Weekly report: May 22-May 29, 2015

Ivan Rios S.

May 29, 2015

1 Goals for the week

- Improve the results obtained by inverting the steps of the trajectory detection algorithm
- Look through literature related to trajectory detection
- Replicate work done by Tommy related to trajectory detection

2 Activities

- Fix problem of trips not showing up for professor Golab's bike.
- Modify the trip detection algorithm to improve the accuracy after swapping steps.
- Analyze the results obtained and compare them with the original algorithm.
- Replicate Tommy's work with the new data.
- Revise the code used in the webpage.

3 What I learnt

There are problems in the code used for the webpage:

- Changes not properly documented.

- No comments on the code to make it easy to read for other people working on it.

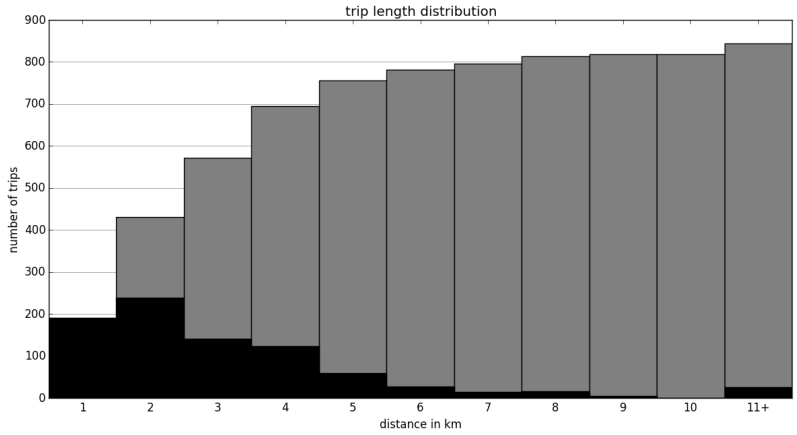
- There are possible points of failure that could cause problems later in the project. I think we need a report from the URA who worked directly on the webpage. I emailed Ran to ask him if he had a summary of his work but I haven't heard back from him; Mikhail has helped me to understand but neither of us is completely familiar with the changes done on the web page code.

- Script that populates the table with information about the trips does not keep track of errors generated during this transmission of data which makes it

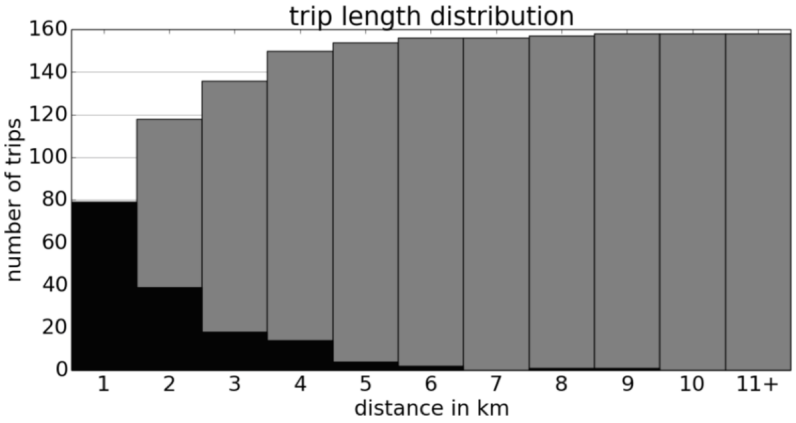
hard to find the root cause of failures. We need to improve this script because the only way to solve the problems right now is to clear those tables and rerun the scripts manually.

- Trips calculated are missing a considerable part of the trajectory (we are reporting shorter distances than the real ones). I have gone through the code changes done my Mikhail and they seem to be properly implemented; we will need to dig in deeper into this problem.

