account for the fact that each cell phone tower is different, with varying radii of reach for each unique tower, and much different sets of data. To normalize such data on a large scale may be counterintuitive.

**Proposal**

Thus, we propose to introduce Pathsim, a more sophisticated approach based on two entities’ movement to make a more accurate prediction on where they are based on the MRO log data as well.

Pathsim works the same as Vecsim in the fact that it is a software installed on the server side. Therefore, user privacy is still protected. It also is a bit more sophisticated in terms of its algorithm to find where the users are. It will conduct more realistic walking paths to figure out where the entities may be moving toward and use this data to find where the users are. Still, localization is not being used in this case as the user does not have to turn on Bluetooth or install any special software. However, Pathsim will be much more accurate in this case since it is able to retrieve more data regarding DPDs and figure out discontinuities in the MRO data better than Vecsim can.