Reviewer Name: Usha Kumar

e-mail: kuma0253@umn.edu

Paper title: T-Drive: Driving Directions based on Taxi trajectories

Feedback on the review report:

Overall this is a well written critical analysis report, accurately describing the problem, its significance to spatial DB, challenges faced in calculating the driving directions at run time. The reviewers also captured the main points like the contribution of this paper very well.

However, while explaining the key concepts introduced in this paper, reviewers chose to list the different terminologies used in the paper. In my opinion this should capture the main essential concepts which are introduced in this paper to solve the problem at hand. Some of the examples are:

- They propose a cloud based solution to aggregate and mine taxi trajectories
- Proposing a solution to build Time dependent landmark graph for intelligence modeling from taxi trajectory
- Travel time estimation using landmark graph construction algorithm.
- Handling temporal aspect of the problem using Variance entropy based clustering
- Proposing a VE-clustering algorithm
- Defining route computation based on landmark graph and VE clustering using a two step routing scheme.

In the assumptions sections, in addition to the assumptions mentioned here the other thing to consider are the variation in season, for example in winter a different route might yield a better results than what was applicable during summer months, hence averaging it over a 3 months period may not be adequence enough to take the seasonal variation into account. Also not sure if the estimates takes into account the processing time taken on the cloud in computing different landmark graphs, this could very well have been omitted by the authors of the paper, hence that can have a significant impact on the estimation result.

The validation methodology and suggested revisions are very well presented here. However, this can also state that a real trajectory data available in the cloud should also be used for evaluation methodology.

Time estimation related to landmark graph is not mentioned in the paper.

Feedback on the Presentation:
It would be nice if we could start the topic with a small introduction on the topic addressing the issues in calculating the driving distance at run time.

1. In slide 2, A better description of the main problem the authors are trying to solve would start the presentation nicely, instead of a problem statement. It would be nice if we could start the topic with a small introduction on the topic addressing the issues in calculating the driving distance at run time.

2. Slide 3 could be changed to highlight just the salient points which will enhance the presentation, for e.g
   - Trajectory data is both spatial and temporal in nature
   - This paper introduces a novel technique of using data mining to spatial and temporal applications.

3. slide 4 is perfectly fine.

4. Slide 5, Major contributions can be little less verbose like,
   - Landmark graph for trajectory preparation
     - Major landmarks are mined from the taxi trajectory data to build shortest path.
   - Variance-Entropy based clustering algorithm for travel time estimation.
   - Evaluation of the proposed methodologies using real world data sets as well as synthetic data sets.

5. Slide 6, Key concepts does not describe the key concepts presented in the paper. This can be replaced with the approaches slide which follows. As that explains the key concepts better. Hence can be removed.


7. Slide 9&10 – authors can also list In-the filed evaluation to present the three evaluation technique used in this paper in one slide, and can present it in a concise form.

8. In slide 10, strengths & weaknesses could be provided in a separate slide.

9. In slide for assumptions, it would be nice if it was less verbose.

10. Suggestions are good and is nice to have here.