- Using the new data generated I got the following results:
Trip distribution:

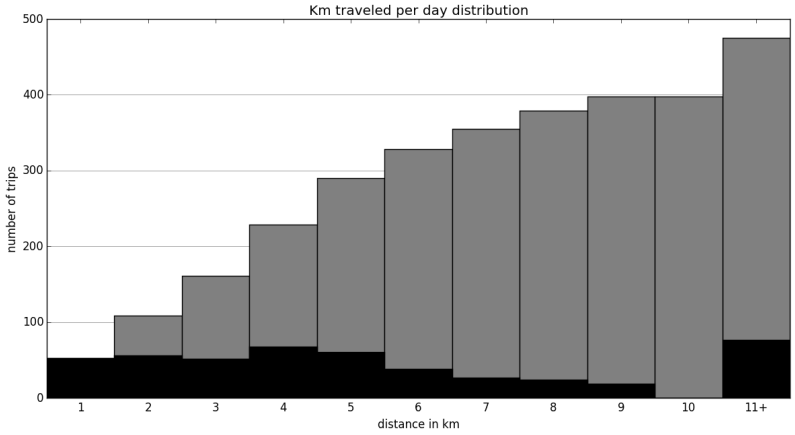


Which compared to Tommy's:



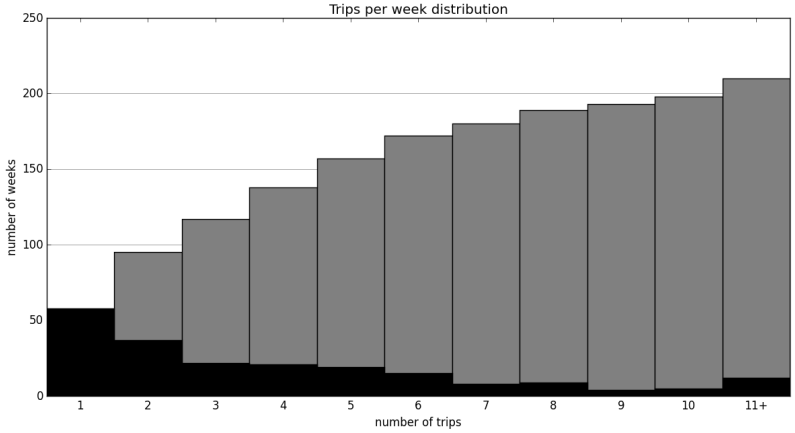
We can see that the distribution has changed, the number of kms people bike were considerably lower at the beginning of the project. This could be caused by the interest people had in the new bikes when they first got them; also, as they know more about the capacity of the battery and become more comfortable using the bike users might be taking longer trips.

Also, I obtained the distribution of kms traveled per day:



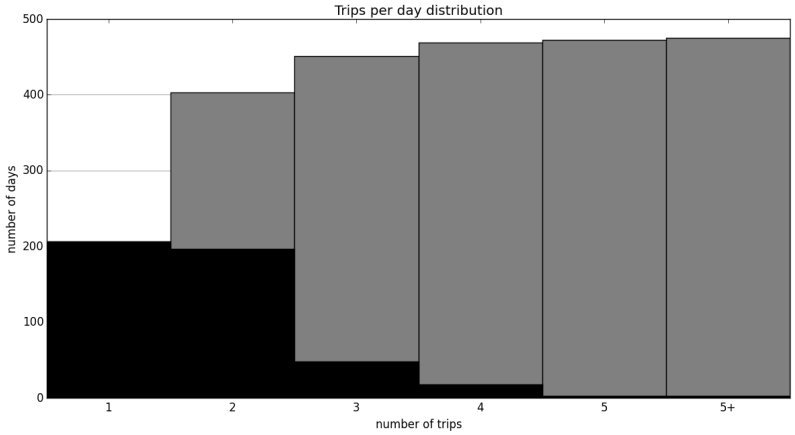
Which shows that bike users are biking generally between 1 and 5 kms per day. However, this graph also shows that trips might now be detected correctly since it would be expected that a bike owner would travel from their origin to a destination and back (e.g. to work and back home); but if we compare the two graphs (trip length distribution and kms traveled per day distribution) we can see that numbers do not match.

