Feedback Seminar II

I am aware not all sections are compelte, the feedback for the none complete sections are there just to make it easier for the readerr who do not have any prior knowledge.

Some general notes: - Tables have no name and caption. You mention to see Table 1. I assumed it was the first table.

Introduction

You describe MAD as a new tool. However, you you do not mention any tools/approaches used before. Maybe a short background part to give the reader more context of how it was executed before andt today?

First paragraph:

It is a good flow and introduces the concepts, and uses good exaples.

Second paragraph:

You mention that the data is retrived from phone users. However, you do not mention anything else regarding the etich until chapter 3.1.

I had to jump in the text to read the ethics part before continue reading the introduction. Is it possible to mention that all user-data was collected with consent?

The OD matrix is introduces. How does it work? how does it look? A short descriprion and a figure to make it easier for the reader to understand.

Third paragraph:

Good general descrription of the study. Once again the data is mentioned, but no mention of the consent part.

Related Works

Understanding mobility using big geolocation data

Good description of what has beed done before. However, you still mentioned the OD matrix with do explonation of how it works.

You mention different angles/perspectives regarding the area, I like it a lot.

One sentence which were outside the flow: "The above mentioned papers", maybe change it to "The papers mentioned above" for a nicer flow?

What is the four step model?

Activity plans from big geolocation data

You begini with a connection back to the prior text. Really nice.

The second paragraph does contain a lot of information. Maybe a small table to helt summarize the models? Maybe a bit more explonation regarding the models?

Routine activities at home and work Developed by a study? Use the Authors name to give more clarity who the autor are.

Synthesizing activity plans from incomplete data

CDR? You introduce a new acronyme without any exploation - Google search: Call Detail Record?

Research gaps

A new general technique, but what model? Describe the model, a gigure would be nice as well

Thesis objectives

You mentioned two points in the introduction, here is it three points? need to stay consistent.

Otherwise, I do like this part, explaining with a usecase and potential stakeholders.

Ethical considerations

Good description regarding the etichs. However, since you mention user collected data earlier in the interoduction, it would be nice to have it in the same place to reduce jumping arount.

Methods

Good short description, short and informative what to expect in the chapter.

Data sources

I would keep it as it is, reference to the data tables would be a nice addition.

Mobile Application Data (MAD)

The text were well written and described all the points with enough information, there are som minor thoughts: - In the MAD processing, you mention *"sevetal data processin steps were implemented"*, which steps? - Table missing Name,

Caption. - You mention that you will not inclde weekends and holidays. Weekends are simple to exclude, but how does it work with the holiday? What id someone have a holiday in none vacation season?

Travel Survey

When you reference the table, maybe: "summarized in the table `\ref{table}`

Table is missing caption and name

Home and work detection

You refer to a model that you use, what is the model? can you describe it with a figure?

Maybe rephrase the first paragraph:

> We employ the model proposed in to estimate our MAD-users' home and work locations. The paper focuses on determining home locations, and we utilize the most frequented area throughout the night for each individual as the projected home location, and we estimate using the same time frame. To estimate work location, we first eliminate each user's predicted home location. Because we are interested in travel patterns, having both home and work in the same region might confound our results. Then we look at the opposite hours that we looked at when we looked at the user's home location, minus some time for the user to commute. As a result, the estimated work location is the location that was visited the most from 09-17 and is not the home location.

- Usually dont begin a sentence with "So", it is more "every day English"
- Last sentence in the first paragraph: "So the estimated. . . "
- X imes and X minutes. This means that they are the same, are they always the same? If not, change to X and Y or some other variable name.
- The paragraph under the placeholder for the picture, does not really connect to tyhe third paragraph.

Matching home location with DeSO-code

Describe waht DeSo-code and DeSO-area, for a reader with no prior knowledge, there is a bit of confusion. It was hard to understand the subchapter.

Polpulation debiasing based on estimated home location

I did not read this part. . .

Transforming the data for Jaccard similarity comparison

You describe how the data is used, workflow, and example. However, I have no idea what the Jaccard is. Jaccard needds to be described what it is and how it works.

Overall, it is very good written, you cover the solution part. However, to fully grasp the solution part, some descriptions regarding the Jaccard is needed.

Running the Jaccard similarity comparison with added probability information

Good flow in the text and explain waht you are doing. Still missing some theoretical parts.

This sentence was not complete: *"The comparison only looks at MAD-users and survey participants living in the same county, since"* ?

Summary

Overall is the text good, and have a flow. There are some minor parts that are more "every data" English.

The text needs some more explanations regarding the theoretical part. I had to google "Jaccard" and "OD matrix" do understand how it worked (or it could be me who missed something?).

The title page would be nice if you included it, it helps to narrow down the scope and get an idea of what you are writing about. The same also goes for the reference list. It would be nice to be able to see the sources, i.e. in chapter 4.2 you mention that you are using the model from the source `\cite{pappalardo2021evaluation}`, it is har do understand how the model is working with no way of looking it up.

You state the aim of the thesis with 2 points in the intrroduction. Howvever, in chapter 3. Thesis Objective, you state 3 points. Which one is the correct one?

It would be nice with figures as a complement to the text to make it more understandable, i.e. the models you are using, Jaccard, OD matrix.