Also, I obtained the number of trips per week:



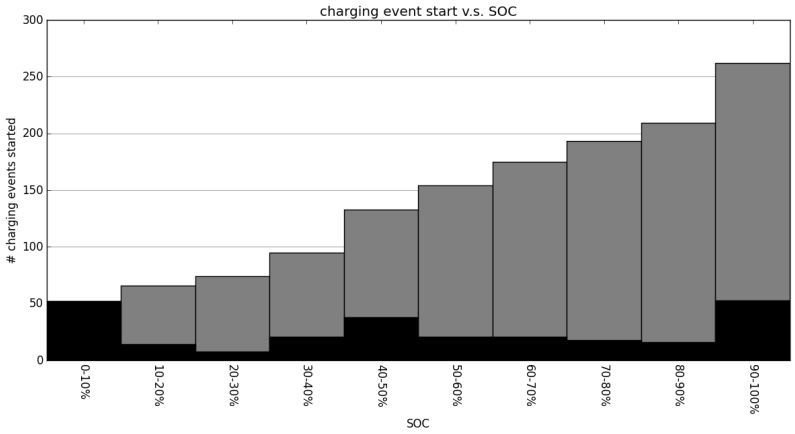
Which shows that most of the weeks people have 1 trip; this suggests that probably the algorithm is not identifying all the trips correctly since generally bikes are used for a 2 ways trip. However, this could also mean that users are taking their bikes for exercise which is reflected as having mostly one trip per week since a regular workout would be considered as one trip.

Similarly, we have the number of trips per day:

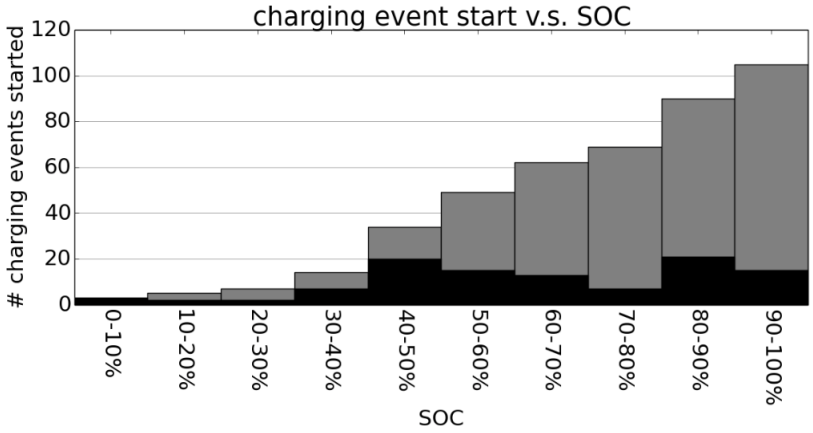


Which leads us to similar conclusions as before. Here, we can see that generally the bikes are being used for either 1 or 2 trips per day. Again, this could mean that bikes have become a way for people to exercise and also, in a similar amount, they have become a mean of transportation for specific purposes (probably to go to work).

Additionally, I obtained the SOC at which people started charging the batteries of the bike:



Which compared to the original graph obtained by Tommy:



We can see that the distribution has changed since people are now charging their batteries when the battery SOC is lower than before. Again, this could be cause because the bike users are more confident using the bike (since they know its capabilities better now) and also because their trips are longer which more more use of the battery charge. Previously, users would charge their battery mostly when the SOC was ˜40%, on the other hand, now they are charging it when the SOC is lower than 10%. This shows a change in the mentality of the users and the confidence that has been building up after using the bicycle for a considerable amount of time.

- Also, this week I read one paper related to trip detection:

Paper: Chung, Eh, and A. Shalaby. "A Trip Reconstruction Tool for GPS-Based Personal Travel Surveys." Transportation Planning and Technology 28.5 (2005): 381-401. Print.

Summary:

The paper describes a methodology and tool developed to identify the modes and trajectories followed by a group of people in the city of Toronto using the GPS traces they generated. This tool uses a GIS based map-matching algorithm which uses information of the different routes available around the city in order to determine the most likely mode (walk, bike, bur or car) and estimate the route used. In order to define the routes followed by the users, the algorithm process the GPS data in 3 parts: first, it pre-process the data by identifying low quality data with HDOP and the number of satellites available when each record was taken; also, they determine if the trip has finished by checking if the variation of latitude or longitude are lower than 0.000005^{o}. Secondly, the algorithm estimates 'links' which is points in the map that could actually be used as a potential route (also called map-matching). Finally, the post-processing stage defines the correct trajectory followed by re-arranging the points in a way that the trajectory is one directional and does have sudden turns potentially caused by noisy GPS data.

4 Proposed goals for next week

- Analyze problems with the trip detection algorithm.
- Make changes to improve the results obtained and compare them with the previous version (Dec 2014) and Tommy's version.
- Go through more literature dedicated to trip detection.